

# Refined Instrumental Variable Methods for Identification of LPV Box-Jenkins Models

Vincent Laurain<sup>a</sup>, Marion Gilson<sup>a</sup>, Roland Tóth<sup>b</sup>, Hugues Garnier<sup>a</sup>

<sup>a</sup>Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, BP 70239, 54506 Vandoeuvre-les-Nancy Cedex, France.

<sup>b</sup>Delft Center for Systems and Control (DCSC), Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.

---

## Abstract

Identification of linear parameter-varying systems in an input-output setting is investigated, focusing on the case when the noise part of the data generating system is an additive colored noise. In the Box-Jenkins and output-error cases, it is shown that the currently available linear regression and instrumental variable methods from the literature are far from being optimal in terms of bias and variance of the estimates. To overcome the underlying problems, a refined instrumental variable method is introduced. The proposed approach is compared to the existing methods via a representative simulation example.

*Key words:* LPV models, system identification, Refined Instrumental Variable, Box-Jenkins models, Input/Output, transfer function

---

## 1 Introduction

The common need for accurate and efficient control of today's industrial applications is driving the system identification field to face the constant challenge of providing better models of physical phenomena. Systems encountered in practice are often nonlinear or have time-varying nature. Dealing with models of such kind without any structure is often found infeasible in practice. This rises the need for system descriptions that form an intermediate step between *Linear Time-Invariant* (LTI) systems and nonlinear/time-varying plants. To cope with these expectations, the model class of *Linear Parameter-Varying* (LPV) systems provides an attractive candidate. In LPV systems the signal relations are considered to be linear just as in the LTI case, but the parameters are assumed to be functions of a measurable time-varying signal, the so-called *scheduling variable*  $p : \mathbb{Z} \mapsto \mathbb{P}$ . The compact set  $\mathbb{P} \subseteq \mathbb{R}^{n_p}$  denotes the *scheduling space*. The LPV system class has a wide representation capability of physical processes and this framework is also supported by a well worked out and industrially reputed control theory. Despite the advances of the LPV control field, identification of such systems is not well developed.

---

*Email addresses:* [vincent.laurain@cran.uhp-nancy.fr](mailto:vincent.laurain@cran.uhp-nancy.fr) (Vincent Laurain), [marion.gilson@cran.uhp-nancy.fr](mailto:marion.gilson@cran.uhp-nancy.fr) (Marion Gilson), [r.toth@tudelft.nl](mailto:r.toth@tudelft.nl) (Roland Tóth), [hugues.garnier@cran.uhp-nancy.fr](mailto:hugues.garnier@cran.uhp-nancy.fr) (Hugues Garnier).

The existing LPV identification approaches are almost exclusively formulated in discrete-time, commonly assuming static dependence on the scheduling parameter (dependence only on the instantaneous value of the scheduling variable), and they are mainly characterized by the type of LPV model structure used: *Input-Output* (IO) [2,3,21] *State-Space* (SS) [14,20,5] or *Orthogonal Basis Functions* (OBFs) based models [19] (see [18] for an overview of existing methods). In the field of system identification, IO models are widely used as the stochastic meaning of estimation is much better understood for such models, for example via the *Prediction-Error* (PE) setting, than for other model structures. Often an important advantage of IO models is that they can be directly derived from physic/chemical laws in their continuous form. Therefore, it is more natural to express a given physical system through an IO operator form or transfer function modeling. A comparison between IO and SS model based approaches can be found in [17] for linear systems.

Among the available identification approaches of IO models, the interest for *Instrumental Variable* (IV) methods has been growing in the last years. The main reason of this increasing interest is that IV methods offer similar performance as extended *Least Square* (LS) methods or other *Prediction Error Minimization* (PEM) methods (see [15,13]) and provide consistent results even for an imperfect noise structure which is the case in most practical applications. These approaches have

been used in many different frameworks such as direct continuous-time [15,10], nonlinear [11] or closed-loop identification [9,8] and lead to optimal estimates in the LTI linear case if the system belongs to the model set defined.

In the LPV case, most of the methods developed for IO models based identification are derived under a linear regression form [21,2,1]. By using the concepts of the LTI PE framework, recursive LS and IV methods have been also introduced [7,4]. However, it has been only recently understood how a PE framework can be established for the estimation of general LPV models [18]. Due to the linear regressor based estimation, the usual model structure in existing methods is assumed to be *auto regressive with exogenous input* (ARX). Even if a non-statistically optimal IV method has been recently introduced in [4] for LPV *Output Error* (OE) models, no method has been proposed so far to deal with colored noise. Moreover, it can be shown that it is generally impossible to reach statistically optimal estimates by using linear regression as presented so far in the literature. These imply, that there is lack of an LPV identification approach, which is capable of efficiently estimating LPV-IO models under colored noise conditions, *e.g.* as in a *Box-Jenkins* (BJ) setting, which is the case in many practical applications.

By aiming at fulfilling this gap, an estimation method is developed in this paper for LPV-IO BJ discrete-time models in the SISO case. The properties of the method are compared to the existing theory showing the increased statistical performance of the estimation.

The paper is organized as follows: in Section 2, the general class of LPV systems in an IO representation form is introduced pointing out the main difficulties presented. In Section 3, linear regression in the LPV prediction error framework is analyzed and it is shown that such an estimation scheme even in a IV setting is statistically not optimal if the noise is not white. Moreover, a reformulation of the dynamical description of LPV data generating plants in the considered setting is introduced which makes possible the extension of LTI-IV methods to the LPV framework. In Section 4, LPV-IV methods are introduced and analyzed, while their performance increase compared to other methods is shown in Section 5. Finally in Section 6, the main conclusions of the paper are drawn and directions of future research are indicated.

## 2 Problem description

### 2.1 System description

Consider the data generating LPV system described by the following equations:

$$\mathcal{S}_o \begin{cases} A_o(p_k, q^{-1})\chi_o(t_k) = B_o(p_k, q^{-1})u(t_k-d) \\ y(t_k) = \chi_o(t_k) + v_o(t_k) \end{cases} \quad (1)$$

where  $p_k$  is the value of the scheduling parameter  $p$  at sample time  $t_k$ ,  $d$  is the delay,  $\chi_o$  is the noise-free out-

put,  $u$  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.*  $q^{-i}u(t_k) = u(t_{k-i})$ .  $A_o(p_k, q^{-1})$  and  $B_o(p_k, q^{-1})$  are polynomials in  $q^{-1}$  of degree  $n_a$  and  $n_b$  respectively:

$$A_o(p_k, q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i^o(p_k)q^{-i}, \quad (2a)$$

$$B_o(p_k, q^{-1}) = \sum_{j=0}^{n_b} b_j^o(p_k)q^{-j}, \quad (2b)$$

where the coefficients  $a_i$  and  $b_j$  are real meromorphic functions ( $f : \mathbb{R}^n \mapsto \mathbb{R}$  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 non-singular on  $\mathbb{P}$ , thus the solutions of  $\mathcal{S}_o$  are well-defined and the process part is completely characterized by the coefficient functions  $\{a_i^o\}_{i=1}^{n_a}$  and  $\{b_j^o\}_{j=0}^{n_b}$ .

