Identification of LPV Output-Error and Box-Jenkins Models via Optimal Refined Instrumental Variable Methods

V. Laurain*, M. Gilson*, R. Tóth**, H. Garnier*

Abstract— **Identification of Linear Parameter-Varying (LPV) models is often addressed in an Input-Output (IO) setting. However, statistical properties of the available algorithms are not fully understood. Most methods apply auto regressive models with exogenous input (ARX) which are unrealistic in most practical applications due to their associated noise structure. A few methods have been also proposed for Output Error (OE) models, however it can be shown that the estimates are not statistically efficient. To overcome this problem, the paper proposes a Refined Instrumental Variable (RIV) method dedicated to LPV Box-Jenkins (BJ) models where the noise part is an additive colored noise. The statistical performance of the algorithm is analyzed and compared with existing methods.**

I. Introduction

Despite the advances of the LPV control field, identification of this system class is not well developed. Existing LPV identification approaches are mainly characterized by the type of LPV model structure used: *Input-Output* (IO) [3], [1], [13] *State Space* (SS) [7], [12], or *Orthogonal Basis Functions* (OBFs) based models [11]. 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. Moreover, 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 [9].

In the LPV case, most of the methods developed for IO models based identification are derived under a linear regression form [13], [4], [3]. By using the concepts of the *linear time invariant* (LTI) PE framework, recursive *least squares* (LS) and *instrumental variable* (IV) methods have been also introduced [4], [3].

Due to the linear regressor based estimation, the usual model structure in existing methods is assumed to be ARX. This assumption is unrealistic in most practical applications as it assumes that both the noise and the process part of the system contain the same dynamics (for the reasoning see [6]). Even if some IV methods have been developed for LPV *Output Error* (OE) models in [3], it can be shown that the obtained estimates are not statistically efficient. The underlying reason is that statistically optimal estimates cannot be reached by using linear regression as presented so far in the

*Centre de Recherche en Automatique de Nancy (CRAN), Nancy-université, CNRS, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France, vincent.laurain@cran.uhp-nancy.fr

**Delft Center for Systems and Control (DCSC), Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, r.toth@tudelft.nl

literature because the identification methods are commonly based on the concept of the LTI framework and hence, they do not consider the non-commutativity of LPV filters. In other words, there is lack of a LPV identification approach, which is able to yield statistically efficient estimates for LPV-IO models under colored noise conditions, e.g. as in a *Box-Jenkins* (BJ) setting, which is the case in many practical applications (see [6]).

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 II, IO representation based models of LPV systems are introduced together with the concept of LPV identification in the prediction-error framework. A reformulation of the dynamical description of LPV data generating plants in the considered setting is presented in Section III which makes possible the extension of LTI-IV methods to the LPV framework. In Section IV, LPV-IV methods are introduced and analyzed, while their performance increase compared to other methods is shown in Section V. Finally in Section VI, the main conclusions of the paper are drawn and directions of future research are indicated.

II. Problem description

A. System description

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

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

where d is the delay, 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. $q^{-i}u(t_k) = u(t_{k-i})$. $A_{\mathrm{o}}(p_k, q^{-1})$ and $B_{\mathrm{o}}(p_k, q^{-1})$ are polynomials in q^{-1} of degree n_{a} and n_{b} respectively:

$$
A_{\mathrm{o}}(p_k, q^{-1}) = 1 + \sum_{i=1}^{n_{\mathrm{a}}} a_i^{\mathrm{o}}(p_k) q^{-i}, \tag{2a}
$$

$$
B_{\mathrm{o}}(p_k, q^{-1}) = \sum_{j=0}^{n_{\mathrm{b}}} b_j^{\mathrm{o}}(p_k) q^{-i}, \tag{2b}
$$

where the coefficients a_i and b_j are real meromorphic functions, e.g. rational functions, with static dependence on

the so-called scheduling parameter $p : \mathbb{Z} \to \mathbb{P}$. It is assumed that these coefficients are non-singular on the scheduling space $\mathbb{P} \subseteq \mathbb{R}^{n_{\mathbb{P}}}$ thus the solutions of $\mathcal{S}_{\mathrm{o}}$ are well-defined and the process part is completely characterized by the coefficient functions $\{a_i^{\mathrm{o}}\}_{i=1}^{n_{\mathrm{a}}}$ and $\{b_j^{\mathrm{o}}\}_{j=0}^{n_{\mathrm{b}}}$.

