Direct Identification of Continuous-Time LPV Input/Output Models

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

Abstract

Controllers in the linear parameter-varying (LPV) framework are commonly designed in continuous-time (CT) requiring accurate and low-order CT models of the system. However, identification of continuous-time LPV models is largely unsolved, representing a gap between the available LPV identification methods and the needs of control synthesis. In order to bridge this gap, direct identification of CT LPV systems in an input-output setting is investigated, focusing on the case when the noise part of the data generating system is an additive discrete-time colored noise process. To provide consistent model parameter estimates in this setting, a refined instrumental variable (IV) approach is proposed and its properties are analyzed based on the prediction-error framework. The benefits of the introduced direct CT-IV approach over identification in the discrete-time case are demonstrated through a representative simulation example inspired by the Rao-Garnier benchmark.

Index Terms

Continuous-time models, LPV models, system identification, refined instrumental variable, Box–Jenkins models, input/output.

I. Introduction

The framework of *linear parameter-varying* (LPV) systems was introduced in the 1990s with the purpose to handle in a simple but efficient way the often nonlinear or time-varying nature of systems encountered in practice. The LPV system class forms an intermediate step between *linear time-invariant* (LTI) systems and nonlinear/time-varying plants as the signal relations in LPV systems 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* $p : \mathbb{Z} \to \mathbb{P}$. Here the compact set $\mathbb{P} \subset \mathbb{R}^{n_\mathbb{P}}$ denotes the *scheduling space*. The scheduling variable p represents

V. Laurain, M. Gilson and H. Garnier are with the Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, BP 70239, 54506 Vandoeuvre-les-Nancy Cedex, France. R. Tóth is with the Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands, email: r.toth@tudelft.nl.

the nonlinear or time-varying nature of the modeled dynamics, *e.g.* p describes the changes in the operating conditions of the plant, time-variation of the system dynamics or external effects like temperature changes. This LPV modeling concept allows for a wide representation capability of physical processes, but the real practical significance of the LPV framework lays in its well worked out and industrially reputed control synthesis approaches, *e.g.* [1]–[3], that have led to many successful applications of LPV control in practice [4]–[7].

However a major drawback of the LPV framework today is that, despite the advances of the LPV control field, identification of such systems is not well developed as the current methods are unable to support practical control design. Commonly LPV controllers are synthesized in *continuous time* (CT) as stability and performance requirements of the closed loop behavior can be more conveniently expressed in CT, like in a mixed-sensitivity setting [8]. Therefore, the current design tools focus on continuous-time LPV controller synthesis requiring accurate and low-order CT models of system. However, LPV identification methods are almost exclusively developed for *discrete-time* (DT) (for a recent survey see [9]), as in this setting it is much easier to handle the estimation of parameter-varying dynamics. The only available CT method (to the authors' knowledge) uses a local, or a so-called multiple-model type of approach, where CT-LTI models of the LPV system are estimated for constant trajectories of p, i.e. at certain operating operating points, by using a frequency domain method and then the resulting models are interpolated over $\mathbb{P}$. However such an approach is unable to capture the global behavior of the system (limited number of identified CT-LTI models, no information about transient behavior from one operating condition to the other) and is affected by interpolation problems described in [10], [11]. Thus, the absence of CT methods obviously represents a gap between the available identification approaches and the needs of LPV control synthesis. As 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, there is a growing need of the LPV framework for efficient identification methods that directly deliver reliable CT models.

In practice, CT systems can only be identified based on sampled measured data records. Thus in general, for delivering a CT model estimate, the available approaches in system identification can be categorized as follows [12]:

- **Indirect approaches:** These methods involve the identification of a DT model in a completely DT setting, followed by the transformation of the DT model estimate to a CT form.

- **Direct approaches:** The methods formulate the identification of the CT model directly based on samples of the measured CT signals.

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 [13], [14]) 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 underlaying problems, consider the following simple CT-LPV model

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t) + p(t)y(t) = bu(t), \tag{1}$$

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: $\frac{\mathrm{d}}{\mathrm{d}t}y(t_k) \approx \frac{y(t_k)-y(t_{k-1})}{\mathtt{T}_\mathrm{s}}$ with $\mathtt{T}_\mathrm{s} > 0$ the sampling period, (1) transforms into:

$$y(t_k) = \frac{1}{1+\mathtt{T}_\mathrm{s}p(t_k)}y(t_{k-1}) + \frac{b\mathtt{T}_\mathrm{s}}{1+\mathtt{T}_\mathrm{s}p(t_k)}u(t_k). \tag{2}$$

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

$$y(t_k) = (1 - \mathtt{T}_\mathrm{s}p(t_{k-1}))y(t_{k-1}) + b\mathtt{T}_\mathrm{s}u(t_{k-1}). \tag{3}$$

This discretized model has only one p-dependent coefficient and the linearity of the dependence is preserved, however now the model equation is dependent on $p(t_{k-1})$ instead of $p(t_k)$. 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 [15]). Furthermore, it is well known in numerical analysis that the forward Euler approximation is more sensitive for the choice of $\mathtt{T}_\mathrm{s}$ in terms of numerical stability than the backward Euler approximation [16]. This means that (3) requires much faster sampling rate than (2) to give a stable approximation of the system and it is more sensitive to parameter uncertainties which rises problems if (3) is used for estimation. Consequently it can be concluded 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 underlaying problems for which there are no general theoretical solutions available.

Unlike an indirect approach, a direct solution offers a way to efficiently overcome the previous problems. Due to the recent technological developments of sampling instruments in terms of achievable sampling rate, use of direct CT approaches in identification has recently regained interest showing a better performance than indirect approaches for both linear and nonlinear models, see *e.g.* [12], [17]–[20]. An exhaustive review of direct estimation methods can be found in [12], [21], [22]. Among the available identification approaches for CT-LTI *input-output* (IO) models, the interest for *instrumental variable* (IV) methods has been growing in the last years [20], [23], [24]. 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 [18], [25]) 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 CT [18], [22], direct nonlinear CT [19] or closed-loop CT identification [26], [27] and lead to optimal estimates in the LTI case if the system belongs to the model set defined.

In this paper we aim to provide the very first step towards bridging the existing gap between LPV control and identification via the introduction of a direct CT identification approach that benefits from the properties of IV methods. It was shown recently in [28] that in order to minimize the classical prediction error for DT-LPV models, a *Multiple Input Single Output* (MISO)-LTI reformulation of the data-generating system is needed. Based on a similar reformulation, the prediction error minimization problem can be clearly stated in the present CT case. Furthermore, the proposed approach extends the recent results from the CT-LTI identification framework using an IV method to face the direct CT-LPV identification problem stated here. The resulting approach not only provides the very first global LPV identification method that is able to provide consistent estimates of LPV-IO models in continuous-time, but it is also applicable in case of colored output noise and has a low computational load. Furthermore, it opens the possibility for closed-loop CT-LPV identification.

The paper is organized as follows: in Section II, the general class of CT-LPV systems in an IO representation form is introduced pointing out the main difficulties of this model class. Additionally, a reformulation of the dynamical description of LPV data generating plants in the considered setting is developed which makes possible the extension of CT LTI-IV methods to the LPV framework. In Section III, CT LPV-IV methods are described and analyzed, while their performance are illustrated in Section IV through a representative simulation example inspired by the Rao-Garnier benchmark. Finally in Section V, 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 CT LPV system described by the following equations

$$\mathcal{S}_{\mathrm{o}} \begin{cases} A_{\mathrm{o}}(p_t, \mathscr{d})\chi_{\mathrm{o}}(t) = B_{\mathrm{o}}(p_t, \mathscr{d})u(t), \\ y(t) = \chi_{\mathrm{o}}(t) + v_{\mathrm{o}}(t), \end{cases} \tag{4}$$

where $\mathscr{d}$ denotes the differentiation operator w.r.t. time, i.e. $\mathscr{d} = \frac{\mathrm{d}}{\mathrm{d}t}$, $p : \mathbb{R} \to \mathbb{P}$ 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 it is uncorrelated to p. A_{o} and B_{o} are polynomials in $\mathscr{d}$ with coefficients a_i^{o} and b_i^{o} that are meromorphic functions [1] of p with no singularity on $\mathbb{P}$

$$A_{\mathrm{o}}(p_t, \mathscr{d}) = \mathscr{d}^{n_a} + \sum_{i=1}^{n_{\mathrm{a}}} a_i^{\mathrm{o}}(p_t)\mathscr{d}^{n_a - i} \qquad \text{and} \qquad B_{\mathrm{o}}(p_t, \mathscr{d}) = \sum_{j=0}^{n_{\mathrm{b}}} b_j^{\mathrm{o}}(p_t)\mathscr{d}^{n_b - j}. \tag{5}$$

Note that a_i^{o} and b_j^{o} are functions of p at time t, which is called *static dependence*. In LPV system theory, a more general p-dependence of coefficients than static is required to establish equivalence of representations. In particular, the coefficients a_i^{o} and b_j^{o} need to depend also on the time derivatives of p, which is called *dynamic dependence* [9]. In order to simplify the upcoming discussion, we restrict our attention to static dependence. Nevertheless, the established results hold also in the case of dynamic dependence of (4) and of the proposed model structure.