Most of existing methods in the literature assume an ARX type of data generating system, which means that the noise process  $v_o$  can be written as

$$e_o(t_k) = A_o(p_k, q^{-1})v_o(t_k), \quad (3)$$

where  $e_o$  is a zero-mean, discrete-time white noise process with a normal distribution  $\mathcal{N}(0, \sigma_o^2)$ , where  $\sigma_o^2$  is the variance. Even if in some specific applications, the dependence of the noise on  $p$  can be considered as a fair assumption, the structure of (3) is often found unrealistic as it assumes that both the noise and the process part of  $\mathcal{S}_o$  contain the same dynamics. In this paper, a more general case is considered where the colored noise associated with the sampled output measurement  $y(t_k)$  is assumed to have a rational spectral density which might have no relation to the actual process dynamics of  $\mathcal{S}_o$ . As a preliminary step towards the case of a  $p$ -dependent noise, it is also assumed that this rational spectral density is not dependent on  $p$ : this corresponds to a more realistic assumption than (3), especially in case of measurement noise. Therefore,  $v_o$  is represented by a discrete-time *autoregressive moving average* (ARMA) model:

$$v_o(t_k) = H_o(q)e_o(t_k) = \frac{C_o(q^{-1})}{D_o(q^{-1})}e_o(t_k), \quad (4)$$

where  $C_o(q^{-1})$  and  $D_o(q^{-1})$  are monic polynomials with constant coefficients and with respective degree  $n_c$  and  $n_d$ . Furthermore, all roots of  $z^{n_d}D_o(z^{-1})$  are inside the unit disc. It can be noticed that in case  $C_o(q^{-1}) = D_o(q^{-1}) = 1$ , (4) defines an OE noise model.

### 2.2 Model considered

Next we introduce a discrete-time LPV Box-Jenkins (BJ) type of model structure that we propose for the identification of the data-generating system (1) with noise model (4). In the proposed model structure, the noise model and the process model are parameterized separately.

### 2.2.1 Process model

The process model is denoted by  $\mathcal{G}_\rho$  and defined in a form of an LPV-IO representation:

$$\mathcal{G}_\rho : (A(p_k, q^{-1}, \rho), B(p_k, q^{-1}, \rho)) = (\mathcal{A}_\rho, \mathcal{B}_\rho) \quad (5)$$

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

$$\mathcal{A}_\rho \begin{cases} A(p_k, q^{-1}, \rho) = 1 + \sum_{i=1}^{n_a} a_i(p_k)q^{-i}, \\ a_i(p_k) = a_{i,0} + \sum_{l=1}^{n_\alpha} a_{i,l}f_l(p_k) \quad i = 1, \dots, n_a \end{cases}$$

$$\mathcal{B}_\rho \begin{cases} B(p_k, q^{-1}, \rho) = \sum_{j=0}^{n_b} b_j(p_k)q^{-j}, \\ b_j(p_k) = b_{j,0} + \sum_{l=1}^{n_\beta} b_{j,l}g_l(p_k) \quad j = 0, \dots, n_b \end{cases}$$

In this parametrization,  $\{f_l\}_{l=1}^{n_\alpha}$  and  $\{g_l\}_{l=1}^{n_\beta}$  are meromorphic functions of  $p$ , with static dependence, allowing the identifiability of the model (pairwise orthogonal functions on  $\mathbb{P}$  for example). The associated model parameters  $\rho$  are stacked columnwise:

$$\rho = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_{n_a} & \mathbf{b}_0 & \dots & \mathbf{b}_{n_b} \end{bmatrix}^\top \in \mathbb{R}^{n_\rho}, \quad (7)$$

where

$$\mathbf{a}_i = \begin{bmatrix} a_{i,0} & a_{i,1} & \dots & a_{i,n_\alpha} \end{bmatrix} \in \mathbb{R}^{n_\alpha+1}$$

$$\mathbf{b}_j = \begin{bmatrix} b_{j,0} & b_{j,1} & \dots & b_{j,n_\beta} \end{bmatrix} \in \mathbb{R}^{n_\beta+1}$$

and  $n_\rho = n_a(n_\alpha + 1) + (n_b + 1)(n_\beta + 1)$ . Introduce also  $\mathcal{G} = \{\mathcal{G}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$ , as the collection of all process models in the form of (5).

### 2.2.2 Noise model

The noise model is denoted by  $\mathcal{H}$  and defined as an LTI transfer function:

$$\mathcal{H}_\eta : (H(q, \eta)) \quad (8)$$

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

$$H(q, \eta) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} = \frac{1 + c_1q^{-1} + \dots + c_{n_c}q^{-n_c}}{1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}}. \quad (9)$$

The associated model parameters  $\eta$  are stacked columnwise in the parameter vector,

$$\eta = \begin{bmatrix} c_1 & \dots & c_{n_c} & d_1 & \dots & d_{n_d} \end{bmatrix}^\top \in \mathbb{R}^{n_\eta}, \quad (10)$$

where  $n_\eta = n_c + n_d$ . Additionally, denote  $\mathcal{H} = \{\mathcal{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$ , the collection of all noise models in the form of (8).

### 2.2.3 Whole model

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

$$\mathcal{M}_\theta \begin{cases} A(p_k, q^{-1}, \rho)\chi(t_k) = B(p_k, q^{-1}, \rho)u(t_{k-d}) \\ v(t_k) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)}e(t_k) \\ y(t_k) = \chi(t_k) + v(t_k) \end{cases} \quad (11)$$

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

$$\mathcal{M} = \{(\mathcal{G}_\rho, \mathcal{H}_\eta) \mid \text{col}(\rho, \eta) = \theta \in \mathbb{R}^{n_\rho+n_\eta}\}. \quad (12)$$

This set corresponds to the set of candidate models in which we seek the model that explains data gathered from  $\mathcal{S}_o$ , the best, under a given identification criterion (cost function).

### 2.3 Predictors and prediction error

Similar to the LTI case, in the LPV prediction error framework, one is concerned about finding a model in a given LPV model structure  $\mathcal{M}$ , which minimizes the statistical mean of the squared prediction error based on past samples of  $(y, u, p)$ . However in the LPV case, no transfer function representation of systems is available. Furthermore, multiplication with  $q$  is not commutative over the  $p$ -dependent coefficients, meaning that  $q^{-1}B(p_k, q^{-1})u(t_k) = B(p_{k-1}, q^{-1})u(t_{k-1})$  which is not equal to  $B(p_k, q^{-1})u(t_{k-1})$ . Therefore to define predictors with respect to models  $\mathcal{M}_\theta \in \mathcal{M}$ , a convolution type representation of the system dynamics, *i.e.* an LPV *Impulse Response Representation* (IRR), is used where the coefficients have dynamic dependence on  $p$  (dependence on future and past samples of  $p$ ) [18]. This means that  $\mathcal{S}_0$  with an asymptotically stable process and noise part is written as

$$y(t_k) = \underbrace{(G_o(q) \diamond p)(t_k)u(t_k)}_{\chi_o(t_k)} + \underbrace{(H_o(q) \diamond p)(t_k)e_o(t_k)}_{v_o(t_k)} \quad (13)$$

where

$$(G_o(q) \diamond p)(t_k) = \sum_{i=0}^{\infty} (\alpha_i^o \diamond p)(t_k)q^{-i}, \quad (14a)$$

$$(H_o(q) \diamond p)(t_k) = 1 + \sum_{i=1}^{\infty} (\beta_i^o \diamond p)(t_k)q^{-i}, \quad (14b)$$

with  $\alpha_i^o \diamond p$  expressing dynamic dependence of  $\alpha_i$  on  $p$ , *i.e.*  $\alpha_i^o \diamond p = \alpha_i(p, qp, q^{-1}p, q^2p, \dots)$ . Now if  $p$  is deterministic and there exists a convergent adjoint  $H_o^\dagger$  of  $H_o$  such that

$$e_o(t_k) = (H_o^\dagger(q) \diamond p)(t_k)v_o(t_k), \quad (15)$$

then it is possible to show (see [18]) that the *one-step ahead predictor* of  $y$  is

$$y(t_k | t_{k-1}) = ((H_o^\dagger(q)G_o(q) \diamond p)(t_k) u(t_k) + ((1 - H_o^\dagger(q)) \diamond p)(t_k) y(t_k)). \quad (16)$$

