LPV system identification under noise corrupted scheduling and output signal observations ⋆

Dario Piga [a], Pepijn Cox [b], Roland Tóth [b], Vincent Laurain [c].

[a] *Dalle Molle Institute for Artificial Intelligence Research, SUPSI, Galleria 1, 6928 Manno, Switzerland.*

[b] *Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands.*

[c] *Centre de Recherche en Automatique de Nancy, Université de Lorraine, CNRS, 2 rue Jean Lamour, 54519 Vandoeuvre-lès-Nancy Cedex, France.*

Abstract

Most of the approaches available in the literature for the identification of *Linear Parameter-Varying* (LPV) systems rely on the assumption that only the measurements of the output signal are corrupted by the noise, while the observations of the scheduling variable are considered to be noise free. However, in practice, this turns out to be an unrealistic assumption in most of the cases, as the scheduling variable is often related to a measured signal and, thus, it is inherently affected by a measurement noise. In this paper, it is shown that neglecting the noise on the scheduling signal, which corresponds to an error-in-variables problem, can lead to a significant bias on the estimated parameters. Consequently, in order to overcome this corruptive phenomenon affecting practical use of data-driven LPV modeling, in this paper we present an identification scheme to compute a consistent estimate of LPV *Input/Output* (I/O) models from noisy output and scheduling signal observations. A simulation example is provided to prove the effectiveness of the proposed methodology.

Key words: Linear Parameter-Varying systems; Parameter estimation; Instrumental Variables; System identification.

1 Introduction

The concept of *linear parameter-varying* (LPV) systems, originally introduced in [19], offers a promising framework for modeling and control of a large class of *nonlinear* (NL) and *time-varying* (TV) systems. LPV systems can be seen as an extension of *linear time-invariant* (LTI) systems as the dynamic relation between the input and the output signals is linear. Unlike in the LTI case, these signal relations are allowed to change over time and they are assumed to depend on a measurable time-varying signal, commonly referred in the literature as *scheduling variable*. Scheduling variables can be external signals like space coordinates, measurable disturbances or changing operating conditions. In this way, the nonlinear and time-varying behavior of the system can be embedded in the solution set of a linear dynamic input-output relationship which varies with the scheduling variable.

The LPV modeling paradigm has evolved rapidly in the last twenty years, and it has been employed in many control applications like aircrafts [13,12], automotive applications [28,16,5], robotic manipulators [7] and induction motors [18]. Motivated by the need of accurate and low-complexity LPV models for control design purposes, significant efforts have been spent in the last years for developing efficient approaches for identification of LPV systems from measured data. In the current literature, the existing LPV identification approaches have been mainly formulated in *discrete-time* (DT) and they are categorized by the used model structure. In particular, identification schemes for LPV models based on a *state-space* (SS) representation are discussed in [22,26,6,14,11,25], identification of LPV models represented in terms of a

⋆ This paper was not presented at any IFAC meeting. Corresponding author Dario Piga. Tel. +41-(0)58-666-6664.

⋆⋆This work has been partly supported by the Netherlands Organization for Scientific Research (NWO, grant no.: 639.021.127) and by the French ministries of Foreign Affairs, Education and Research and the French-Dutch Academy (PHC Van Gogh project, n. 29342QL).

Email addresses: dario.piga@supsi.ch (Dario Piga), p.b.cox@tue.nl (Pepijn Cox), r.toth@tue.nl (Roland Tóth), vincent.laurain@univ-lorraine.fr (Vincent Laurain).

series expansion is dealt with in [24], while identification of LPV *input-output* (IO) models is addressed in [2,3,1,27,9]. The reader is referred to [23] for an overview on the existing LPV identification methods.