In terms of identification we can assume that sampled measurements of (y, p, u) are available with a sampling period $\mathtt{T}_{\mathrm{s}} > 0$. Hence, we will denote the discrete-time samples of these signals as $u(t_k) = u(k\mathtt{T}_{\mathrm{s}})$, where $k \in \mathbb{Z}$. The basic idea to solve the noisy CT modeling problem of (4) is to assume that the CT noise process $v_{\mathrm{o}}(t)$ can be considered at the sampling instances as a DT noise process filtered by a DT transfer function. In this paper, a practically general case is considered where the colored noise associated with the sampled output measurement

[1] A function f is called meromorphic if $f = \frac{g}{h}$ where g, h are holomorphic (analytic) functions and h is not the zero function.

$y(t_k)$ is assumed to have a rational spectral density which might has no relation to the actual process dynamics of $\mathcal{S}_\mathrm{o}$. As a preliminary step towards the case of a p-dependent noise, it is also assumed that this spectral density is not dependent on p, like in case of a measurement noise. Therefore, v_o is represented by a discrete-time *autoregressive-moving-average* (ARMA) model

$$v_\mathrm{o}(t_k) = H_\mathrm{o}(q)e_\mathrm{o}(t_k) = \frac{C_\mathrm{o}(q^{-1})}{D_\mathrm{o}(q^{-1})}e_\mathrm{o}(t_k), \tag{6}$$

where $e_\mathrm{o}(t_k)$ is a DT zero-mean white noise process, q^{-1} is the backward time shift operator, *i.e.* $q^{-i}u(t_k) = u(t_{k-i})$. C_o and D_o are monic polynomials having constant coefficients. This formulation of the noise model (6) avoids the rather difficult mathematical problem of treating sampled CT random process in terms of a filtered piece-wise constant CT noise source (see [29], [30]). In terms of (6), $y(t_k)$ can be written as

$$y(t_k) = \chi_\mathrm{o}(t_k) + v_\mathrm{o}(t_k), \tag{7}$$

which corresponds to a so called hybrid Box-Jenkins system concept already used in CT identification of LTI systems (see [29], [30], [19]). Furthermore, in terms of (6), exactly the same noise assumption is made as in the classical DT Box-Jenkins models (see [31]).

B. Model structure considered

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

$$\mathcal{G}_\rho : \left(A(p_t, d, \rho), B(p_t, d, \rho)\right), \tag{8}$$

where the p-dependent polynomials A and B given as

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

are parameterized as

$$a_i(p_t) \;=\; a_{i,0} + \sum_{l=1}^{n_\alpha} a_{i,l} f_l(p_t), \quad i = 1, \ldots, n_\mathrm{a}, \tag{9a}$$

$$b_j(p_t) \;=\; b_{j,0} + \sum_{l=1}^{n_\beta} b_{j,l} g_l(p_t), \quad j = 0, \ldots, n_\mathrm{b}, \tag{9b}$$

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 (they can be chosen for example as linearly independent functions on $\mathbb{P}$). The associated model parameters are stacked column wise:

$$\rho = [\ \mathtt{a}_1 \ \ \ldots \ \ \mathtt{a}_{n_\mathrm{a}} \ \ \mathtt{b}_0 \ \ \ldots \ \ \mathtt{b}_{n_\mathrm{b}}\ ]^\top \in \mathbb{R}^{n_\rho}, \tag{10}$$

where $\mathtt{a}_i = [\ a_{i,0} \ \ a_{i,1} \ \ \ldots \ \ a_{i,n_\alpha}\ ] \in \mathbb{R}^{n_\alpha+1}$, $\mathtt{b}_j = [\ b_{j,0} \ \ b_{j,1} \ \ \ldots \ \ b_{j,n_\beta}\ ] \in \mathbb{R}^{n_\beta+1}$ and $n_\rho = n_\mathrm{a}(n_\alpha+1)+(n_\mathrm{b}+1)(n_\beta+1)$. Introduce also $\mathscr{G} = \{\mathcal{G}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$, as the collection of all process models in the form of (8).

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

$$\mathcal{H}_\eta : (H(q,\eta)), \tag{11}$$

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

$$H(q,\eta) = \frac{C(q^{-1},\eta)}{D(q^{-1},\eta)} = \frac{1 + c_1 q^{-1} + \ldots + c_{n_\mathrm{c}} q^{-n_\mathrm{c}}}{1 + d_1 q^{-1} + \ldots + d_{n_\mathrm{d}} q^{-n_\mathrm{d}}}. \tag{12}$$

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

$$\eta = [\ c_1 \ \ \ldots \ \ c_{n_\mathrm{c}} \ \ d_1 \ \ \ldots \ \ d_{n_\mathrm{d}}\ ]^\top \in \mathbb{R}^{n_\eta}, \tag{13}$$

where $n_\eta = n_\mathrm{c} + n_\mathrm{d}$. Additionally, denote $\mathscr{H} = \{\mathcal{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$, the collection of all noise models in the form of (11).

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

$$\mathcal{M}_\theta \begin{cases} A(p_k, d, \rho)\chi(t) = B(p_k, d, \rho)u(t) \\ v(t_k) = \dfrac{C(q^{-1},\eta)}{D(q^{-1},\eta)} e(t_k) \\ y(t_k) = \chi(t_k) + v(t_k) \end{cases} \tag{14}$$

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

$$\mathscr{M} = \left\{ (\mathcal{G}_\rho, \mathcal{H}_\eta) \ \mid \ \mathrm{col}(\rho,\eta) = \theta \in \mathbb{R}^{n_\rho + n_\eta} \right\}. \tag{15}$$

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. 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 $\mathscr{M}$, 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 $\mathscr{d}$ is not commutative over the p-dependent coefficients [9], meaning that

$$\mathscr{d}\left(B(p,\mathscr{d})u(t)\right) = \sum_{j=0}^{n_\mathrm{b}} \left(\frac{\partial b_j}{\partial p}(p_t)\mathscr{d}p_t\right)\mathscr{d}^{\,n_b-j}u(t) + B(p,\mathscr{d})\mathscr{d}u(t) \tag{16}$$

which is not equal to $B(p,\mathscr{d})\mathscr{d}u(t)$.

In the DT case, in order 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 [9], [28]. Considering the CT case, no IRR has been developed yet and thus the same concept cannot be used to define the predictors.

1) System reformulation and prediction error: Under the assumed noise conditions and for DT-LPV-IO models, it was shown in [28] that an efficient way to deal with the LPV identification problem in the PEM framework is to express the LPV system as a MISO LTI model. Therefore, based on the same idea, if the system belongs to the model set defined and with a deterministic p signal, it is possible to express the CT LPV system as a CT MISO LTI system by rewriting the signal relations of (4) as

$$\underbrace{\chi_\mathrm{o}^{(n_a)}(t) + \sum_{i=1}^{n_\mathrm{a}} a_{i,0}^\mathrm{o}\chi_\mathrm{o}^{(n_a-i)}(t)}_{F_\mathrm{o}(\mathscr{d})\chi_\mathrm{o}(t)} + \sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{n_\alpha} a_{i,l}^\mathrm{o}\underbrace{f_l(p(t))\chi_\mathrm{o}^{(n_a-i)}(t)}_{\chi_{i,l}^\mathrm{o}(t)} = \sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{n_\beta} b_{j,l}^\mathrm{o}\underbrace{g_l(p(t))u^{(n_b-j)}(t)}_{u_{j,l}(t)} \tag{17}$$

where $g_0(t)=1$ and the superscript (n) for a signal, like $u^{(n)}$, denotes the n^{th} time-derivative of the signal, *e.g.* $u^{(n)}(t) = \mathscr{d}^{\,n}u(t)$. Furthermore, $F(\mathscr{d}) = \mathscr{d}^{\,n_a} + \sum_{i=1}^{n_\mathrm{a}} a_{i,0}\mathscr{d}^{\,n_a-i}$ while $u^{(n)}(t_k)$ represents the value of the signal $u^{(n)}(t)$ sampled at time instance t_k.