In case the noise model is not dependent on  $p$ , like in (4),  $(H_o(q) \diamond p)(t_k) = \frac{C_o(q^{-1})}{D_o(q^{-1})}$  and  $(H_o^\dagger(q) \diamond p)(t_k) = \frac{D_o(q^{-1})}{C_o(q^{-1})}$ . With respect to a parameterized model structure, we can define the *one-step ahead prediction error* as

$$\varepsilon_\theta(t_k) = y(t_k) - \hat{y}(t_k | t_{k-1}), \quad (17)$$

where

$$\hat{y}(t_k | t_{k-1}) = ((H^\dagger(q, \theta)G(q, \theta) \diamond p)(t_k) u(t_k) + ((1 - H^\dagger(q, \theta)) \diamond p)(t_k) y(t_k)) \quad (18)$$

with  $G(q, \theta)$  and  $H(q, \theta)$  the IRR's of the process and noise part respectively. Denote  $\mathcal{D}_N = \{y(t_k), u(t_k), p(t_k)\}_{k=1}^N$  a data sequence of  $\mathcal{S}_o$ . Then to provide an estimate of  $\theta$  based on the minimization of  $\varepsilon_\theta$ , an identification criterion  $W(\mathcal{D}_N, \theta)$  can be introduced, like the *least squares* criterion

$$W(\mathcal{D}_N, \theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon_\theta^2(t_k), \quad (19)$$

such that the parameter estimate is

$$\hat{\theta}_N = \arg \min_{\theta \in \mathbb{R}^{n_\rho + n_\eta}} W(\mathcal{D}_N, \theta). \quad (20)$$

#### 2.4 Persistency of excitation

In order to estimate an adequate model in a given model set, most PEM algorithms like least squares or instrumental variable methods require that a *persistency of excitation* condition with respect to  $\mathcal{D}_N$  collected from the system is satisfied. Such condition is required to guarantee consistency and convergence of the algorithm providing estimates. However, in the LPV case it turns out that persistency of excitation in terms of  $(u, p)$  for a given order, as it is understood in the LTI case (see [12]), does not guarantee consistency of the estimated parameters. The reason is that even if identifiability of the given parametrization is satisfied under the considered identification criterion, statistically global minimum of the criterion function is not guaranteed with respect to such data. This means that the terminology of persistency of excitation with order  $n$  is ill-defined in the LPV case. Instead, the informativity of the data sets (see [6]) with respect to the assumed coefficient parametrization and model order is needed to be satisfied in order to ensure consistency and convergence of the estimation. However, conditions of informative data sets have not been investigated directly in the LPV literature. For some preliminary work with conservative conditions see [3,21]. The

question whether a data set is informative in the LPV case remains open. In terms of the upcoming analysis, it is assumed that the considered data sets satisfy this property. However in practice, the absence of a solid criterion restricts the user to the paradigm to excite as much as possible the system in order to guarantee consistency and convergence of the estimation.

#### 2.5 Identification problem statement

Based on the previous considerations, the identification problem addressed in the sequel can now be defined.

**Problem 1** *Given a discrete time LPV data generating system  $\mathcal{S}_o$  defined as in (1) and a data set  $\mathcal{D}_N$  collected from  $\mathcal{S}_o$ . Based on the LPV-BJ model structure  $\mathcal{M}_\theta$  defined by (11), estimate the parameter vector  $\theta$  using  $\mathcal{D}_N$  under the following assumptions:*

- A1**  $\mathcal{S}_o \in \mathcal{M}$ , i.e. there exists a  $\mathcal{G}_o \in \mathcal{G}$  and a  $\mathcal{H}_o \in \mathcal{H}$  such that  $(\mathcal{G}_o, \mathcal{H}_o)$  is equal to  $\mathcal{S}_o$ .
- A2** In the parametrization  $\mathcal{A}_\rho$  and  $\mathcal{B}_\rho$ ,  $\{f_l\}_{l=1}^{n_\alpha}$  and  $\{g_l\}_{l=1}^{n_\beta}$  are chosen such that  $(\mathcal{G}_o, \mathcal{H}_o)$  is identifiable for any trajectory of  $p$ .
- A3**  $u(t_k)$  is not correlated to  $e_o(t_k)$ .
- A4**  $\mathcal{D}_N$  is informative with respect to  $\mathcal{M}$ .
- A5**  $\mathcal{S}_o$  is globally BIBO stable, i.e. for any trajectory of  $p : \mathbb{R} \mapsto \mathbb{P}$  and any bounded input signal  $u$ , the output of  $\mathcal{S}_o$  is bounded [18].

### 3 On the use of linear regression framework and statistical optimality

LPV-IO parametric identification methods proposed in the literature so far are based on LS methods such as least squares or instrumental variables [3,4]. The currently accepted view in the literature is that if the system belongs to the model set defined in (12), then  $y(t_k)$  can be written in the linear regression form:

$$y(t_k) = \varphi^\top(t_k) \rho + \tilde{v}(t_k) \quad (21)$$

with  $\rho$  as defined in (7) and

$$\varphi(t_k) = \begin{bmatrix} -y(t_{k-1}) & \dots & -y(t_{k-n_\alpha}) & -y(t_{k-1})f_1(p_k) & \dots \\ & & -y(t_{k-n_\alpha})f_{n_\alpha}(p_k) & u(t_{k-d}) & \dots & u(t_{k-n_b-d}) \\ & & & u(t_{k-d})g_1(p_k) & \dots & u(t_{k-n_b-d})g_{n_\beta}(p_k) \end{bmatrix}^\top \quad (22a)$$

$$\tilde{v}(t_k) = A(p_k, q^{-1}, \rho)v(t_k). \quad (22b)$$

In this section it is shown why such a linear regression cannot lead to statistically optimal (unbiased and minimal variance) estimates when the model structure is an LPV Box-Jenkins. Let us first introduce the adjoint  $A^\dagger$  of  $A$ , such that  $\chi = A^\dagger(p_k, q^{-1}, \rho)u \Leftrightarrow A(p_k, q^{-1}, \rho)\chi = u$ . Note that the adjoint always exists in a IRR sense with respect to an asymptotically stable  $A$ . In the LTI case,  $A^\dagger = \frac{1}{A}$ , however, in the LPV case,  $A^\dagger \neq \frac{1}{A}$  due to the non-commutativity of the multiplication by  $q$ .