Most of the LPV identification methods rely on the assumption that only the output measurements are corrupted by noise, while the observations of the scheduling signal are considered to be noise free. However, in practice, this turns out to be an unrealistic assumption in most of the cases, as the scheduling variable is often related to a measured signal and, thus, it is inherently affected by a measurement noise. Assuming noise on the scheduling variables corresponds to a so-called *error-in-variables* (EIV) problem for which several identification methods have been proposed in the LTI context [29,8,20]. In the LPV setting, in order to model a broader class of behaviors, the dependency on the scheduling variable is mostly considered to be nonlinear. This renders the EIV problem more difficult than in the LTI case since the stochastic noise affecting the scheduling observations is distorted by nonlinear functions. To the best of the authors' knowledge, the only contributions available in the literature addressing the identification of LPV systems from noise-corrupted measurements of both the output and the scheduling signal are given in [4] and [3]. The contribution in [4] focuses on the identification of LPV-IO systems in the *set-membership* (or *bounded-error*) identification context. Specifically, a convex relaxation approach is used to evaluate an outer-bounding box of the so-called *feasible parameter set* (FPS), that is the set of all system parameters consistent with the measurements, the error bounds and the assumed model structure, under the assumption that the noise corrupting the output and scheduling signal observations is bounded. The worst-case realization of the noise sequence is computed and hard bounds on the values which can be assumed by the system parameters are obtained. Unfortunately, the selection of a nominal model within the computed approximation of the FPS remains an open problem. Furthermore, because of high computational complexity, the applicability of the approach in [4] is limited to small/medium scale LPV identification problems. The work [3] proposes an *instrumental-variable* (IV) based approach for identification of LPV-IO systems in a statistical framework. The advantage of using IV-based estimate in the LPV setting is the ability to cope with realistic assumptions on the output noise instead of applying nonlinear optimization. The IV method delivers unbiased estimates for a large generality of noise scenarios with the sole condition that all disturbances can be expressed as a zero mean noise process. Furthermore, refined extension of the IV method in [9] offer a solution to minimize the variance of the IV estimate for even *Box-Jenkins* (BJ) noise models under the condition of noise-free scheduling variables. In case of noisy dependency, it is shown in [3] that unbiased estimates can be computed if i) the instrument is uncorrelated with the scheduling variable noise and ii) the scheduling dependency is linear. The latter condition limits the applicability of the IV-based scheme in [3]. Furthermore, as discussed in [9], in order to minimize the variance of the IV estimate, the chosen instrument should be correlated with the noise-free observations of the output and of the scheduling signal. Although the refined IV algorithm proposed in [9] can be used to iteratively compute an approximation of the noise-free output, computing a variable which is correlated with the noise-free scheduling signal, and at the same time uncorrelated with the measurement noise, can be a difficult task in many applications, in particular when the scheduling signal is not directly manipulatable.

In this paper, a bias-corrected, IV-based method is developed for the identification of LPV models from noise-corrupted measurements of the output and of the scheduling signal. The noise process associated to the output is assumed to be colored, zero-mean and with unknown distribution, while the measurements of the scheduling signal are supposed to be affected by a white Gaussian noise. The advantages of the proposed method w.r.t. the approach in [3] are twofold:

- it provides a consistent estimate of LPV models with polynomial dependence on the scheduling variable;
- the used instrument is only required to be uncorrelated with the noise corrupting the output observations. Thus, an approximation of the noise-free scheduling signal does not need to be computed.

The paper is organized as follows. Section 2 is devoted to the problem formulation. In Section 3, the asymptotic properties of the IV approach, developed in [3], are discussed. A bias-corrected IV-based method is presented in Section 4 to compute a consistent estimate of the system parameters under the assumption that the variance of the noise corrupting the scheduling signal measurements is *a-priori* known. This assumption is relaxed in Section 5 to extend the applicability of the developed method for a more general setting. The effectiveness of the presented identification approach is shown in Section 6 through a simulation example and the obtained results are compared with the ones obtained by the IV method in [3].