Note that in this way, the time variation of the coefficients is transposed onto the signals

$$\chi_{i,l}^\mathrm{o}(t) = f_l(p(t))\chi_\mathrm{o}^{(n_a-i)}(t), \qquad \{i,l\}\in\{1\ldots n_\mathrm{a}, 1\ldots n_\alpha\}, \tag{18a}$$

$$u_{j,l}(t) = g_l(p(t))u^{(n_b-j)}(t), \qquad \{j,l\}\in\{1\ldots n_\mathrm{b}, 1\ldots n_\alpha\}. \tag{18b}$$

Therefore, the process part of the LPV-BJ model is rewritten as a *Multiple-Input Single-Output* (MISO) system with $(n_\mathrm{b}+1)(n_\beta+1)+n_\mathrm{a}n_\alpha$ inputs $\{\chi_{i,l}^\mathrm{o}\}_{i=1,l=1}^{n_\mathrm{a},n_\alpha}$ and $\{u_{j,l}\}_{j=0,l=0}^{n_\mathrm{b},n_\beta}$. By using (17), (14) can be rewritten in terms of the sampled output signal $y(t_k)$ as

$$y(t_k) = \underbrace{-\left(\sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{n_\alpha}\frac{a_{i,l}^\mathrm{o}}{F_\mathrm{o}(\mathscr{d})}\chi_{i,l}^\mathrm{o}\right)(t_k) + \left(\sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{n_\beta}\frac{b_{j,l}^\mathrm{o}}{F_\mathrm{o}(\mathscr{d})}u_{k,j}\right)(t_k)}_{G_\mathrm{o}(\chi_\mathrm{o},u,t_k)} + v_\mathrm{o}(t_k), \tag{19}$$

which is a sampled LTI representation of the system defined in (4).

Given the assumption that $v_{\mathrm{o}}(t_k) = H_{\mathrm{o}}(q)e_{\mathrm{o}}(t_k)$ and $C(q^{-1})$ is a monic polynomial, (6) can be rewritten in the form

$$v_{\mathrm{o}}(t_k) = e_{\mathrm{o}}(t_k) + \sum_{i=1}^{\infty} h_i e_{\mathrm{o}}(t_{k-i}). \tag{20}$$

This shows that the knowledge of $\{v_{\mathrm{o}}(\tau)\}_{\tau \le t_{k-1}}$ implies the knowledge of $\{e_{\mathrm{o}}(\tau)\}_{\tau \le t_{k-1}}$. Therefore, by using the traditional approach [31], the prediction of $v_{\mathrm{o}}(t_k)$ is considered as the conditional expectation of $v_{\mathrm{o}}(t_k)$ based on $\{e_{\mathrm{o}}(\tau)\}_{\tau \le t_{k-1}}$ which is according to (20):

$$\hat{v}(t_k) = \hat{v}(t_k \,|\, t_{k-1}) = \mathbb{E}\{v_{\mathrm{o}}(t_k) \mid \{e_{\mathrm{o}}(\tau)\}_{\tau \le k-1}\} = \sum_{i=1}^{\infty} h_i e_{\mathrm{o}}(t_{k-i}), \tag{21}$$

where $\mathbb{E}$ is the expectation operator. Assuming that H_{o} has a stable inverse such that $e_{\mathrm{o}}(t_k) = H_{\mathrm{o}}^{-1}(q)v_{\mathrm{o}}(t_k)$, then the classical one-step-ahead predictor can be given as [31]

$$\hat{v}(t_k) = v(t_k \,|\, t_{k-1}) = v_{\mathrm{o}}(t_k) - e_{\mathrm{o}}(t_k) = (1 - H_{\mathrm{o}}^{-1}(q))v_{\mathrm{o}}(t_k). \tag{22}$$

Consequently, for the considered LPV system formulated as in (19), the one-step-ahead predictor of $y(t_k)$ (defined as the conditional expectation $\hat{y}(t_k \,|\, t_{k-1})$ on $\{y(t_i)\}_{i \le k-1}$, $\{u(t_i), \chi_{\mathrm{o}}(t_i)\}_{i \le k}$) is given by

$$\begin{aligned}
\hat{y}(t_k) =& H_{\mathrm{o}}^{-1}(q)G_{\mathrm{o}}(\chi_{\mathrm{o}}, u, t_k) + (1 - H_{\mathrm{o}}^{-1}(q))(y(t_k)),\\
\hat{y}(t_k) =& H_{\mathrm{o}}^{-1}(q)\Bigg(-\Bigg(\sum_{i=1}^{n_{\mathrm{a}}}\sum_{l=1}^{n_{\alpha}} \frac{a_{i,l}^{\mathrm{o}}}{F_{\mathrm{o}}(\mathrm{d})}\chi_{i,l}^{\mathrm{o}}\Bigg)(t_k) + \Bigg(\sum_{j=0}^{n_{\mathrm{b}}}\sum_{l=0}^{n_{\beta}} \frac{b_{j,l}^{\mathrm{o}}}{F_{\mathrm{o}}(\mathrm{d})}u_{k,j}\Bigg)(t_k)\Bigg)\\
&+ (1 - H_{\mathrm{o}}^{-1}(q))(y(t_k)).
\end{aligned} \tag{23}$$

2) Prediction Error Model: Using the same idea as in Subsection II-C1, the LPV model from (14) can also be expressed in a MISO LTI form [28]:

$$y_\theta(t_k) = -\Bigg(\sum_{i=1}^{n_{\mathrm{a}}}\sum_{l=1}^{n_{\alpha}} \frac{a_{i,l}}{F(\mathrm{d},\rho)}\chi_{i,l}\Bigg)(t_k) + \Bigg(\sum_{j=0}^{n_{\mathrm{b}}}\sum_{l=0}^{n_{\beta}} \frac{b_{j,l}}{F(\mathrm{d},\rho)}u_{k,j}\Bigg)(t_k) + H(q,\eta)e(t_k). \tag{24}$$

Therefore, similarly to the LTI case, the *one-step-ahead prediction error* can be expressed and defined as [31]:

$$\varepsilon_\theta(t_k) = y(t_k) - \hat{y}_\theta(t_k), \tag{25}$$

where $\hat{y}_\theta(t_k)$ is the *one-step-ahead predictor* based on the model (14) written as in (24) and is defined as (see (23)):

$$\hat{y}_\theta(t_k) = H^{-1}(q,\eta)\left(-\left(\sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{n_\alpha}\frac{a_{i,l}}{F(\mathscr{d},\rho)}\chi_{i,l}\right)(t_k) + \left(\sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{n_\beta}\frac{b_{j,l}}{F(\mathscr{d},\rho)}u_{k,j}\right)(t_k)\right) + \left(1 - H^{-1}(q,\eta)\right)y(t_k). \quad (26)$$

3) Prediction error minimization: Denote $\mathcal{D}_N = \{y(t_k), u(t_k), p(t_k)\}_{k=1}^N$ a data sequence of $\mathcal{S}_\mathrm{o}$. Then to provide an estimate of θ based on the minimization of ε_θ, an identification criterion $W(\mathcal{D}_N,\theta)$ can be introduced, like the *least square* criterion

$$W(\mathcal{D}_N,\theta) = \frac{1}{N}\sum_{k=1}^{N}\varepsilon_\theta^2(t_k), \quad (27)$$

such that the parameter estimate is

$$\hat{\theta}_N = \arg\min_{\theta\in\mathbb{R}^{n_\rho+n_\eta}} W(\mathcal{D}_N,\theta). \quad (28)$$

4) CT filtering and sampled data: The hybrid representation of the model (14) is a combined operation of CT filtering and DT filtering which implicitly appears in the formulation of (26). In order to clearly define the coexistence of DT and CT filtering in (26), a detailed investigation and discussion about the assumptions and the structure of the model are needed. In this paper, we considered the practically feasible situation such that only sampled measurements of the CT signals (y,p,u) are available. In order to apply a CT filter on sampled data one can either interpolate the samples to obtain a continuous-time 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 continuous-time systems. For simulation purposes, DT approximation of the system can efficiently be dealt with by using powerful numerical algorithms available [16]. 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 [32], which has also been shown in case of LPV systems with static dependence [9]. This provides the hypothesis, also used in [29], [30], 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 $\mathtt{T}_\mathrm{s}k$. Therefore, under these inter-sampling conditions, the following operation is well-defined [22]:

$$\left(F(\mathcal{d})y\right)(t_k) = F(\mathcal{d})y(t_k), \tag{29}$$

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, in terms of (26), the coefficients of the parsimonious CT model.