### 3.1 The conclusion brought in [4]

By considering (21) and the associated extended regressor in (22a), it is well known that the LS method leads to optimal estimate only if the noise model is ARX ( $\tilde{v}(t_k)$  is a white noise). This condition implies that  $v(t_k) = A^\dagger(p_k, q^{-1}, \rho)e(t_k)$  and is not fulfilled in many practical situations as  $v_o$  is often not related directly to the process itself and might not depend on  $p_k$ . Therefore it is proposed in [4] to use an IV method where the instrument is built using the simulated data generated from an estimated auxiliary ARX model:

#### Algorithm 1 (One-step IV method)

**Step 1** Estimate an ARX model by the LS method (minimizing (19)) using the extended regressor (22a).

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

In general, instrumental variable methods have the particularity to produce unbiased estimates if the instrument is not correlated to the measurement noise. Based on the numerical simulation given in [4], the following conclusions have been proposed:

- In case  $\mathcal{S}_o$  corresponds to an LPV-OE model ( $v_o = e_o$ ), Algorithm 1 leads to an unbiased estimate.
- The variance of the estimated parameters is much larger than in a LS estimation as it is well-known.
- The estimation result can be improved if one uses a multi-step algorithm such as in [12].

### 3.2 Existing methods and optimal estimates

In the present paper, the authors only partially agree with the conclusions stated in [4]. It is true that the results can be improved and that the IV estimates are unbiased but this paper claims that:

- Even by using multi-step algorithm of [12], the optimal estimate cannot be reached with the linear regression form (21).
- For LPV-BJ models, estimates that are close to the statistically optimal solution can be reached by using IV methods and the variance of the estimated parameters is close to variance of the LS estimator in given situations.

In the following part it is shown why these statements hold true. In order to show why statistically optimal estimation of the model (11) cannot be reached under the viewpoint (21), it is necessary to revisit the result of optimal prediction error in the LTI case.

#### 3.2.1 The LTI case

In analogy with (11), consider the LTI-BJ model as

$$\mathcal{M}_\theta^{\text{LTI}} \begin{cases} A(q^{-1}, \rho)\chi(t_k) = B(q^{-1}, \rho)u(t_{k-d}) \\ v(t_k) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)}e(t_k), \\ y(t_k) = \chi(t_k) + v(t_k). \end{cases} \quad (23)$$

where  $A(q^{-1}, \rho)$  and  $B(q^{-1}, \rho)$  are polynomials in  $q^{-1}$  with constant real coefficients and have degree  $n_a$  and  $n_b$  respectively and  $e$  is a white noise with  $e(t_k) \in \mathcal{N}(0, \sigma^2)$ .  $y(t_k)$  can be written in the linear regression form:

$$y(t_k) = \varphi^\top(t_k)\rho + \tilde{v}(t_k), \quad (24)$$

with

$$\rho = \begin{bmatrix} a_1 & \dots & a_{n_a} & b_0 & \dots & b_{n_b} \end{bmatrix}^\top \in \mathbb{R}^{n_a+n_b+1}$$

$$\varphi = \begin{bmatrix} y(t_{k-1}) & \dots & y(t_{k-n_a}) & u(t_{k-d}) & \dots & u(t_{k-n_b-d}) \end{bmatrix}^\top$$

and

$$\tilde{v}(t_k) = A(q^{-1}, \rho)v(t_k). \quad (25)$$

Following the conventional PEM approach of the LTI framework (which is maximum likelihood estimation because of the normal distribution assumption on  $e(t_k)$ ), the prediction error  $\varepsilon_\theta(t_k)$  of (24) with respect to (23) is

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)A(q^{-1}, \rho)} \left( A(q^{-1}, \rho)y(t_k) - B(q^{-1}, \rho)u(t_k) \right), \quad (26)$$

where the filter  $D(q^{-1}, \eta)/C(q^{-1}, \eta)$  can be recognized as the inverse of the ARMA( $n_c, n_d$ ) noise model in (23). The polynomial operators commute and therefore  $\varepsilon_\theta(t_k)$  is equivalent to the error function  $\varepsilon_*(t_k)$  defined as:

$$\varepsilon_*(t_k) = A(p_k, q^{-1}, \rho)y_f(t_k) - B(p_k, q^{-1}, \rho)u_f(t_k), \quad (27)$$

where  $y_f = Q(q^{-1}, \theta)y$  and  $u_f = Q(q^{-1}, \theta)u$  represent the outputs of the prefiltering operation with

$$Q(q^{-1}, \theta) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)A(q^{-1}, \rho)}. \quad (28)$$

Therefore (24) is equivalent to:

$$y_f(t_k) = \varphi_f^\top(t_k)\rho + \tilde{v}_f(t_k) \quad (29)$$

with

$$\tilde{v}_f(t_k) = A(q^{-1}, \rho)v_f(t_k) = e(t_k). \quad (30)$$

In other words, if the optimal filter (28) is known *a priori*, it is possible to filter the data such that the estimation problem is reduced to the maximum likelihood estimation. This implies that a simple LS algorithm applied to the data prefiltered with (28) leads to the statistically optimal estimate under minor conditions.

### 3.2.2 The LPV case

Following the above introduced PEM approach in the LPV case (which is again maximum likelihood estimation because  $e_o(t_k) \in \mathcal{N}(0, \sigma_o^2)$ ), the prediction error  $\varepsilon_\theta(t_k)$  of (21) with respect to (11) is

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)} A^\dagger(p_k, q^{-1}, \rho) \left( A(p_k, q^{-1}, \rho) y(t_k) - B(p_k, q^{-1}, \rho) u(t_k) \right) \quad (31)$$

where  $D(q^{-1}, \eta)/C(q^{-1}, \eta)$  can be again recognized as the inverse of the ARMA( $n_c, n_d$ ) noise model of (11). In contrast to the LTI case, the polynomial operators do not commute as it has been shown in Section 2.3. Hence, no filter can be chosen such that both conditions

$$\begin{aligned} A(p_k, q^{-1}, \rho) y_f(t_k) &= \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)} A^\dagger(p_k, q^{-1}, \rho) A(p_k, q^{-1}, \rho) y(t_k) \\ B(p_k, q^{-1}, \rho) u_f(t_k) &= \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)} A^\dagger(p_k, q^{-1}, \rho) B(p_k, q^{-1}, \rho) u(t_k) \end{aligned}$$

are fulfilled simultaneously. Consequently, no filtering of the data can lead to a regression equation

$$y_f(t_k) = \varphi_f^\top(t_k) \rho + \tilde{v}_f(t_k) \quad (32)$$

which is equivalent to (21) and where  $\tilde{v}_f$  is white. In other words, by choosing  $\varphi$  such as in (22a) and therefore by assuming (21) (as in [3] and [4]) it is not possible to transform the estimation problem of (11) into a maximum likelihood estimation problem. The latter implies that no method proposed so far in the literature for solving the estimation of LPV-IO models or LTI-IO models can lead to optimal estimate in the LPV-BJ case by assuming the regression form (21). As a consequence, the existing theory needs to be modified in order to solve the identification problem stated in Section 2.5.