Most methods in the literature assume an ARX type of data generating system, but this assumption is often unrealistic in practical applications as it is based on the concept that both the noise and the process part of $\mathcal{S}_{\mathrm{o}}$ contain the same dynamics. Often the colored noise associated with the sampled output measurement $y(t_k)$ has a rational spectral density which has no dependence on p. Therefore, it is a more realistic assumption that v_{o} is represented by a discrete-time *autoregressive moving average* (ARMA) model:

$$v_{\mathrm{o}}(t_k) = H_{\mathrm{o}}(q)e_{\mathrm{o}}(t_k) = \frac{C_{\mathrm{o}}(q^{-1})}{D_{\mathrm{o}}(q^{-1})}e_{\mathrm{o}}(t_k), \qquad (3)$$

where $C_{\mathrm{o}}(q^{-1})$ and $D_{\mathrm{o}}(q^{-1})$ are monic polynomials with constant coefficients and with respective degree n_{c} and n_{d}. Furthermore, all roots of $z^{n_{\mathrm{d}}}D_{\mathrm{o}}(z^{-1})$ are inside the unit disc. It can be noticed that in case $C_{\mathrm{o}}(q^{-1}) = D_{\mathrm{o}}(q^{-1}) = 1$, (3) defines an OE noise model.

B. 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 (3). In the proposed model structure the noise model and the process model is parameterized separately.

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

$$\mathcal{G}_\rho : \left(A(p_k, q^{-1}, \rho), B(p_k, q^{-1}, \rho) \right) = (\mathcal{A}_\rho, \mathcal{B}_\rho) \qquad (4)$$

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

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

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

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 ρ are stacked columnwise in the parameter vector,

$$\rho = \begin{bmatrix} \mathtt{a}_1 & \ldots & \mathtt{a}_{n_{\mathrm{a}}} & \mathtt{b}_0 & \ldots & \mathtt{b}_{n_{\mathrm{b}}} \end{bmatrix}^\top \in \mathbb{R}^{n_\rho}, \quad (6)$$

where

$$\mathtt{a}_i = \begin{bmatrix} a_{i,0} & a_{i,1} & \ldots & a_{i,n_\alpha} \end{bmatrix} \in \mathbb{R}^{n_\alpha + 1}$$
$$\mathtt{b}_j = \begin{bmatrix} b_{j,0} & b_{j,1} & \ldots & b_{j,n_\beta} \end{bmatrix} \in \mathbb{R}^{n_\beta + 1}$$

and $n_\rho = n_{\mathrm{a}}(n_\alpha + 1) + (n_{\mathrm{b}} + 1)(n_\beta + 1)$. 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 (4).

2) Noise model: The noise model is denoted by $\mathcal{H}$ and defined as a LTI transfer function:

$$\mathcal{H}_\eta : (H(q, \eta)) \qquad (7)$$

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

$$H(q, \eta) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} = \frac{1 + c_1 q^{-1} + \ldots + c_{n_{\mathrm{c}}} q^{-n_{\mathrm{c}}}}{1 + d_1 q^{-1} + \ldots + d_{n_{\mathrm{d}}} q^{-n_{\mathrm{d}}}}. \quad (8)$$

The associated model parameters η are stacked columnwise in the parameter vector,

$$\eta = \begin{bmatrix} c_1 & \ldots & c_{n_{\mathrm{c}}} & d_1 & \ldots & d_{n_{\mathrm{d}}} \end{bmatrix}^\top \in \mathbb{R}^{n_\eta}, \quad (9)$$

where $n_\eta = n_{\mathrm{c}} + n_{\mathrm{d}}$. Additionally, denote $\mathcal{H} = \{\mathcal{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$, the collection of all noise models in the form of (7).