D. Identification problem statement

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

Problem 1: Given a CT-LPV data generating system $\mathcal{S}_\mathrm{o}$ defined as in (4) and a data set $\mathcal{D}_N$ collected from $\mathcal{S}_\mathrm{o}$. Based on the hybrid LPV-BJ model structure $\mathcal{M}_\theta$ defined by (14), 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 (9a-b) $\{f_l\}_{l=1}^{n_\alpha}$ and $\{g_l\}_{l=1}^{n_\beta}$ are chosen such that $(\mathcal{G}_\mathrm{o}, \mathcal{H}_\mathrm{o})$ is identifiable.

A3 $u(t_k)$ and $p(t_k)$ are 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 globally BIBO stable, i.e. for any trajectory of $p : \mathbb{R} \rightarrow \mathbb{P}$ and any bounded input signal u, the output of $\mathcal{S}_\mathrm{o}$ is bounded [9].

III. Refined Instrumental variable for LPV systems

Based on the MISO-LTI formulation (24), it becomes theoretically possible to achieve optimal PEM using linear regression [28]. This allows to extend the *Refined Instrumental Variable* (RIV) approach of the LTI framework to provide an efficient way of identifying hybrid LPV-BJ models.

A. Linear Regression for CT LPV-BJ models

Using the LTI model (14) reformulated as in (24), $y(t_k)$ can be written in the regression form:

$$y^{(n_a)}(t_k) = \varphi^\top(t_k)\rho + \tilde{v}(t_k), \tag{30}$$

where,

$$\begin{aligned}
\varphi(t_k) &= \left[\, -y^{(n_a-1)}(t_k) \;\; \dots \;\; -y(t_k) \;\; -\chi_{1,1}(t_k) \;\; \dots \;\; -\chi_{n_\mathrm{a},n_\alpha}(t_k) \;\; u_{0,0}(t_k) \;\; \dots \;\; u_{n_\mathrm{b},n_\beta}(t_k) \,\right]^\top \\
\rho &= \left[\, a_{1,0} \;\; \dots \;\; a_{n_\mathrm{a},0} \;\; a_{1,1} \;\; \dots \;\; a_{n_\mathrm{a},n_\alpha} \;\; b_{0,0} \;\; \dots \;\; b_{n_\mathrm{b},n_\beta} \,\right]^\top
\end{aligned}$$

and $\tilde{v}(t_k) = F(d, \rho)v(t_k)$. The extended regressor in (30) contains the noise-free output terms $\{\chi_{i,k}\}$. Therefore, by momentary assuming that $\{\chi_{i,l}(t_k)\}_{i=1,l=0}^{n_\mathrm{a},n_\alpha}$ are known *a priori*, the prediction error $\varepsilon_\theta(t_k)$ for (30) is given in terms of (25) as:

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1},\eta)}{C(q^{-1},\eta)}\left[\frac{1}{F(d,\rho)}\left(F(d,\rho)y(t_k) - \Big[-\sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{n_\alpha} a_{i,l}\chi_{i,l}(t_k) + \sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{n_\mathrm{b}} b_{j,l}u_{j,l}(t_k)\Big]\right)\right]. \tag{31}$$

In the given context, the filters $\frac{D(q^{-1},\eta)}{C(q^{-1},\eta)}$ and $F(d,\rho)$ in (31) commute (see Section II-C4). The later allows us to rewrite *the one-step-ahead prediction error* (31) associated with (30) as

$$\varepsilon_\theta(t_k) = \left(F(d,\rho)y_\mathrm{f}\right)(t_k) - \sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{n_\alpha} a_{i,l}\chi^\mathrm{f}_{i,l}(t_k) + \sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{n_\beta} b_{j,l}u^\mathrm{f}_{k,j}(t_k), \tag{32}$$

where $y_\mathrm{f}(t_k)$, $u^\mathrm{f}_{j,l}(t_k)$ and $\chi^\mathrm{f}_{i,l}(t_k)$ represent the outputs of a hybrid prefiltering operation, involving the CT and DT filters (see [33]):

$$Q_\mathrm{c}(d,\rho) = \frac{1}{F(d,\rho)} \quad \text{and} \quad Q_\mathrm{d}(q^{-1},\eta) = \frac{D(q^{-1},\eta)}{C(q^{-1},\eta)}. \tag{33}$$

In other words:

$$y_\mathrm{f}(t_k) = \frac{D(q^{-1},\eta)}{C(q^{-1},\eta)}\left[\left(\frac{1}{F(d,\rho)}y\right)(t_k)\right]. \tag{34}$$

Based on (32), the associated linear-in-the-parameters model takes the form [33]:

$$y_\mathrm{f}^{(n_a)}(t_k) = \varphi_\mathrm{f}^\top(t_k)\rho + \tilde{v}_\mathrm{f}(t_k), \tag{35}$$

where

$$\begin{aligned}
\varphi_\mathrm{f}(t_k) &= [\, -y_\mathrm{f}^{(n_a-1)}(t_k) \ \ldots \ -y_\mathrm{f}(t_k) \ -\chi^\mathrm{f}_{1,1}(t_k) \ \ldots \ -\chi^\mathrm{f}_{n_\mathrm{a},n_\alpha}(t_k) \ u^\mathrm{f}_{0,0}(t_k) \ \ldots \ u^\mathrm{f}_{n_\mathrm{b},n_\beta}(t_k)]^\top \\
\tilde{v}_\mathrm{f}(t_k) &= Q_\mathrm{d}(q^{-1},\eta)Q_\mathrm{c}(d,\rho)\tilde{v}(t_k) = e(t_k).
\end{aligned}$$

B. The refined instrumental variable approach

Under the assumption that both the inverse noise model $Q_\mathrm{d}(q^{-1},\eta)$ and the CT filter $Q_\mathrm{c}(d,\rho)$ and consequently $\{\chi_{i,l}(t_k)\}_{i=1,l=0}^{n_\mathrm{a},n_\alpha}$ are known *a priori*, traditional parametric estimation methods from the LTI framework could provide efficient estimates of ρ and η. However, in a practical situation, $Q_\mathrm{d}(q^{-1},\eta)$ and $Q_\mathrm{c}(d,\rho)$ are unknown and need to be estimated as well.

Furthermore, it is important to notice here that the regressors in (35) and (30) contain some time-derivatives of y and u which, in the assumed framework considering sampled data, can only be approximated. It is well-known that the approximation of derivatives requires a low pass filtering of y and u. The most commonly used filters for this purpose are Poisson's filters, or state-variable filters [21]. The drawback of these filters is that they require the choice of a design variable. However, in the proposed approach $F(d, \rho)$ achieves this stable low-pass filtering directly. Therefore, it is a particular strength of the presented reformulation (35) is that the estimated filter $F(d, \rho)$ is not only used for the minimization of the prediction error but it also provides the filtering for the approximation of the time derivatives. In order to estimate the parameter vector in (35) without the prior knowledge of $Q_\mathrm{d}(q^{-1}, \eta)$ and $Q_\mathrm{c}(d, \rho)$, the *refined instrumental variable* (RIV) method is proposed due to the following reasons:

- RIV methods lead to optimal estimates in the LTI case if $\mathcal{S}_\mathrm{o} \in \mathcal{M}$ (see [23], [20], [33]).
- In a practical identification scenario, $\mathcal{G}_\mathrm{o} \in \mathcal{G}$ might be fulfilled due to first principle or expert's knowledge, however, it is commonly fair to assume that $H_\mathrm{o} \notin \mathcal{H}$. In such case, RIV methods have the advantage of providing consistent estimates whereas methods such as the extended LS are biased and more advanced PEM methods need robust initialization [25].
- The RIV algorithm has been successfully used for models with similar hybrid structure, like in the case of linear models [17], [20] and nonlinear ones [19].