2 Problem Description

2.1 The data-generating system

Consider a discrete-time *single-input single-output* (SISO) linear-parameter-varying system $\mathcal{S}_\mathrm{o}$ described by the difference equation

$$\mathcal{A}_\mathrm{o}(p_\mathrm{o}(k), q^{-1}) y_\mathrm{o}(k) = \mathcal{B}_\mathrm{o}(p_\mathrm{o}(k), q^{-1}) u(k), \tag{1a}$$
$$y(k) = y_\mathrm{o}(k) + v_\mathrm{o}(k), \tag{1b}$$

where $u(k) \in \mathbb{R}$ is the input signal, $p_\mathrm{o}(k) \in \mathbb{R}^{n_\mathrm{P}}$ is the value of the scheduling variable at time k, $y_\mathrm{o}(k) \in \mathbb{R}$ and $y(k) \in \mathbb{R}$ are the noise-free and the noise-corrupted output, respectively, $v_\mathrm{o}(k) \in \mathbb{R}$ is an additive zero-mean colored noise with bounded spectral density. For the sake of exposition, only the case of scalar scheduling variable $p_o(k)$ (i.e., $n_p = 1$) will be discussed in the paper. The terms $\mathcal{A}_o(p_o(k), q^{-1})$ and $\mathcal{B}_o(p_o(k), q^{-1})$ in (1) are polynomials in the backward shift operator q^{-1} (i.e., $q^{-1}u(k) = u(k-1)$) of finite degree n_a and n_b, respectively, i.e.,

$$\mathcal{A}_o(p_o(k), q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i^o(p_o(k))q^{-i}, \tag{2a}$$

$$\mathcal{B}_o(p_o(k), q^{-1}) = \sum_{j=0}^{n_b} b_j^o(p_o(k))q^{-j}, \tag{2b}$$

where $a_i^o : \mathbb{R} \to \mathbb{R}$ and $b_j^o : \mathbb{R} \to \mathbb{R}$ are polynomial functions in the scheduling variable $p_o(k)$ of maximum degree n_g, i.e.,

$$a_i^o(p_o(k)) = a_{i,0}^o + \sum_{s=1}^{n_g} a_{i,s}^o p_o^s(k), \tag{3a}$$

$$b_j^o(p_o(k)) = b_{j,0}^o + \sum_{s=1}^{n_g} b_{j,s}^o p_o^s(k), \tag{3b}$$

where $a_{i,s}^o$ and $b_{j,s}^o$ (with $s \in \{0, \ldots, n_g\}$) are unknown real constants. For the sake of simplicity, in (3), the polynomials a_i^o and b_j^o are considered with the same degree n_g. Nevertheless, the discussion reported in the sequel can be straightforwardly extended to the case when a_i^o and b_j^o have different degrees.

In order to describe the data-generating system $\mathcal{S}_o$ in a compact form, let us introduce the following matrix notation:

$$\begin{aligned}
&\underline{a}_i^o = \left[\ a_{i,0}^o \ \ a_{i,1}^o \ \ \ldots \ \ a_{i,n_g}^o\ \right]^\top, \\
&\underline{b}_j^o = \left[\ b_{j,0}^o \ \ b_{j,1}^o \ \ \ldots \ \ b_{j,n_g}^o\ \right]^\top, \\
&\theta_o = \left[\ (\underline{a}_1^o)^\top \ \ldots \ (\underline{a}_{n_a}^o)^\top \ (\underline{b}_0^o)^\top \ \ldots \ (\underline{b}_{n_b}^o)^\top\ \right]^\top, \\
&\mathbf{p}_o(k) = \left[1 \ \ p_o(k) \ \ p_o^2(k) \ldots \ \ p_o^{n_g}(k)\right]^\top, \\
&\chi_o(k) = [-y_o(k-1) \ \ldots \ -y_o(k-n_a) \ \ u(k) \ldots \ u(k-n_b)]^\top, \\
&\phi_o(k) = \chi_o(k) \otimes \mathbf{p}_o(k),
\end{aligned}$$