3) Whole model: With respect to a given process and noise part $(\mathcal{G}_\rho, \mathcal{H}_\eta)$, the parameters can be collected as

$$\theta = \begin{bmatrix} \rho \\ \eta \end{bmatrix}, \qquad (10)$$

and the signal relations of the LPV-BJ model, denoted in the sequel as $\mathcal{M}_\theta$, are defined as:

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

Based on the previously defined 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

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

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

C. 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 $\mathcal{S}_{\mathrm{o}}$ defined as (1) and a data set $\mathcal{D}_N = \{y(t_k), u(t_k), p_k\}_{k=0}^N$ collected from $\mathcal{S}_{\mathrm{o}}$. Based on the LPV-BJ model structure $\mathcal{M}_\theta$ defined by (11), estimate the parameter vector θ using $\mathcal{D}_N$ under the following assumptions:

A1 $\mathcal{S}_{\mathrm{o}} \in \mathcal{M}$, i.e. there exits a $\mathcal{G}_{\mathrm{o}} \in \mathcal{G}$ and a $\mathcal{H}_{\mathrm{o}} \in \mathcal{H}$ such that $(\mathcal{G}_{\mathrm{o}}, \mathcal{H}_{\mathrm{o}})$ is equal to $\mathcal{S}_{\mathrm{o}}$.

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 $(\mathcal{G}_{\mathrm{o}}, \mathcal{H}_{\mathrm{o}})$ is identifiable for any trajectory of p.

A3 $u(t_k)$ is not correlated to $e_{\mathrm{o}}(t_k)$.

A4 $\mathcal{D}_N$ is informative with respect to $\mathcal{M}$.

A5 $\mathcal{S}_{\mathrm{o}}$ is uniformly frozen stable, i.e. for any $\mathrm{p} \in \mathbb{P}$, the roots of $z^{n_{\mathrm{a}}}(A(\mathrm{p}, z^{-1}))$ are in the unit disc [10].

III. REFORMULATION OF THE SYSTEM EQUATIONS

All methods for LPV-IO parametric identification proposed in the literature so far are based on linear regression methods such as least squares or instrumental variables [2] [3]. 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:

$$y(t_k) = \varphi^\top(t_k)\rho + \tilde{v}(t_k) \tag{13}$$

with ρ as defined in (6),

$$\varphi(t_k) = \begin{bmatrix} -\mathbf{y}(t_k)f_0(p_k) \\ \vdots \\ -\mathbf{y}(t_k)f_{n_\alpha}(p_k) \\ \mathbf{u}(t_k)g_0(p_k) \\ \vdots \\ \mathbf{u}(t_k)g_{n_\beta}(p_k) \end{bmatrix} \in \mathbb{R}^{n_\rho}, \tag{14a}$$

$$\mathbf{y}(t_k) = \begin{bmatrix} y(t_{k-1}) \\ \vdots \\ y(t_{k-n_a}) \end{bmatrix}, \quad \mathbf{u}(t_k) = \begin{bmatrix} u(t_{k-d}) \\ \vdots \\ u(t_{k-n_b-d}) \end{bmatrix}, \tag{14b}$$

and

$$\tilde{v}(t_k) = A(p_k, q^{-1}, \rho)v(t_k). \tag{14c}$$

For an LPV-ARX system, a LS algorithm based on (13) leads to statistically optimal estimates as it minimizes the prediction error on $\tilde{v}$ which is a white noise in this case. However, for LPV-BJ systems, the LS estimate is biased. Moreover, in opposition to the LTI case, it can be proven that in general it is impossible to yield statistically optimal estimates using (13). The proof is only briefly explained due to space restrictions. For LTI-BJ systems, statistically optimal methods require an optimal prefiltering of the data contained in the regressor [8]. However, for LPV systems, the property of commutativity in filtering does not hold as $q^{-1}A(p_k, q^{-1}) \neq A(p_k, q^{-1})q^{-1}$ (see [11]) and it can be shown that no optimal filtering can be applied to the regressor (14a). In other word, no existing method in the LPV literature can deal in a statistical optimal way with the identification of LPV-BJ data generating systems.