### 3.3 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 (11) as

$$\mathcal{M}_\theta \left\{ \begin{aligned} & \underbrace{\chi(t_k) + \sum_{i=1}^{n_a} a_{i,0} \chi(t_{k-i})}_{F(q^{-1})\chi(t_k)} + \underbrace{\sum_{i=1}^{n_a} \sum_{l=1}^{n_\alpha} a_{i,l} f_l(p_k) \chi(t_{k-i})}_{\chi_{i,l}(t_k)} \\ &= \sum_{j=0}^{n_b} \sum_{l=0}^{n_\beta} b_{j,l} g_l(p_k) \underbrace{u(t_{k-d-j})}_{u_{j,l}(t_k)} \\ & v(t_k) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} e(t_k) \\ & y(t_k) = \chi(t_k) + v(t_k) \end{aligned} \right. \quad (33)$$

where  $F(q^{-1}) = 1 + \sum_{i=1}^{n_a} a_{i,0} q^{-i}$  and  $g_0(t_k) = 1$ . Note that in this way, the LPV-BJ model is rewritten as a

*Multiple-Input Single-Output* (MISO) system with  $(n_b + 1)(n_\beta + 1) + n_a n_\alpha$  inputs  $\{\chi_{i,l}\}_{i=1, l=1}^{n_a, n_\alpha}$  and  $\{u_{j,l}\}_{j=0, l=0}^{n_b, n_\beta}$  as represented in Fig. 1. Given the fact that the polynomial operator commutes in this representation ( $F(q^{-1})$  does not depend on  $p_k$ ), (33) can be rewritten as

$$\begin{aligned} y(t_k) &= - \sum_{i=1}^{n_a} \sum_{l=1}^{n_\alpha} \frac{a_{i,l}}{F(q^{-1})} \chi_{i,l}(t_k) \\ &+ \sum_{j=0}^{n_b} \sum_{l=0}^{n_\beta} \frac{b_{j,l}}{F(q^{-1})} u_{j,l}(t_k) + H(q) e(t_k), \end{aligned} \quad (34)$$

which is an LTI representation. As (34) is an equivalent form of the model (11), thus under the Assumption A1, it holds that the data generating system  $\mathcal{S}_o$  has also a MISO-LTI interpretation.

## 4 Refined Instrumental variable for LPV systems

Based on the MISO-LTI formulation (34), it becomes possible in theory to achieve optimal PEM using linear regression. This allows to extend the *Refined Instrumental Variable* (RIV) approach of the LTI identification framework to provide an efficient way of identifying LPV-BJ models.

### 4.1 Optimal PEM for LPV-BJ models

Using (34),  $y(t_k)$  can be written in the regression form:

$$y(t_k) = \varphi^\top(t_k) \rho + \tilde{v}(t_k) \quad (35)$$

where,

$$\begin{aligned} \varphi(t_k) &= \left[ -y(t_{k-1}) \dots -y(t_{k-n_a}) \quad -\chi_{1,1}(t_k) \dots \right. \\ &\quad \left. -\chi_{n_a, n_\alpha}(t_k) \quad u_{0,0}(t_k) \dots u_{n_b, n_\beta}(t_k) \right]^\top \\ \rho &= \left[ a_{1,0} \dots a_{n_a,0} \quad a_{1,1} \dots a_{n_a, n_\alpha} \quad b_{0,0} \dots b_{n_b, n_\beta} \right]^\top \\ \tilde{v}(t_k) &= F(q^{-1}, \rho) v(t_k). \end{aligned}$$

It is important to notice that (35) and (21) are not equivalent. The extended regressor in (35) contains the noise-free output terms  $\{\chi_{i,l}\}$ . Therefore, by momentarily assuming that  $\{\chi_{i,l}(t_k)\}_{i=1, l=0}^{n_a, n_\alpha}$  are known *a priori*, the conventional PEM approach on (35) leads to the prediction error  $\varepsilon_\theta(t_k)$  given as:

$$\begin{aligned} \varepsilon_\theta(t_k) &= \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta) F(q^{-1}, \rho)} \left( F(q^{-1}, \rho) y(t_k) - \right. \\ &\quad \left. - \sum_{i=1}^{n_a} \sum_{l=1}^{n_\alpha} a_{i,l} \chi_{i,l}(t_k) + \sum_{j=0}^{n_b} \sum_{l=0}^{n_\beta} b_{j,l} u_{j,l}(t_k) \right) \end{aligned} \quad (36)$$

where  $D(q^{-1}, \eta)/C(q^{-1}, \eta)$  can be recognized again as the inverse of the ARMA( $n_c, n_d$ ) noise model in (11).

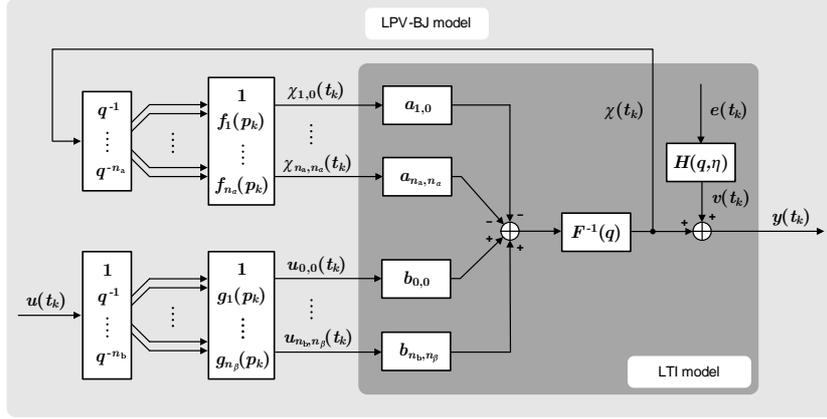


Fig. 1. MISO LTI interpretation of the LPV-BJ model

However, since the system written as in (34) is equivalent to a LTI system, the polynomial operators commute and (36) can be considered in the alternative form

$$\varepsilon_\theta(t_k) = F(q^{-1}, \rho)y_f(t_k) - \sum_{i=1}^{n_\alpha} \sum_{l=1}^{n_\alpha} a_{i,l} \chi_{i,l}^f(t_k) + \sum_{j=0}^{n_\beta} \sum_{l=0}^{n_\beta} b_{j,l} u_{j,l}^f(t_k) \quad (37)$$

where  $y_f(t_k)$ ,  $u_{j,l}^f(t_k)$  and  $\chi_{i,l}^f(t_k)$  represent the outputs of the prefiltering operation, using the filter (see [24]):

$$Q(q^{-1}, \theta) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)F(q^{-1}, \rho)}. \quad (38)$$

Based on (37), the associated linear-in-the-parameters model takes the form [24]:

$$y_f(t_k) = \varphi_f^\top(t_k)\rho + \tilde{v}_f(t_k), \quad (39)$$

where

$$\varphi_f(t_k) = \begin{bmatrix} -y_f(t_{k-1}) \dots -y_f(t_{k-n_\alpha}) -\chi_{1,1}^f(t_k) \dots \\ -\chi_{n_\alpha, n_\alpha}^f(t_k) \ u_{0,0}^f(t_k) \dots u_{n_\beta, n_\beta}^f(t_k) \end{bmatrix}^\top$$

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

#### 4.2 The refined instrumental variable estimate

Many methods of the LTI identification framework can be used to provide an efficient estimate of  $\rho$  given (39) where  $\tilde{v}_f(t_k)$  is a white noise. Here, the RIV method is chosen for the following reasons:

- RIV methods lead to optimal estimates in the LTI case if  $\mathcal{S}_o \in \mathcal{M}$ , see [16]. This statement is true as well for usual prediction error methods such as the extended LS approach.

- In practical situation of identification,  $\mathcal{G}_o \in \mathcal{G}$  might be fulfilled due to first principle or expert's knowledge. However, it is commonly fair to assume that  $H_o \notin \mathcal{H}$ . In such case, RIV methods has the advantage that they still provide consistent estimates whereas methods such as extended LS are biased and more advanced PEM methods needs robust initialization [13].