Aiming at the extension of the RIV approach for the estimation of hybrid LPV-BJ models, consider the relationship between the process input and output signals as in (30). Based on this form, the extended-IV estimate is given as [20]:

$$\hat{\rho}_{\mathtt{XIV}}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N} \sum_{k=1}^{N} \zeta_\mathrm{f}(t_k) \varphi_\mathrm{f}^\top(t_k) \right] \rho - \left[\frac{1}{N} \sum_{t=1}^{N} \zeta_\mathrm{f}(t_k) y_\mathrm{f}^{(n_a)}(t_k) \right] \right\|_W^2 ,$$

where $\zeta(t_k)$ is the instrument, $\|x\|_W^2 = x^T W x$, with W a positive definite weighting matrix and the filtered variables ζ_f, φ_f and y_f are constructed using a stable prefilter. If $G_\mathrm{o} \in \mathcal{G}$, the extended-IV estimate is consistent under the following two conditions[2]:

C1 $\quad \bar{\mathbb{E}}\{\zeta_\mathrm{f}(t_k) \varphi_\mathrm{f}^\top(t_k)\}$ is full column rank.

C2 $\quad \bar{\mathbb{E}}\{\zeta_\mathrm{f}(t_k) \tilde{v}_\mathrm{f}(t_k)\} = 0$.

Moreover it has been shown in [23], [24] and [20] that the minimum variance estimator can be achieved if:

C3 $\quad W = I$.

[2]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 [31].

C4 ζ is chosen as the noise-free version of the extended regressor in (30) and is therefore defined in the present LPV case as:

$$\zeta(t_k)=\Big[-\chi^{(n_\mathrm{a}-1)}(t_k) \;\ldots -\chi(t_k) \; -\chi_{1,1}(t_k) \;\ldots -\chi_{n_\mathrm{a},n_\alpha}(t_k) \; u_{0,0}(t_k) \;\ldots\; u_{n_\mathrm{b},n_\beta}(t_k)\Big]^\top$$

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 The used hybrid filter is chosen as the filter chain (33).

While conditions C1, C2, C3 and C5 are quite straight-forward to fulfill (see [23], [24]), both the construction of a suitable instrument that fulfills C4 and of an optimal filter fulfilling C6 are not trivial in practice. The RIV algorithm 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. Thus, the RIV is a suitable method to i) efficiently estimate ρ in (35) when $S_\mathrm{o} \in \mathcal{M}$ and ii) consistently estimate ρ in a practical situation when $H_\mathrm{o} \notin \mathcal{H}$.

Nonetheless, it has to be added that even in the CT LTI case, the convergence of the iterative CT RIV algorithm has not been proven so far and is only empirically assumed [20]. An additional concern is that even if conditions C1-C6 are fulfilled, assuming that the properties of the approach established in the LTI case apply to the estimation of the reformulated LPV model would mean that the noise-free output terms are *a priori* known. Therefore, even if the presented method considerably lowers the variance in the estimated parameters, the optimality of the estimates cannot be guaranteed.

C. The `LPV-RIVC` *and* `LPV-SRIVC` *(*Simplified *RIVC) Algorithms*

Based on the previous considerations, the iterative scheme of the LPV-*Refined Instrumental Variable for Continuous-time models* (`LPV-RIVC`) as well as the simplified version (`LPV-SRIVC`) can be given in the considered hybrid LPV framework.

1) The `LPV-RIVC` *Algorithm:* The following algorithm is designed for hybrid LPV-BJ models.

Algorithm 1 (`LPV-RIVC`*):*

Step 1 The usual initialization for `CT-RIV` algorithm is a DT model estimate issued from an `LS` method or a `DT-RIV` algorithm. In the LPV case however, the transformation of a DT model into a CT model is not trivial. Consequently, the initial estimate proposed for the `LPV-RIVC` algorithm is an `LTI-RIVC` estimate of $\mathcal{M}_\theta$, i.e. $\hat{\theta}^{(0)} = [\ (\hat{\rho}^{(0)})^\top \quad (\hat{\eta}^{(0)})^\top\ ]^\top$ is given. Set $\tau = 0$.

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

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

where $\hat{\rho}^{(\tau)}$ is estimated in the previous iteration. Based on $\mathcal{M}_{\hat{\theta}^{(\tau)}}$, deduce $\hat{\chi}(t_k)$ which is bounded according to Assumption A5. Moreover in terms of Assumption A3, $\chi(t)$ is not correlated to the noise.

Step 3 Compute the estimated continuous-time filter $\hat{Q}_\mathrm{c}(\mathrm{d}, \hat{\rho}^{(\tau)})$:

$$\hat{Q}_\mathrm{c}(\mathrm{d}, \hat{\rho}^{(\tau)}) = \frac{1}{F(\mathrm{d}, \hat{\rho}^{(\tau)})}, \tag{36}$$

where $F(\mathrm{d}, \hat{\rho}^{(\tau)})$ is as given in (17).

Step 4 Use the CT filter $\hat{Q}_\mathrm{c}(\mathrm{d}, \hat{\rho}^{(\tau)})$ as well as $\hat{\chi}(t_k)$ in order to generate the estimates of the derivatives which are needed later to construct the regressor:

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

$$\{\hat{Q}_\mathrm{c}(\mathrm{d}, \hat{\rho}^{(\tau)})u_{j,l}(t_k)\}_{j=0,l=0}^{n_\mathrm{b},n_\beta}, \qquad \{\hat{Q}_\mathrm{c}(\mathrm{d}, \hat{\rho}^{(\tau)})\hat{\chi}_{i,l}(t_k)\}_{i=1,l=0}^{n_\mathrm{a},n_\alpha}.$$

Step 5 Compute the estimated discrete-time filter:

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

Step 6 The needed filtered signals $\{u_{j,l}^\mathrm{f}(t_k)\}_{j=0,l=0}^{n_\mathrm{b},n_\beta}$, $y_\mathrm{f}(t_k)$ and $\{\chi_{i,l}^\mathrm{f}(t_k)\}_{i=1,l=0}^{n_\mathrm{a},n_\alpha}$ are computed by applying the DT filter $\hat{Q}_\mathrm{d}$ on the estimated derivatives obtained in Step 4.

Step 7 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:

$$\hat{\varphi}_\mathrm{f}(t_k) = \Big[\ -y_\mathrm{f}^{(n_a-1)}(t_k) \quad \ldots \quad -y_\mathrm{f}(t_k) \quad -\hat{\chi}_{1,1}^\mathrm{f}(t_k) \quad \ldots \quad -\hat{\chi}_{n_\mathrm{a},n_\alpha}^\mathrm{f}(t_k) \quad u_{0,0}^\mathrm{f}(t_k) \quad \ldots \quad u_{n_\mathrm{b},n_\beta}^\mathrm{f}(t_k)\ \Big]^\top,$$

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

Note that the generation of $\chi(t)$ using Step 2 guarantees that the instrument is not correlated to the noise and therefore implies C2.

Step 8 The IV optimization problem can now be stated in the form

$$\hat{\rho}^{(\tau+1)}(N) = \arg\min_{\rho\in\mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(t_k)\hat{\varphi}_{\mathrm{f}}^{\top}(t_k) \right] \rho - \left[\frac{1}{N} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(t_k) y_{\mathrm{f}}^{(n_a)}(t_k) \right] \right\|^2 . \tag{37}$$

where the solution is obtained as

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

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 9 An estimate of the noise signal v is obtained as

$$\hat{v}(t_k) = y(t_k) - \hat{\chi}(t_k, \hat{\rho}^{(\tau)}). \tag{38}$$

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 [20]).

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.

2) The `LPV-SRIVC` *Algorithm:* Based on a similar concept for the estimation of CT LPV-OE models, the so-called *simplified* `LPV-RIVC` (`LPV-SRIVC`) method can also be developed. This method is based on a model structure (14) with $C(q^{-1},\eta) = D(q^{-1},\eta) = 1$ and consequently in this case $Q_{\mathrm{d}}(q^{-1},\eta) = 1$. Therefore the `LPV-SRIVC` algorithm remains the same as the `LPV-RIVC` algorithm except Step 5, 6 and 10 of Algorithm 1 are skipped. Naturally, the `LPV-SRIVC` does not lead to the statistically optimal PEM for hybrid LPV-BJ models, however it still leads to consistent estimates. Moreover, a CT LPV-OE models does not involve any DT filtering and consequently, their structure is fully CT unlike for hybrid LPV-BJ models.