The existing theory needs to be modified in order to solve the identification problem stated in Section II-C. One particular way is to rewrite the signal relations of (11) into the form:

$$\mathcal{M}_\theta \begin{cases} \underbrace{\chi(t_k) + \sum_{i=1}^{n_a} a_{i,0}\chi(t_{k-i})}_{F(q^{-1})\chi(t_k)} + \sum_{i=1}^{n_a}\sum_{l=1}^{n_\alpha} a_{i,l}\underbrace{f_l(p_k)\chi(t_{k-i})}_{\chi_{i,l}(t_k)} = \\ \qquad \sum_{j=0}^{n_b}\sum_{l=0}^{n_\beta} b_{j,l}\underbrace{g_l(p_k)u(t_{k-d-j})}_{u_{j,l}(t_k)} \\ v(t_k) = \dfrac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)}e(t_k) \\ y(t_k) = \chi(t_k) + v(t_k) \end{cases} \tag{15}$$

where $F(q^{-1}) = 1 + \sum_{i=1}^{n_a} a_{i,0}q^{-i}$. 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 $\{u_{j,l}\}_{j=0,l=0}^{n_b,n_\beta}$. Given the fact that the polynomial operator commutes in this representation ($F(q^{-1})$ does not depend on p_k), (15) can be rewritten as

$$y(t_k) = -\sum_{i=1}^{n_a}\sum_{l=1}^{n_\alpha} \frac{a_{i,l}}{F(q^{-1})}\chi_{i,l}(t_k) \\ + \sum_{j=0}^{n_b}\sum_{l=0}^{n_\beta} \frac{b_{j,l}}{F(q^{-1})}u_{k,j}(t_k) + H(q)e(t_k), \tag{16}$$

which is a LTI representation. As (16) 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.

IV. REFINED INSTRUMENTAL VARIABLE APPROACH FOR LPV SYSTEMS

Based on the MISO-LTI formulation (16), it becomes possible to achieve optimal prediction error minimization (PEM) using linear regression and to extend the *Refined Instrumental Variable* (RIV) approach of the LTI identification framework to provide an efficient way of identifying LPV-BJ models.

A. Optimal PEM for LPV-BJ models

Using (16), $y(t_k)$ can be written in the regression form:

$$y(t_k) = \varphi^\top(t_k)\rho + \tilde{v}_r(t_k) \tag{17}$$

where,

$$\varphi(t_k) = \begin{bmatrix} -y(t_{k-1}) & \dots & -y(t_{k-n_a}) & -\chi_{1,1}(t_k) & \dots \\ & -\chi_{n_a,n_\alpha}(t_k) & u_{0,0}(t_k) & \dots & u_{n_b,n_\beta}(t_k) \end{bmatrix}^\top$$

$$\rho = \begin{bmatrix} a_{1,0} & \dots & a_{n_a,0} & a_{1,1} & \dots & a_{n_a,n_\alpha} & b_{0,0} & \dots & b_{n_b,n_\beta} \end{bmatrix}^\top$$

and

$$\tilde{v}_r(t_k) = F(q^{-1}, \rho)v(t_k). \tag{18}$$

It is important to notice that (17) and (13) are not equivalent. The extended regressor in (17) contains the noise-free output terms $\{\chi_{i,k}\}$. By following the conventional PEM approach on (17), the prediction error $\varepsilon_\theta(t_k)$ is given as:

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)F(q^{-1}, \rho)}\bigg(F(q^{-1}, \rho)y(t_k) - \\ -\sum_{i=1}^{n_a}\sum_{l=1}^{n_\alpha} a_{i,l}\chi_{i,l}(t_k) + \sum_{j=0}^{n_b}\sum_{l=0}^{n_b} b_{j,l}u_{j,l}(t_k)\bigg) \tag{19}$$

where $D(q^{-1}, \eta)/C(q^{-1}, \eta)$ can be recognized as the inverse of the ARMA(n_c, n_d) noise model in (11). However, since the system written as in (16) is equivalent to a LTI system, the

polynomial operators commute and (19) can be considered in the alternative form

$$\varepsilon_\theta(t_k) = F(q^{-1}, \rho) y_{\mathrm{f}}(t_k) - \sum_{i=1}^{n_{\mathrm{a}}} \sum_{l=1}^{n_\alpha} a_{i,l} \chi_{i,l}^{\mathrm{f}}(t_k)$$
$$+ \sum_{j=0}^{n_{\mathrm{b}}} \sum_{l=0}^{n_\beta} b_{j,l} u_{j,l}^{\mathrm{f}}(t_k) \quad (20)$$

where $y_{\mathrm{f}}(t_k)$, $u_{k,j}^{\mathrm{f}}(t_k)$ and $\chi_{i,l}^{\mathrm{f}}(t_k)$ represent the outputs of the prefiltering operation, using the filter (see [16]):

$$Q(q^{-1}, \theta) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta) F(q^{-1}, \rho)}. \quad (21)$$

Based on (20), the associated linear-in-the-parameters model takes the form [16]:

$$y_{\mathrm{f}}(t_k) = \varphi_{\mathrm{f}}^\top(t_k) \rho + \tilde{v}_{\mathrm{f}}(t_k), \quad (22)$$

where

$$\varphi_{\mathrm{f}}(t_k) = \begin{bmatrix} -y_{\mathrm{f}}(t_{k-1}) & \dots & -y_{\mathrm{f}}(t_{k-n_{\mathrm{a}}}) & -\chi_{1,1}^{\mathrm{f}}(t_k) & \dots \\ -\chi_{n_{\mathrm{a}}, n_\alpha}^{\mathrm{f}}(t_k) & u_{0,0}^{\mathrm{f}}(t_k) & \dots & u_{n_{\mathrm{b}}, n_\beta}^{\mathrm{f}}(t_k) \end{bmatrix}^\top$$

and

$$\tilde{v}_{\mathrm{f}}(t_k) = F(q^{-1}, \rho) v_{\mathrm{f}}(t_k) =$$
$$F(q^{-1}, \rho) \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta) F(q^{-1}, \rho)} v(t_k) = e(t_k). \quad (23)$$

B. The refined instrumental variable estimate

Many methods of the LTI identification framework can be used to provide an efficient estimate of ρ given (22) where $\tilde{v}_{\mathrm{f}}(t_k)$ is a white noise. Here, the refined instrumental variable method is chosen as it leads to optimal estimates if $\mathcal{S}_{\mathrm{o}} \in \mathcal{M}$ [8] and provides consistent estimates in case $H_{\mathrm{o}} \notin \mathcal{H}$.

Aiming at the application of the RIV approach for the estimation of LPV-BJ models, consider the relationship between the process input and output signals as in (17). Based on this form, the extended-IV estimate can be given as [8]:

$$\hat{\rho}_{\mathrm{XIV}}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N} \sum_{k=1}^{N} L(q) \zeta(t_k) L(q) \varphi^\top(t_k) \right] \right.$$
$$\left. - \left[\frac{1}{N} \sum_{t=1}^{N} L(q) \zeta(t_k) L(q) y(t_k) \right] \right\|_W^2, \quad (24)$$

where $\zeta(t_k)$ is the instrumental vector, $\|x\|_W^2 = x^T W x$, with W a positive definite weighting matrix and $L(q)$ is a stable prefilter. If $G_{\mathrm{o}} \in \mathcal{G}$, the extended-IV estimate is consistent under the following conditions[1]:

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 [8] and [14] that the minimum variance estimator can be achieved if:

---

[1] 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 [6].