Aiming at the extension of the RIV approach for the estimation of LPV-BJ models, consider the relationship between the process input and output signals as in (35). Based on this form, the extended-IV estimate can be given as [16]:

$$\hat{\rho}_{\text{XIV}}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \left[ \frac{1}{N} \sum_{k=1}^N L(q)\zeta(t_k)L(q)\varphi^\top(t_k) \right] \rho - \left[ \frac{1}{N} \sum_{t=1}^N L(q)\zeta(t_k)L(q)y(t_k) \right] \right\|_W^2, \quad (40)$$

where  $\zeta(t_k)$  is the instrument,  $\|x\|_W^2 = x^\top W x$ , with  $W$  a positive definite weighting matrix and  $L(q)$  is a stable prefilter. if  $G_o \in \mathcal{G}$ , the extended-IV estimate is consistent under the following conditions<sup>1</sup>:

- C1**  $\bar{\mathbb{E}}\{L(q)\zeta(t_k)L(q)\varphi^\top(t_k)\}$  is full column rank.
- C2**  $\bar{\mathbb{E}}\{L(q)\zeta(t_k)L(q)\tilde{v}(t_k)\} = 0$ .

Moreover it has been shown in [16] and [22] that the minimum variance estimator can be achieved if:

- C3**  $W = I$ .
- C4**  $\zeta$  is chosen as the noise-free version of the extended regressor in (35) and is therefore defined in the present LPV case as:

$$\zeta(t_k) = \begin{bmatrix} -\chi(t_{k-1}) \dots -\chi(t_{k-n_\alpha}) -\chi_{1,1}(t_k) \dots \\ -\chi_{n_\alpha, n_\alpha}(t_k) \ u_{0,0}(t_k) \dots u_{n_\beta, n_\beta}(t_k) \end{bmatrix}^\top$$

<sup>1</sup> The notation  $\bar{\mathbb{E}}\{\cdot\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbb{E}\{\cdot\}$  is adopted from the prediction error framework of [12].

**C5**  $\mathcal{G}_o \in \mathcal{G}$  and  $n_\rho$  is equal to the minimal number of parameters required to represent  $\mathcal{G}_o$  with the considered model structure.

**C6**  $L(q)$  is chosen as in (38).

#### 4.3 Remarks on the use of the RIV approach

- Full column rank of  $\mathbb{E}\{L(q)\varphi(t_k)L(q)\varphi^\top(t_k)\}$  follows under Assumption A4 [3]. To fulfill C1 under A4, the discussion can be found in [16].
- In a practical situation none of  $F(q^{-1}, \rho)$ ,  $C(q^{-1}, \eta)$ ,  $D(q^{-1}, \eta)$  or  $\{a_{i,l}(\rho)\}_{i=1, l=0}^{n_a, n_\alpha}$ ,  $\{b_{j,l}(\rho)\}_{j=0, l=0}^{n_b, n_\beta}$  is known *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 instrumental variables (which guarantees C2), as well as the associated prefilters. This auxiliary model is based on the parameter estimates obtained at the previous iteration. Consequently, if convergence occurs, C4 and C6 are fulfilled.
- Convergence of the iterative RIV algorithm has not been proved so far and is only empirically assumed [23].
- The considered LPV model can be reformulated in a LTI-MISO form only under the condition that the noise-free output terms are *a priori* known (see Section 3.3). Therefore, even if the presented method considerably lowers the variance in the estimated parameters, the optimality cannot be guaranteed.

#### 4.4 Iterative LPV-RIV Algorithm

Based on the previous considerations, the iterative scheme of the RIV algorithm can be extended to the LPV case as follows.

##### Algorithm 2 (LPV-RIV)

**Step 1** Assume that as an initialization, an ARX estimate of  $\mathcal{M}_\theta$  is available by the LS approach, i.e.  $\hat{\theta}^{(0)} = [(\hat{\rho}^{(0)})^\top (\hat{\eta}^{(0)})^\top]^\top$  is given. Set  $\tau = 0$ .

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

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

where  $\hat{\rho}^{(\tau)}$  is estimated in the previous iteration. Based on  $\mathcal{M}_{\hat{\rho}^{(\tau)}}$ , deduce  $\{\hat{\chi}_{i,l}(t_k)\}_{i=1, l=0}^{n_a, n_\alpha}$  as given in (33). According to Assumption A5 each  $\hat{\chi}_{i,l}$  is bounded.

**Step 3** Compute the estimated filter:

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

and the associated filtered signals  $\{u_{j,l}^f(t_k)\}_{j=0, l=0}^{n_b, n_\beta}$ ,  $y_f(t_k)$  and  $\{\chi_{i,l}^f(t_k)\}_{i=1, l=0}^{n_a, n_\alpha}$ .

**Step 4** Build the filtered estimated regressor  $\hat{\varphi}_f(t_k)$  and in terms of C4 the filtered instrument  $\hat{\zeta}_f(t_k)$  as:

$$\begin{aligned} \hat{\varphi}_f(t_k) &= \begin{bmatrix} -y_f(t_{k-1}) \dots -y_f(t_{k-n_a}) -\hat{\chi}_{1,1}^f(t_k) \\ \dots -\hat{\chi}_{n_a, n_\alpha}^f(t_k) u_{0,0}^f(t_k) \dots u_{n_b, n_\beta}^f(t_k) \end{bmatrix}^\top \\ \hat{\zeta}_f(t_k) &= \begin{bmatrix} -\hat{\chi}_f(t_{k-1}) \dots -\hat{\chi}_f(t_{k-n_a}) -\hat{\chi}_{1,1}^f(t_k) \\ \dots -\hat{\chi}_{n_a, n_\alpha}^f(t_k) u_{0,0}^f(t_k) \dots u_{n_b, n_\beta}^f(t_k) \end{bmatrix}^\top \end{aligned}$$

**Step 5** The IV optimization problem can now be stated in the form

$$\hat{\rho}^{(\tau+1)}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \begin{bmatrix} \frac{1}{N} \sum_{k=1}^N \hat{\zeta}_f(t_k) \hat{\varphi}_f^\top(t_k) \\ \frac{1}{N} \sum_{k=1}^N \hat{\zeta}_f(t_k) y_f(t_k) \end{bmatrix} \rho \right\|^2 \quad (41)$$

where the solution is obtained as

$$\hat{\rho}^{(\tau+1)}(N) = \left[ \sum_{k=1}^N \hat{\zeta}_f(t_k) \hat{\varphi}_f^\top(t_k) \right]^{-1} \sum_{k=1}^N \hat{\zeta}_f(t_k) y_f(t_k).$$

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

$$\hat{v}(t_k) = y(t_k) - \hat{\chi}(t_k, \hat{\rho}^{(\tau)}). \quad (42)$$

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 [23]).

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

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 (11) with  $C(q^{-1}, \eta) = D(q^{-1}, \eta) = 1$  and consequently, Step 6 of Algorithm 2 can be skipped. Naturally, the LPV-SRIV does not minimize statistically optimal PEM for LPV-BJ models, however it still has a certain degree of robustness as it is shown in Section 5.

## 5 Simulation Example

As a next step, the performance of the proposed and of the existing methods in the literature are compared based on a representative simulation example.