IV. Simulation Example

As a next step, the performance of the proposed algorithms are presented on a representative simulation example. It is important to note that to the best of the authors' knowledge, the presented method is the first approach able to handle the case of colored output measurement noise for CT LPV models. Therefore, the results obtained for the presented algorithm cannot be compared to other algorithms.

A. Data generating system

The system taken into consideration is inspired by a benchmark example proposed by Rao and Garnier in [17]. It has been widely used since then to demonstrate the performance of direct continuous-time identification methods [17]–[19], [34], [35]. 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 $\mathcal{S}_\mathrm{o}$ is a function of p. This phenomenon often occurs in mechatronic applications such as for instance, wafer scanners [36]. 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:

$$\mathcal{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(\mathrm{d},p) &= \mathrm{d}^4 + (2\zeta_2\omega_2(p) + 2\zeta_1\omega_1)\,\mathrm{d}^3 + (\omega_1^2 + \omega_2^2(p) + 4\zeta_2\zeta_1\omega_2(p)\omega_1)\,\mathrm{d}^2 \\ & \quad + (2\zeta_2\omega_2(p)\omega_1^2 + 2\zeta_1\omega_2^2(p)\omega_1)\mathrm{d} + \omega_2^2(p)\omega_1^2 \\ B_\mathrm{o}(\mathrm{d},p) &= -T\omega_2^2(p)\omega_1^2\mathrm{d} + \omega_2^2(p)\omega_1^2 \\ H_\mathrm{o}(q) &= \dfrac{1}{1-q^{-1}+0.2q^{-2}} \end{cases} \tag{39}$$

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

$$G_\mathrm{RG}(\mathrm{d}) = \frac{-T\mathrm{d}+1}{(\frac{\mathrm{d}^2}{\omega_1^2} + 2\zeta_1\frac{\mathrm{d}}{\omega_1} + 1)(\frac{\mathrm{d}^2}{\omega_2^2(0)} + 2\zeta_2\frac{\mathrm{d}}{\omega_2(0)} + 1)}. \tag{40}$$

Using the given numerical values, $\mathcal{S}_\mathrm{o}$ takes the following form

$$\mathcal{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(\mathrm{d},p) = \mathrm{d}^4 + a_1^\mathrm{o}(p)\mathrm{d}^3 + a_2^\mathrm{o}(p)\mathrm{d}^2 + a_3^\mathrm{o}(p)\mathrm{d} + a_4^\mathrm{o}(p) \\ B_\mathrm{o}(\mathrm{d},p) = b_0^\mathrm{o}(p)\mathrm{d} + b_1^\mathrm{o}(p) \\ H_\mathrm{o}(q) = \dfrac{1}{1-q^{-1}+0.2q^{-2}} \end{cases} \tag{41}$$

where

$$a_1^{\text{o}}(p) = 5 + 0.25p, \qquad a_2^{\text{o}}(p) = 408 + 3p + 0.25p^2, \tag{42a}$$

$$a_3^{\text{o}}(p) = 416 + 108p + p^2, \qquad a_4^{\text{o}}(p) = 1600 + 800p + 100p^2, \tag{42b}$$

$$b_0^{\text{o}}(p) = -6400 - 3200p - 400p^2, \qquad b_1^{\text{o}}(p) = 1600 + 800p + 100p^2. \tag{42c}$$

The Bode plot of 20 frozen behaviors of $\mathcal{S}_{\text{o}}$ is depicted in Figure 1 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.

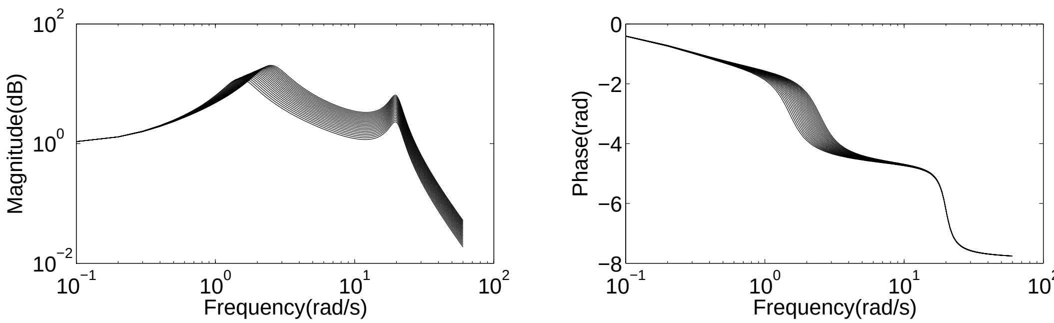

Fig. 1. Bode plot of the frozen behaviors of the true LPV system for 20 values of the scheduling variable p (between -1 to 1)

To obtain data records for identification purposes, the input signal u is chosen as a uniformly distributed sequence $\mathcal{U}(-1, 1)$, while the scheduling variable is chosen as $p(t) = \sin(\pi t)$. Furthermore, the sampling period is chosen as $T_{\text{s}} = 1\text{ms}$, and the simulation time is $T_{\max} = 10\text{s}$ which, considering the currently available acquisition possibilities, is a fair assumption.

B. Model structures

In the sequel both the `LPV-RIVC` and the `LPV-SRIVC` algorithms are studied for the identification of the data generating system $\mathcal{S}_{\text{o}}$. The proposed `LPV-RIVC` method is applicable to the hybrid LPV-BJ model and assumes the following model structure:

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

where

$$a_1(p) = a_{1,0} + a_{1,1}p, . \qquad a_2(p) = a_{2,0} + a_{2,1}p + a_{2,2}p^2, \tag{43a}$$

$$a_3(p) = a_{3,0} + a_{3,1}p + a_{3,2}p^2, \qquad a_4(p) = a_{4,0} + a_{4,1}p + a_{4,2}p^2, \tag{43b}$$

$$b_0(p) = b_{0,0} + b_{0,1}p + b_{0,2}p^2, \qquad b_1(p) = b_{1,0} + b_{1,1}p + b_{1,2}p^2. \tag{43c}$$

while the `LPV-SRIVC` method is applicable to CT LPV-OE models and assumes the following model structure:

$$\mathcal{M}_{\texttt{LPV-SRIVC}} \begin{cases} A(d, p) = d^4 + a_1(p)d^3 + a_2(p)d^2 + a_3(p)d + a_4(p) \\ B(d, p) = b_0(p)d + b_1(p) \\ H(q) = 1 \end{cases} \tag{44}$$

with $a_1(p)$, $a_2(p)$, $a_3(p)$, $a_4(p)$, $b_0(p)$, $b_1(p)$ as given in (43a-f). Note that to demonstrate the achievable performance with the proposed methods we assume that information about the plant in terms of model order and structural dependency is known priori.

In terms of identification, the model $\mathcal{M}_{\texttt{LPV-RIVC}}$ corresponds to the case $\mathcal{S}_\mathrm{o} \in \mathcal{M}$ while the model $\mathcal{M}_{\texttt{LPV-SRIVC}}$ corresponds to the more realistic practical assumption $\mathcal{G}_\mathrm{o} \in \mathcal{G}$ and $\mathcal{H}_\mathrm{o} \notin \mathcal{H}$. Therefore, 17 parameters are to be estimated by the `LPV-SRIVC` algorithm and 19 by the `LPV-RIVC` algorithm. To provide representative results, a *Monte Carlo* (MCs) simulation of $N_\mathrm{MC} = 200$ random realizations is used with a *Signal-to-Noise Ratio* (SNR) of 20dB where

$$\mathrm{SNR} = 10 \log \frac{P_{\chi_\mathrm{o}}}{P_{v_\mathrm{o}}}, \tag{45}$$

and P_x is the power of signal x. The MC results obtained using both algorithms are presented in Table I. It can be clearly seen that the estimates are unbiased which conforms to the theory. The standard deviation for the nominal part of the coefficients ($a_{*,0}(p)$ and $b_{*,0}(p)$) remains low whereas it raises considerably for coefficients $a_{*,2}(p)$ and $b_{*,2}(p)$.