C3 $W = I$.
C4 ζ is chosen as the noise-free version of the extended regressor in (17) and is therefore defined in the present LPV case as:

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

C5 $\mathcal{G}_{\mathrm{o}} \in \mathcal{G}$ and n_ρ is equal to the minimal number of parameters required to represent $\mathcal{G}_{\mathrm{o}}$ with the considered model structure.
C6 $L(q)$ is chosen as $Q(q^{-1}, \theta_{\mathrm{o}})$ in (21).

C. Remarks on the use of the RIV approach

- Full column rank of $\bar{\mathbb{E}}\{L(q)\varphi(t_k)L(q)\varphi^\top(t_k)\}$ follows under Assumption A4 [2]. To fulfill C1 under A4, the discussion can be found in [8].
- In a practical situation none of $F(q^{-1}, \rho)$, $\{a_{i,l}(\rho)\}_{i=1,l=0}^{n_{\mathrm{a}}, n_\alpha}$, $\{b_{j,l}(\rho)\}_{j=0,l=0}^{n_{\mathrm{b}}, n_\beta}$, $C(q^{-1}, \eta)$, $D(q^{-1}, \eta)$ 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. It is important to note that convergence of the iterative RIV has not been proved so far and is only empirically assumed [15], [5].

D. Iterative LPV-RIV Algorithm

Based on the previous considerations, the proposed RIV algorithm for LPV systems is given:

Algorithm 1 (LPV-RIV):

Step 1 An ARX estimate of $\mathcal{M}_\theta$ is computed using the LS approach, $\hat{\theta}^{(0)} = \begin{bmatrix} (\hat{\rho}^{(0)})^\top & (\hat{\eta}^{(0)})^\top \end{bmatrix}^\top$. Set $\tau = 0$.

Step 2 Compute an estimate of $\chi(t_k)$ by simulating the auxiliary model:

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

based on the estimated parameters $\hat{\rho}^{(\tau)}$ of the previous iteration. Deduce the output terms $\{\hat{\chi}_{i,l}(t_k)\}_{i=1,l=0}^{n_{\mathrm{a}}, n_\alpha}$ as given in (15) using the model $\mathcal{M}_{\hat{\theta}^{(\tau)}}$.

Step 3 Compute the estimated filter:

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

along with the associated filtered signals $\{u_{j,l}^{\mathrm{f}}(t_k)\}_{j=0,l=0}^{n_{\mathrm{b}}, n_\beta}$, $y_{\mathrm{f}}(t_k)$ and $\{\chi_{i,l}^{\mathrm{f}}(t_k)\}_{i=1,l=0}^{n_{\mathrm{a}}, n_\alpha}$.

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

$$\hat{\varphi}_{\mathrm{f}}(t_k) = \left[-y_{\mathrm{f}}(t_{k-1}) \ \cdots \ -y_{\mathrm{f}}(t_{k-n_{\mathrm{a}}}) \ -\hat{\chi}_{1,1}^{\mathrm{f}}(t_k) \right.$$
$$\left. \cdots \ -\hat{\chi}_{n_{\mathrm{a}},n_\alpha}^{\mathrm{f}}(t_k) \ u_{0,0}^{\mathrm{f}}(t_k) \ \cdots \ u_{n_{\mathrm{b}},n_\beta}^{\mathrm{f}}(t_k) \right]^\top$$

$$\hat{\zeta}_{\mathrm{f}}(t_k) = \left[-\hat{\chi}_{\mathrm{f}}(t_{k-1}) \ \cdots \ -\hat{\chi}_{\mathrm{f}}(t_{k-n_{\mathrm{a}}}) \ -\hat{\chi}_{1,1}^{\mathrm{f}}(t_k) \right.$$
$$\left. \cdots \ -\hat{\chi}_{n_{\mathrm{a}},n_\alpha}^{\mathrm{f}}(t_k) \ u_{0,0}^{\mathrm{f}}(t_k) \ \cdots \ u_{n_{\mathrm{b}},n_\beta}^{\mathrm{f}}(t_k) \right]^\top$$