### 5.1 Data generating system

The system taken into consideration is inspired by the example in [4] and is mathematically described as

$$\mathcal{S}_o \begin{cases} A_o(q, p_k) = 1 + a_1^o(p_k)q^{-1} + a_2^o(p_k)q^{-2} \\ B_o(q, p_k) = b_0^o(p_k)q^{-1} + b_1^o(p_k)q^{-2} \\ H_o(q) = \frac{1}{1 - q^{-1} + 0.2q^{-2}} \end{cases} \quad (43)$$

where  $v(t_k) = H_o(q)e(t_k)$  and

$$a_1^o(p_k) = 1 - 0.5p_k - 0.1p_k^2, \quad (44a)$$

$$a_2^o(p_k) = 0.5 - 0.7p_k - 0.1p_k^2, \quad (44b)$$

$$b_0^o(p_k) = 0.5 - 0.4p_k + 0.01p_k^2, \quad (44c)$$

$$b_1^o(p_k) = 0.2 - 0.3p_k - 0.02p_k^2. \quad (44d)$$

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  $u(t_k)$  is taken as a white noise with a uniform distribution  $\mathcal{U}(-1, 1)$  and with length  $N = 4000$  to generate data sets  $\mathcal{D}_N$  of  $\mathcal{S}_o$ .

## 5.2 Model structures

In the sequel, the *One Step Instrumental Variable* (OSIV) method presented in [4] and the conventional *Least Square* (LS) method such as the one used in [3] are compared to the proposed IV approaches. Both methods assume the following model structure:

$$\mathcal{M}_\theta^{\text{LS, OSIV}} \begin{cases} A(p_k, q^{-1}, \rho) = 1 + a_1(p_k)q^{-1} + a_2(p_k)q^{-2} \\ B(p_k, q^{-1}, \rho) = b_0(p_k)q^{-1} + b_1(p_k)q^{-2} \\ H(p_k, q, \rho) = A^\dagger(p_k, q^{-1}, \rho) \end{cases}$$

$$a_1(p_k) = a_{1,0} + a_{1,1}p_k + a_{1,2}p_k^2 \quad (45a)$$

$$\text{where } a_2(p_k) = a_{2,0} + a_{2,1}p_k + a_{2,2}p_k^2 \quad (45b)$$

$$b_0(p_k) = b_{0,0} + b_{0,1}p_k + b_{0,2}p_k^2 \quad (45c)$$

$$b_1(p_k) = b_{1,0} + b_{1,1}p_k + b_{1,2}p_k^2 \quad (45d)$$

In contrast with these model structures, the proposed LPV *Refined Instrumental Variable* method (LPV-RIV) represents the situation  $\mathcal{S}_o \in \mathcal{M}$  and assumes the following LPV-BJ model:

$$\mathcal{M}_\theta^{\text{LPV-RIV}} \begin{cases} A(p_k, q^{-1}, \rho) = 1 + a_1(p_k)q^{-1} + a_2(p_k)q^{-2} \\ B(p_k, q^{-1}, \rho) = b_0(p_k)q^{-1} + b_1(p_k)q^{-2} \\ H(p_k, q, \eta) = \frac{1}{1 + d_1q^{-1} + d_2q^{-2}} \end{cases}$$

with  $a_1(p_k)$ ,  $a_2(p_k)$ ,  $b_0(p_k)$ ,  $b_1(p_k)$  as given in (45a-d), while the LPV *Simplified Refined Instrumental Variable* method (LPV-SRIV) represents the case when  $\mathcal{G}_o \in \mathcal{G}$ ,  $\mathcal{H}_o \notin \mathcal{H}$  and assumes the following LPV-OE model:

$$\mathcal{M}_\theta^{\text{LPV-SRIV}} \begin{cases} A(p_k, q^{-1}, \rho) = 1 + a_1(p_k)q^{-1} + a_2(p_k)q^{-2} \\ B(p_k, q^{-1}, \rho) = b_0(p_k)q^{-1} + b_1(p_k)q^{-2} \\ H(p_k, q, \eta) = 1 \end{cases}$$

The robustness of the proposed and existing algorithms are investigated with respect to different *signal-to-noise ratios*  $\text{SNR} = 10 \log \frac{P_{\chi_o}}{P_{e_o}}$ , where  $P_{\chi_o}$  and  $P_{e_o}$  are the average power of signals  $\chi_o$  and  $e_o$  respectively. To provide representative results, a Monte-Carlo simulation of  $N_{\text{MC}} = 100$  runs with new noise realization is accomplished at different noise levels: 15dB, 10dB, 5dB and 0dB. For the Monte-Carlo simulation at  $\text{SNR} = 15\text{dB}$ ,

Table 1 and 2 show the detailed results about mean and standard deviation of the estimated parameters. In some practical application, only one realization is accessible and therefore it is not possible to compute the uncertainty through *Monte-Carlo simulation* (MCS). In this latter case it is important to be able to determine the *standard error* (SE) on the estimated parameters with a *single realization* (SR). Therefore the results of SR are also given in these tables. Note that it is possible to compute the SR standard error  $\text{SE} = \text{diag}(\hat{P}_\rho)^{1/2}$  from the covariance matrix  $\hat{P}_\rho = \hat{\sigma}_e^2 (\sum_{k=1}^N \hat{\zeta}_f(t_k) \hat{\zeta}_f^\top(t_k))^{-1}$ . With respect to the considered methods, Table 3 shows the norm of the bias (BN)  $\|\rho_o - \mathbb{E}(\hat{\rho})\|_2$  and variance (VN) of the estimated parameter vector  $\|\mathbb{E}(\hat{\rho} - \mathbb{E}(\hat{\rho}))\|_2$ , where  $\mathbb{E}$  is the mean operator over the Monte-Carlo simulation and  $\|\cdot\|_2$  is the  $\mathcal{L}_2$  norm. 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 3 that the IV methods are unbiased according to the theoretical results. It might not appear clearly for the OSIV method when using SNR under 10dB but considering the variances induced, the bias is only due to the relatively low number of simulation runs. Under 10dB, the results of the OSIV cannot be considered as relevant as they induce such large variances. In the present BJ system, the OSIV method does not lead to satisfying results and cannot be used in practical applications. It can be seen that for SNR down to 5dB, the LPV-RIV produces variance in the estimated parameters which are very close to the one obtained with the LS method, not mentioning that the bias has been completely suppressed. Although the statistical optimality of the algorithm cannot be proved, this latter result shows on this example, that the LPV-RIV algorithm dramatically improves the accuracy of the estimates. The suboptimal LPV-SRIV method offers satisfying results, considering that the noise model is not correctly assumed. The variance in the estimated parameters is twice as much as in the LPV-RIV case and it is close to the variance of the LS method. Finally, it can be pointed out that the number of iterations is high in comparison to the linear case for RIV methods (typically, 4 iterations are needed in a second order linear case). Tables 1 and 2 show that detailed results lead to the same conclusion as when looking at Table 3. It can be finally seen from 2 that the LPV-RIV method estimates accurately the noise model and that the standard error obtained from a single realization is well correlated to the standard deviation obtained through Monte-Carlo simulation.