In Figure 2, the 200 simulated model outputs are plotted together with validation data set (generated under the same excitation conditions as the one used for estimation). It can be seen that despite the large variance in the estimated parameters $a_{*,2}(p)$ and $b_{*,2}(p)$, the simulated outputs remain close to the true noise-free output signal (considering the level of noise corresponding to a $\mathrm{SNR} = 20\mathrm{dB}$). It appears therefore, that these parameter values have a low contribution to the observed output signal under these excitation conditions. Thus, in terms of minimization of the squared prediction error, their role is less significant which results in a relatively large variance of their estimates under noisy conditions. This fact underlines that experiment design is needed

TABLE I

MONTE CARLO SIMULATION RESULTS WITH ADDITIVE COLORED MEASUREMENT NOISE FOR $SNR = 20$ dB

		LPV-RIVC		LPV-SRIVC	
Name	True Value	mean	st. dev.	mean	st. dev.
$a_{1,0}$	5	4.99	0.058	4.99	0.059
$a_{1,1}$	0.25	0.249	0.081	0.249	0.082
$a_{2,0}$	408	407.94	1.27	407.95	1.29
$a_{2,1}$	3	2.93	1.47	2.92	1.47
$a_{2,2}$	0.25	0.268	2.35	0.251	2.37
$a_{3,0}$	416	415.64	14.22	415.61	14.19
$a_{3,1}$	108	107.72	11.20	107.76	11.13
$a_{3,2}$	1	1.14	30.62	1.17	30.40
$a_{4,0}$	1600	1598.9	44.58	1599.1	44.67
$a_{4,1}$	800	799.31	27.22	799.38	27.13
$a_{4,2}$	100	101.31	79.85	101.08	80.01
$b_{0,0}$	-6400	-6396.4	47.95	-6396.4	48.03
$b_{0,1}$	-3200	-3195.2	65.72	-3195.2	65.76
$b_{0,2}$	-400	-398.92	70.40	-399.14	70.30
$b_{1,0}$	1600	1593.8	366.9	1593.8	365.7
$b_{1,1}$	800	805.02	219.77	802.43	221.7
$b_{1,2}$	100	104.86	749.61	102.95	743.9
d_1	-1	-0.999	0.0096	X	X
d_2	0.2	0.202	0.0094	X	X

to minimize the variance of the parameters estimates. However, there is a lack of input design methods suited for the identification of LPV systems (see [28], [9] and references therein) as the concept of persistency of excitation is not well understood yet for LPV models. A procedure for input design for CT-LTI models such as in [35] might be suitable to solve the estimation problem of the $a_{*,2}(p)$ and $b_{*,2}(p)$ parameters but optimal input design is not investigated for the considered system here.

Consequently, the quality of the estimated model cannot be judged only from the parameter values. Therefore, the Bode plot of each estimated model ($N_{\mathrm{MC}} = 200$) at 20 frozen values of p from -1 to 1 are depicted in Figure 3. It can be clearly observed from the Bode plot that the large variance of the parameters $a_{*,2}(p)$ and $b_{*,2}(p)$ plays an important role in the low frequency

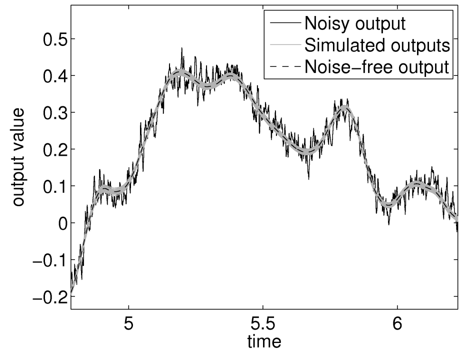

Fig. 2. 200 simulated model outputs together with a noise-free and a single MCs output

area only whereas it does not affect the quality of the estimated models for frequencies above 0.5 rad/s. Therefore in order to analyze the distribution of responses at low frequencies, the density of curves at frequency $\omega = 0.1$ rad/s is displayed on the left hand-side of Figure 3 both for the magnitude and phase. 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. In terms of this description, it becomes clear that the estimated models are normally distributed and centered on the true model. Moreover, it can be seen from Table I that the `LPV-RIVC` and `LPV-SRIVC` give very similar variances in the estimated parameters. This reinforces the hypothesis of a sub-optimal excitation.

Finally, the most interesting advantage of the direct CT estimation in LPV framework is shown in Figure 4. In this figure, the poles of the 200 estimated models at 20 fixed values of p between -1 and 1 are plotted against the poles of the true model. This figure also displays in the top part the density of the real part of the poles. Using the same idea as in Figure 3, the intensity of each displayed pole is related to the number of poles among all MC simulations which have the same real part. Using this representation, it can clearly be 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 in the close neighborhood of the imaginary part of the “fixed pole”. This means that by using a parsimonious CT model, the estimated models perfectly captures the time-varying nature of $\mathcal{S}_\text{o}$ as well as the transient behavior from one operating condition to the other, which is a significant achievement considering the complexity of the studied system and

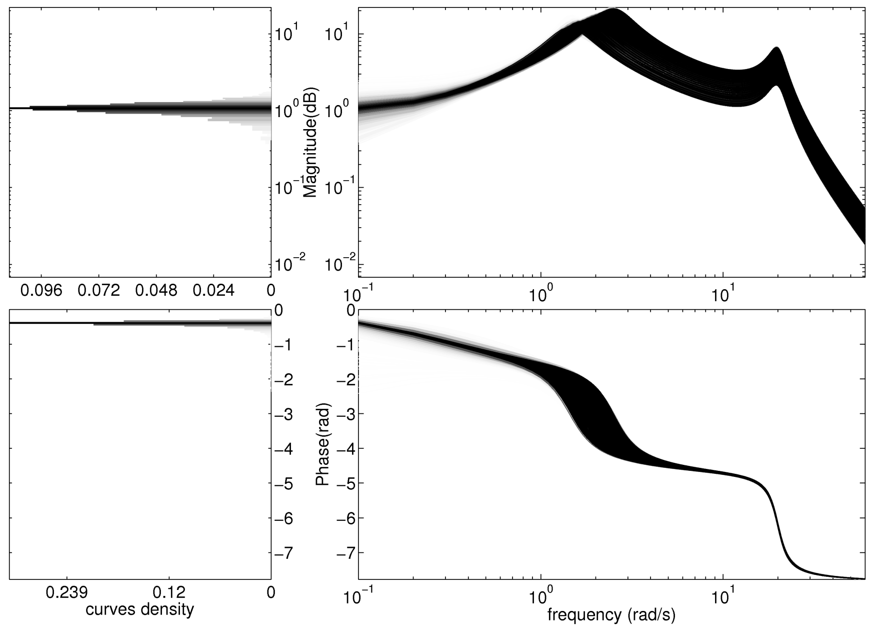

Fig. 3. Bode plots of the estimated models

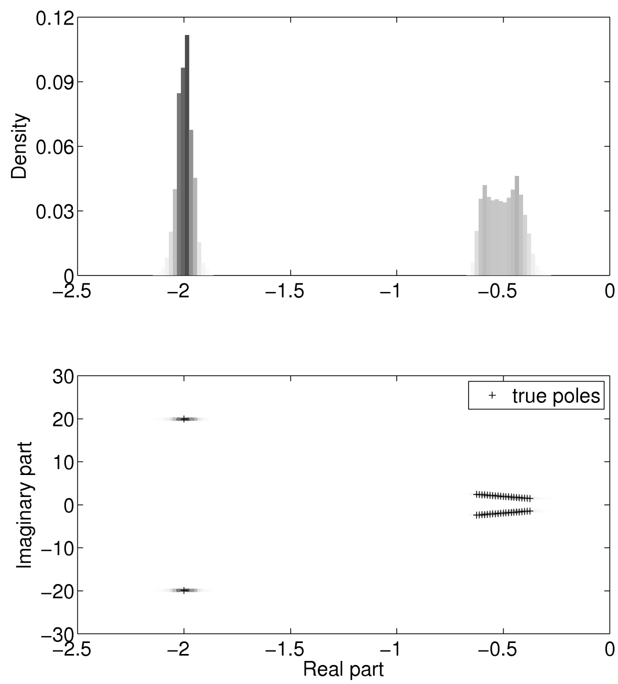

Fig. 4. Poles of the estimated models and of the true system

the level of additive noise.