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

$$\hat{\rho}^{(\tau+1)}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(t_k) \hat{\varphi}_{\mathrm{f}}^\top(t_k) \right] \rho \right.$$
$$\left. - \left[\frac{1}{N} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(t_k) y_{\mathrm{f}}(t_k) \right] \right\|^2 \quad (25)$$

This results in the solution of the IV estimation equations:

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

where $\hat{\rho}^{(\tau+1)}(N)$ 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

$$\hat{v}(t_k) = y(t_k) - \hat{\chi}(t_k, \hat{\rho}^{(\tau)}). \quad (26)$$

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 [15]).

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.

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 1 can be skipped. Naturally, the LPV-SRIV is not statistically optimal for LPV-BJ models, however it still has a certain degree of robustness as it is shown in Section V.

V. SIMULATION EXAMPLE

The performance of the proposed method is compared to the existing methods in the literature, based on a representative simulation example.

A. Data generating system

The system taken into consideration is inspired by [3] and is mathematically described as follows:

$$\mathcal{S}_{\mathrm{o}} \begin{cases} A_{\mathrm{o}}(q, p_k) = 1 + a_1^{\mathrm{o}}(p_k)q^{-1} + a_2^{\mathrm{o}}(p_k)q^{-2} \\ B_{\mathrm{o}}(q, p_k) = b_0^{\mathrm{o}}(p_k)q^{-1} + b_1^{\mathrm{o}}(p_k)q^{-2} \\ H_{\mathrm{o}}(q) = \dfrac{1}{1 - q^{-1} + 0.2q^{-2}} \end{cases} \quad (27)$$

where $v(t_k) = H_{\mathrm{o}}(q)e(t_k)$ and

$$a_1^{\mathrm{o}}(p_k) = 1 - 0.5p_k - 0.1p_k^2 \quad (28\mathrm{a})$$
$$a_2^{\mathrm{o}}(p_k) = 0.5 - 0.7p_k - 0.1p_k^2 \quad (28\mathrm{b})$$
$$b_0^{\mathrm{o}}(p_k) = 0.5 - 0.4p_k + 0.01p_k^2 \quad (28\mathrm{c})$$
$$b_1^{\mathrm{o}}(p_k) = 0.2 - 0.3p_k - 0.02p_k^2 \quad (28\mathrm{d})$$

In the upcoming examples, the scheduling signal p is considered as a periodic function of time:

$$p_k = 0.5 \sin(0.35\pi k) + 0.5, \quad (29)$$

and $u(t_k)$ 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 $\mathcal{S}_{\mathrm{o}}$.

B. Model structures

In the sequel, the instrumental variable method presented in [3] named here *One Step Instrumental Variable* (OSIV) and the conventional *Least Square* (LS) method such as the one used in [2] are compared to the proposed methods. Both methods assume the following model structure:

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

where

$$a_1(p_k) = a_{1,0} + a_{1,1}p_k + a_{1,2}p_k^2 \quad (30\mathrm{a})$$
$$a_2(p_k) = a_{2,0} + a_{2,1}p_k + a_{2,2}p_k^2 \quad (30\mathrm{b})$$
$$b_0(p_k) = b_{0,0} + b_{0,1}p_k + b_{0,2}p_k^2 \quad (30\mathrm{c})$$
$$b_1(p_k) = b_{1,0} + b_{1,1}p_k + b_{1,2}p_k^2 \quad (30\mathrm{d})$$

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

$$\mathcal{M}_\theta^{\mathrm{LPV-RIV}} \begin{cases} A(p_k, q^{-1}, \rho) = 1 + a_1(p_k)q^{-1} + a_2(p_k)q^{-2} \\ B(p_k, q^{-1}, \rho) = b_0(p_k)q^{-1} + b_1(p_k)q^{-2} \\ H(p_k, q, \eta) = \dfrac{1}{1 + d_1 q^{-1} + d_2 q^{-2}} \end{cases}$$

with $a_1(p_k)$, $a_2(p_k)$, $b_0(p_k)$, $b_1(p_k)$ as given in (30a-d), while the LPV *Simplified Refined Instrumental Variable* method (LPV-SRIV) represents the case when $\mathcal{G}_{\mathrm{o}} \in \mathcal{G}$, $\mathcal{H}_{\mathrm{o}} \notin \mathcal{H}$ and assumes the following LPV-OE model:

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

C. Results

In this example, the robustness of the proposed and existing algorithms are investigated with respect to different *signal-to-noise ratios* (SNR). To provide representative results, a Monte-Carlo simulation of $N_{\mathrm{MC}} = 100$ runs with new noise realization is accomplished at different noise levels from 15dB down to 0dB.

TABLE I
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	10.8586
	VN		1.3353	179.4287	590.7869	11782
LPV-SRIV	BN		0.0072	0.0426	0.1775	0.2988
	VN		0.0149	0.0537	0.4425	0.4781
	MIN		22	22	25	30
LPV-RIV	BN		0.0068	0.0184	0.0408	0.1649
	VN		0.0063	0.0219	0.0696	0.2214
	MIN		31	30	30	32

With respect to the considered methods, Table I shows the norm of the bias and variance of the estimated parameter vector:

$$\text{Bias norm} = ||\rho_o - \bar{\mathbb{E}}(\hat{\rho})||_2 \qquad (31a)$$

$$\text{Variance norm} = ||\bar{\mathbb{E}}(\hat{\rho} - \bar{\mathbb{E}}(\hat{\rho}))||_2 \qquad (31b)$$

where $\bar{\mathbb{E}}$ is the mean operator over the Monte-Carlo simulation and $||.||_2$ is the $\mathcal{L}_2$ norm. The table also displays the mean number of iterations (MIN) the algorithms needed to converge to the estimated parameter vector.

It can be seen from Table I that the IV methods are unbiased according to the theoretical results whereas LS method shows a strong bias. 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. 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. The suboptimal LPV-SRIV methods 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 in close range 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).

VI. CONCLUSION

It has been shown that there is a lack of efficient methods in the literature capable of handling the estimation of LPV models with a noise model different than an ARX structure. The underlying reason is that in the LPV case, the conventional formulation of least squares estimation cannot lead to statistically optimal parameter estimates. As a solution, the LPV identification problem is reformulated and a method to estimate efficiently LPV-BJ models 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 illustrated that the proposed procedure is robust to noise and outperforms the existing methods. As continuation of the presented work, extensions of the method to closed-loop and continuous-time LPV system identification are intended.

REFERENCES

[1] H. Abbas and H. Werner. An instrumental variable technique for open-loop and closed-loop identification of input-ouput LPV models. In *Proceedings of the European Control Conference 2009*, pages 2646–2651, Budapest, Hungary, 23-26 August 2009.

[2] B. Bamieh and L. Giarré. Identification of linear parameter-varying models. *International Journal of Robust and Nonlinear Control*, 12:841–853, 2002.

[3] 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.

[4] 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.

[5] 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.

[6] L. Ljung. *System identification : theory for the user - Second Edition*. Prentice-Hall, 1999.

[7] 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.

[8] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer-Verlag, New York, 1983.

[9] P. Stoica and M. Jansson. MIMO system identification: state-space and subspace approximations versus transfer function and instrumental variables. *IEEE Transaction on Signal Processing*, 48(11):3087 – 3099, 2000.

[10] R. Tóth. *Modeling and Identification of Linear Parameter-Varying Systems: An Orthonormal Basis Function Approach*. PhD thesis, Delft University of Technology, 2008.

[11] 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*, pages 1328–1333, St. Malo, France, 6-8 July 2009.

[12] 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.

[13] 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.

[14] P. C. Young. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, 1984.

[15] 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.

[16] P. C. Young, H. Garnier, and M. Gilson. *Identification of continuous-time models from sampled data*, chapter Refined instrumental variable identification of continous-time hybrid Box-Jenkins models. Springer-Verlag, London, 2008.