## 6 Conclusion

This paper highlighted the lack of efficient methods in the literature to handle the estimation of LPV Box-Jenkins models. It has been shown that the conventional formulation of least squares estimation cannot lead to statistically optimal parameter estimates. As a solution,

Table 1

Mean and standard deviation of the estimated  $A$  polynomial parameters at SNR = 15dB

method	true value	$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{2,0}$	$a_{2,1}$	$a_{2,2}$
LS	mean	-0.3794	2.2373	-2.0584	-0.1085	-0.0755	-0.4786
	std	0.0219	0.0663	0.0591	0.0125	0.0600	0.0558
OSIV	mean	1.0259	-0.6161	0.0205	0.5092	-0.7510	-0.0377
	std	0.3023	1.0330	0.8605	0.1227	0.4348	0.3986
LPV-SRIV MCS	mean	1.0003	-0.5013	-0.0971	0.5007	-0.7047	-0.0943
	std	0.0313	0.1022	0.0893	0.0106	0.0650	0.074
LPV-SRIV SR	$\hat{\rho}$	0.9801	-0.3743	-0.2120	0.4978	-0.7154	-0.0736
	SE	0.0377	0.1567	0.1486	0.0171	0.1010	0.1099
LPV-RIV MCS	mean	0.9999	-0.5020	-0.0989	0.5005	-0.7050	-0.0962
	std	0.0170	0.0589	0.0610	0.0084	0.0488	0.0523
LPV-RIV SR	$\hat{\rho}$	0.9947	-0.5053	-0.0506	0.4981	-0.7303	-0.0350
	SE	0.0120	0.0479	0.0435	0.0050	0.0330	0.0368

Table 2

Mean and standard deviation of the estimated  $B$  and  $D$  polynomial parameters at SNR = 15dB

method	true value	$b_{0,0}$	$b_{0,1}$	$b_{0,2}$	$b_{1,0}$	$b_{1,2}$	$b_{2,2}$	$d_1$	$d_2$
LS	mean	0.5043	-0.4045	0.0085	-0.3201	0.7890	-0.7335	X	X
	std	0.0039	0.0233	0.0219	0.0097	0.0284	0.0238	X	X
OSIV	mean	0.4986	-0.3991	0.0110	0.2096	-0.3409	0.0181	X	X
	std	0.0115	0.0564	0.0503	0.1151	0.3731	0.2922	X	X
LPV-SRIV MCS	mean	0.4996	-0.3998	0.0101	0.1997	-0.3004	-0.0190	X	X
	std	0.0038	0.0183	0.0171	0.0104	0.0367	0.0300	X	X
LPV-SRIV SR	$\hat{\rho}$	0.4998	-0.3783	-0.0108	0.1996	-0.2885	-0.0273	X	X
	SE	0.0044	0.0221	0.0216	0.0138	0.0561	0.0493	X	X
LPV-RIV MCS	mean	0.4998	-0.3993	0.0092	0.1998	-0.3008	-0.0194	-1.003	0.2042
	std	0.0020	0.0106	0.0106	0.0055	0.0228	0.0214	0.0171	0.0172
LPV-RIV SR	$\hat{\rho}$	0.5020	-0.4142	0.0245	0.1971	-0.2889	-0.0219	X	X
	SE	0.0016	0.0075	0.0072	0.0042	0.0165	0.0141	X	X

Table 3

Estimator bias and variance norm at different SNR

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

the LPV identification problem is reformulated and a method to estimate efficiently LPV-BJ models with  $p$ -independent noise process was proposed. The introduced method has been compared to the existing methods of the literature both in terms of theoretical analysis and in terms of a representative numerical example. The presented example has shown that the proposed procedure is robust to noise and outperforms the existing methods. Even if RIV based methods are in theory unbiased for  $p$ -dependant noise models, the case was not studied in the present paper and will be considered in further development of this approach. As continuation of the

presented work, extensions of the method to closed-loop and continuous-time LPV system identification are also intended.

## References

- [1] B. Bamieh and L. Giarré. Identification of linear parameter varying models. In *Proceedings of the 38th IEEE Conference on Decision and Control*, pages 1505–1510, Phoenix, USA, December 1999.
- [2] B. Bamieh and L. Giarré. Identification for a general class of LPV models. In *Proceedings of the 11th IFAC Symposium on System Identification*, Santa Barbara, California, USA, June 2000.
- [3] B. Bamieh and L. Giarré. Identification of linear parameter-varying models. *International Journal of Robust and Nonlinear Control*, 12:841–853, 2002.
- [4] M. Butcher, A. Karimi, and R. Longchamp. On the consistency of certain identification methods for linear parameter varying systems. In *Proceedings of the 17th IFAC World Congress*, pages 4018–4023, Seoul, Korea, July 2008.
- [5] P. L. dos Santos, J. A. Ramos, and J. L. M. de Carvalho. Identification of linear parameter varying systems using an iterative deterministic-stochastic subspace approach. In *Proc. of the European Control Conference*, pages 4867–4873, Kos, Greece, July 2007.

- [6] M. Gevers, A. S. Bazanella, and L. Miskovic. Informative data: how to get just sufficiently rich? In *proceedings of the 47th IEEE Conference on Decision and Control*, pages 1962–1967, Cancun, Mexico, December 2008.
- [7] L. Giarré, D. Bauso, P. Falugi, and B. Bamieh. LPV model identification for gain scheduling control: An application to rotating stall and surge control problem. *Control Engineering Practice*, 14:351–361, 2006.
- [8] M. Gilson and P. Van den Hof. Instrumental variable methods for closed-loop system identification. *Automatica*, 41, Issue 2:241–249, 2005.
- [9] M. Gilson, H. Garnier, P. C. Young, and P. Van den Hof. Refined instrumental variable methods for closed-loop system identification methods. In *proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, June 2009.
- [10] H. Garnier and L. Wang (Editors). *Identification of Continuous-time Models from Sampled Data*. Springer-Verlag, London, 2008.
- [11] V. Laurain, M. Gilson, H. Garnier, and P. C. Young. Refined instrumental variable methods for identification of Hammerstein continuous time Box-Jenkins models. In *proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec 2008.
- [12] L. Ljung. *System identification : theory for the user - Second Edition*. Prentice-Hall, 1999.
- [13] L. Ljung. Experiments with identification of continuous time models. In *proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009.
- [14] M. Lovera and G. Mercère. Identification for gain-scheduling: a balanced subspace approach. In *Proc. of the American Control Conference*, pages 858–863, New York City, USA, July 2007.
- [15] G. P. Rao and H. Garnier. Identification of continuous-time systems: direct or indirect ? *Systems Science*, 30, Issue 3:25–50, 2004.
- [16] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer-Verlag, New York, 1983.
- [17] P. Stoica and M. Jansson. MIMO system identification: state-space and subspace approximations versus transfer function and instrumental variables. *IEEE Transactions on Signal Processing*, 48(11):3087 – 3099, 2000.
- [18] R. Tóth. *Modeling and Identification of Linear Parameter-Varying Systems: An Orthonormal Basis Function Approach*. PhD thesis, Delft University of Technology, 2008.
- [19] R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof. An LPV identification framework based on orthonormal basis functions. In *proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009.
- [20] J. W. van Wingerden and M. Verhaegen. Subspace identification of bilinear and LPV systems for open- and closed-loop data. *Automatica*, 45:372–381, 2009.
- [21] X. Wei and L. Del Re. On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque. In *Proceedings of the 14th IFAC Symposium on System Identification*, pages 517–522, Newcastle, Australia, March 2006.
- [22] P. C. Young. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, 1984.
- [23] P. C. Young. The refined instrumental variable method: Unified estimation of discrete and continuous-time transfer function models. *Journal Européen des Systèmes Automatisés*, 42:149–179, 2008.
- [24] P. C. Young, H. Garnier, and M. Gilson. Identification of continuous-time models from sampled data. In *Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models*. H. Garnier and L. Wang (editor), pages 91–131, Springer-Verlag, London, 2008.