with $\otimes$ denoting the Kronecker product. Based on the notation introduce above, the data-generating system in (1) is rewritten as

$$y(k) = \phi_o^\top(k)\theta_o + v_o(k). \tag{4}$$

2.2 Scheduling signal measurements

The observations $p(k)$ of the scheduling signal $p_o(k)$ are affected by an additive measurement noise $\eta_o(k)$, i.e.,

$$p(k) = p_o(k) + \eta_o(k), \tag{5}$$

where $\eta_o(k)$ is a Gaussian distributed white noise process with zero-mean and finite variance σ_η^2, i.e., $\eta_o(k) \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\mathbb{E}\{\eta_o(k)\eta_o(t)\} = \sigma_\eta^2 \delta_{k,t}$, with $\delta_{k,t}$ denoting the Kronecker delta. Furthermore, we assume that $\eta_o(k)$ is independent of the noise $v_o(k)$ corrupting the measurements of the output signal $y(k)$, i.e., $\mathbb{E}[\eta_o(k)v_o(t)] = 0$ for all time indexes k and t. Note that considering $\eta_o(k)$ independent of $v_o(k)$ is a realistic assumption, since usually the scheduling variable is considered to be an external signal of the system, and thus the noise $\eta_o(k)$ corrupting the measurements of $p_o(k)$ is independent of the noise $v_o(t)$ which in turns corrupts the output observations.

2.3 The considered model structure

The following parameterized model structure $\mathcal{M}(\theta)$ is used to identify $\mathcal{S}_o$, i.e., estimate the parameters θ_o:

$$y(k) = -\sum_{i=1}^{n_a} a_i(p(k))y(k-i) + \sum_{j=0}^{n_b} b_j(p(k))u(k-j) + \varepsilon(k), \tag{6}$$

with $\varepsilon(k) \in \mathbb{R}$ denoting the residual term. The functions $a_i : \mathbb{R} \to \mathbb{R}$ and $b_j : \mathbb{R} \to \mathbb{R}$, which depend on the noisy observation $p(k)$ of the scheduling signal, are polynomials parameterized as

$$a_i(p(k)) = a_{i,0} + \sum_{s=1}^{n_g} a_{i,s} p^s(k), \tag{7a}$$

$$b_j(p(k)) = b_{j,0} + \sum_{s=1}^{n_g} b_{j,s} p^s(k). \tag{7b}$$

Based on a notation similar to the one introduced in Section 2.1, the parametric model $\mathcal{M}(\theta)$ in (7) is written in the linear regression form:

$$y(k) = \phi^\top(k)\theta + \varepsilon(k), \tag{8}$$

where $\theta \in \mathbb{R}^{n_\theta}$ (with $n_\theta = (n_g + 1)(n_a + n_b + 1)$) is a vector stacking the parameters that characterize the model $\mathcal{M}(\theta)$, i.e.,

$$\theta = \left[\ \underline{a}_1^\top \ \ldots \ \underline{a}_{n_a}^\top \ \underline{b}_0^\top \ \ldots \ \underline{b}_{n_b}^\top\ \right]^\top, \tag{9}$$

with $\underline{a}_i$ and $\underline{b}_j$ defined similarly as before and $\phi(k)$ being the observed regressor at time k defined as

$$\phi(k) = \chi(k) \otimes \mathbf{p}(k), \tag{10}$$

with

$$\begin{aligned}
&\chi(k) = [-y(k-1) \ \ldots \ -y(k-n_a) \ \ u(k) \ \ldots \ u(k-n_b)]^\top, \\
&\mathbf{p}(k) = \left[1 \ \ p(k) \ \ p^2(k) \ \ldots \ \ p^{n_g}(k)\right]^\top.
\end{aligned}$$

2.4 The problem setting

The parameters θ_o, describing the LPV system to be identified, are estimated based on an observed data sequence $\mathcal{D}_N = \{u(k), p(k), y(k)\}_{k=1}^N$ of the system $\mathcal{S}_o$. The identification scheme developed in this work is based on a proper modification of the instrumental-variable approach proposed in [3]. Before introducing the developed method, the asymptotic behaviour of the IV-based estimates for LPV systems are analyzed in the following section.