C. Comments

Note that trying to identify a discrete-time model for the considered CT system is quite a tedious task and requires the following critical issues to be addressed:

- In the LTI case, it has been shown that direct CT identification methods are better suited for the identification of the Rao-Garnier benchmark [18]. In the presented LPV context, this means that any frozen behavior of $\mathcal{S}_{\mathrm{o}}$ (for any fixed trajectory of p) is better identified using a direct CT identification method.
- Using the trapezoidal integration method [16], the system described in (1) is discretized as:

$$y(t_k) = \frac{2 - \mathtt{T}_{\mathrm{s}}p(t_{k-1})}{2 + \mathtt{T}_{\mathrm{s}}p(t_k)} y(t_{k-1}) + \frac{\mathtt{T}_{\mathrm{s}}b}{2 + \mathtt{T}_{\mathrm{s}}p(t_k)} u(t_k) + \frac{\mathtt{T}_{\mathrm{s}}b}{2 + \mathtt{T}_{\mathrm{s}}p(t_k)} u(t_{k-1}). \tag{46}$$

  By analyzing this simple discretization scheme, it is clear that applying it to (39) results in i) the augmentation of the number of parameters to be estimated, ii) non-linear-in-the-parameters dependence on p iii) some dynamic dependence on p (appearance of $p(t_{k-1}), p(t_{k-2}) \ldots$ terms).

Estimating such a DT model, where the dependences on p are non-linear-in-the-parameters, is hardly feasible using the existing parametric methods as it requires a customized nonlinear optimization approach. Alternatively, non-parametric methods can be applied or the model can be approximated by a simplified model structure. The latter relaxes the assumption $\mathcal{G}_{\mathrm{o}} \in \mathcal{G}$: in other words, the time-varying property of the LPV system (in this case, a "moving pole") would not anymore be directly linked to the coefficients of the considered DT model and consequently, could not be clearly identified.

V. Conclusion

Due to the lack of methods dedicated to the case of direct continuous-time identification of LPV models, there exists a clear gap between the available identification approaches and the practical needs of control synthesis. In order to bridge this gap, a novel method has been proposed in this paper for the identification of hybrid LPV-BJ models and CT LPV-OE models with a p-independent noise process.

The presented method is based on a particular MISO-LTI reformulation of the data equations which enables the use of Refined IV-based methods for LPV IO models in the error-prediction-minimization framework. The proposed algorithm has been tested on a representative numerical simulation example inspired by the Rao-Garnier benchmark. The presented example has shown that the proposed procedure is robust to noise and can reasonably well estimate the system in case of an imperfect noise model. Furthermore, it was motivated that in the given LPV framework and for relatively complicated systems, a direct CT estimation method is an attractive approach for capturing the true time-varying nature of the studied system. In this paper, only the case of p-independent noise models has been investigated. Even if refined IV-based methods are theoretically unbiased for p-dependent noise models, the investigation of scheduling dependent noise models remain as a topic for future research.

REFERENCES

[1] P. Apkarian and P. Gahinet, "A convex characterization of gain-scheduled $\mathcal{H}_\infty$ controllers," *IEEE Transactions on Automatic Control*, vol. 40, no. 5, pp. 853–864, 1995.

[2] C. W. Scherer, "Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control for time-varying and linear parametrically-varying systems," *Int. Journal of Robust and Nonlinear Control*, vol. 6, no. 9-10, pp. 929–952, 1996.

[3] F. Wu and K. Dong, "Gain-scheduling control of LFT systems using parameter-dependent Lyapunov functions," *Automatica*, vol. 42, pp. 39–50, 2006.

[4] A. Marcos and G. J. Balas, "Development of linear-parameter-varying models for aircraft," *Journal of Guidance, Control and Dynamics*, vol. 27, no. 2, pp. 218–228, 2004.

[5] K. Z. Østergaard, J. Stousturp, and P. Barth, "Rate bounded linear parameter varying control of a wind turbine in full load operation," in *17th IFAC World Congress*, Seoul, Korea, July 2008.

[6] F. Wijnheijmer, G. Naus, W. Post, M. Steinbuch, and P. Teerhuis, "Modeling and LPV control of an electro-hydraulic servo system," in *IEEE International Conf. on Control Applications*, Munich, Germany, Oct. 2006, pp. 3116–3120.

[7] M. Dettori and C. Scherer, "LPV design for a CD player: an experimental evaluation of performance," in *40th IEEE Conf. on Decision and Control*, Orlando, Florida, USA, Dec. 2001, pp. 4711–4716.

[8] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice-Hall, 1998.

[9] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, ser. Lecture Notes in Control and Information Sciences, Vol. 403. Springer-Germany, 2010.

[10] R. Murray-Smith, T. A. Johansen, and R. Shorten, "On the interpretation of local models in blended multiple model structures," *Int. Journal of Control*, vol. 72, no. 7-8, pp. 620–628, 1999.

[11] R. Tóth, F. Felici, P. S. C. Heuberger, and P. M. J. Van den Hof, "Discrete time LPV I/O and state space representations, differences of behavior and pitfalls of interpolation," in *European Control Conf.*, Kos, Greece, July 2007, pp. 5418–5425.

[12] G. P. Rao and H. Unbehauen, "Identification of continuous-time systems," *IEE Proceedings Control Theory & Appl.*, vol. 153, no. 2, March 2006.

[13] R. Tóth, M. Lovera, P. S. C. Heuberger, and P. M. J. Van den Hof, "Discretization of linear fractional representations of LPV systems," in *48th IEEE Conf. on Decision and Control*, Shanghai, China, Dec. 2009, pp. 7424–7429.

[14] R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof, "On the discretization of LPV state-space representations," *To appear in IET Control Theory & Applications*, 2010.

[15] ——, "Flexible model structures for LPV identification with static scheduling dependency," in *47th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 4522–4527.

[16] K. E. Atkinson, *An Introduction to Numerical Analysis*. John Wiley and Sons, New York, 1989.

[17] G. Rao and H. Garnier, "Numerical illustrations of the relevance of direct continuous-time model identification," in *15th IFAC World Congress*, Barcelona, Spain, 2002.

[18] ——, "Identification of continuous-time systems: direct or indirect ?" *Systems Science*, vol. 30(3), pp. 25–50, 2004.

[19] V. Laurain, M. Gilson, H. Garnier, and P. C. Young, "Refined instrumental variable methods for identification of Hammerstein continuous-time Box-Jenkins models," in *47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec 2008.

[20] P. Young, "The refined instrumental variable method: unified estimation of discrete and continuous-time transfer function models," *Journal Européen des Systèmes Automatisés*, vol. 42, pp. 149–179, 2008.

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

[22] H. Garnier and L. (Editors), *Identification of Continuous-time Models from Sampled Data*, H. Garnier and L. Wang, Eds. London: Springer-Verlag, 2008.

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

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

[25] L. Ljung, "Experiments with identification of continuous-time models," in *15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009.

[26] M. Gilson, H. Garnier, P. C. Young, and P. Van den Hof, "Optimal instrumental variable method for closed-loop identification," *IET Control Theory and Applications*, vol. To appear, 2010.

[27] M. Gilson and P. M. J. Van den Hof, "Instrumental variable methods for closed-loop system identification," *Automatica*, vol. 41, pp. 241–249, 2005.

[28] V. Laurain, M. Gilson, R. Tóth, and H. Garnier, "Refined instrumental variable methods for identification of LPV output-error and Box-Jenkins models," *Automatica*, vol. 46, Issue 6, pp. 959–967, June 2010.

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

[30] R. Johansson, "Identification of continuous–time models," *IEEE Transactions on Signal Processing*, vol. 42(4), pp. 3887–897, 1994.

[31] L. Ljung., *System Identification : Theory for the User - Second edition*. Prentice-Hall, 1999.

[32] G. Goodwin, S. Graebe, and M. Salgado, *Control System Design*. Prentice Hall, 2000.

[33] P. Young, H. Garnier, and M. Gilson, "Refined instrumental variable identification of continous-time hybrid Box-Jenkins models," in *Identification of continuous-time models from sampled data. H. Garnier and L. Wang (editor)*, Springer-Verlag, London, 2008, pp. 91–131.

[34] L. Ljung, "Initialisation aspects for subspace and output-error identification methods," in *European Control Conference (ECC'2003)*, Cambridge (U.K.), September 2003.

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

[36] M. G. Wassink, M. van de Wal, C. W. Scherer, and O. Bosgra, "LPV control for a wafer stage: Beyond the theoretical solution," *Control Engineering Practice*, vol. 13, pp. 231–245, 2004.