3 Asymptotic behaviour of the IV estimate

The basic IV estimate, originally introduced in [21] for identification of LTI systems and extended to LPV system identification in [3], can be seen as a modification of the *least-squares* (LS) estimate:

$$\hat{\theta}_{\mathrm{LS}} = \left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)\phi^\top(k)\right)^{-1}\frac{1}{N}\sum_{k=1}^{N}\phi(k)y(k), \qquad (11)$$

with the following expression:

$$\hat{\theta}_{\mathrm{IV}} = \underbrace{\left(\frac{1}{N}\sum_{k=1}^{N}z(k)\phi^\top(k)\right)}_{\Gamma_N}{}^{-1}\frac{1}{N}\sum_{k=1}^{N}z(k)y(k), \qquad (12)$$

where $z(k) \in \mathbb{R}^{n_\theta}$ is the so-called *instrument* and it is chosen by the user to satisfy the following conditions:

C1 $\Gamma_*^{-1} = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^{N} z(k)\phi^\top(k)\right)^{-1}$ exists;

C2 the instrument $z(k)$ is not correlated with the noise $v_o(k)$ corrupting the output measurements, i.e., $\mathbb{E}\{z(k)v_o(k)\} = 0$ for all k.

In order to analyze the asymptotic behaviour of the IV estimate $\hat{\theta}_{\mathrm{IV}}$, let us first decompose the output signal $y(k)$ as follows:

$$\begin{aligned}
y(k) &= \phi_o^\top(k)\theta_o + v_o(k) = \\
&= [\chi_o(k)\otimes\mathbf{p}_o(k)]^\top\theta_o + v_o(k) = \\
&= [\chi(k)\otimes\mathbf{p}(k)]^\top\theta_o + [\chi(k)\otimes\mathbf{p}_o(k) - \chi(k)\otimes\mathbf{p}(k)]^\top\theta_o \\
&\quad + [\chi_o(k)\otimes\mathbf{p}_o(k) - \chi(k)\otimes\mathbf{p}_o(k)]^\top\theta_o + v_o(k) = \\
&= [\chi(k)\otimes\mathbf{p}(k)]^\top\theta_o + [(\chi_o(k)-\chi(k))\otimes\mathbf{p}_o(k)]^\top\theta_o + \\
&\quad + v_o(k) + [\chi(k)\otimes(\mathbf{p}_o(k)-\mathbf{p}(k))]^\top\theta_o.
\end{aligned} \qquad (13)$$

Based on Eq. (13), the parameter estimate $\hat{\theta}_{\mathrm{IV}}$ can be decomposed as follows:

$$\begin{aligned}
\hat{\theta}_{\mathrm{IV}} &= \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)y(k) = \\
&= \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)\left[\chi(k)\otimes\mathbf{p}(k)\right]^\top\theta_o && (14a)\\
&\quad + \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)[(\chi_o(k)-\chi(k))\otimes\mathbf{p}_o(k)]^\top\theta_o && (14b)\\
&\quad + \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)v_o(k) && (14c)\\
&\quad + \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)[\chi(k)\otimes(\mathbf{p}_o(k)-\mathbf{p}(k))]^\top\theta_o. && (14d)
\end{aligned}$$

The asymptotical behaviour of each term in equation (14) is now analyzed. First, note that, because of the definition of the regressor $\phi(k)$ (see Eq. (10)), the term (14a) is equal to θ_o. As far as the terms (14b) and (14c) are concerned, they converge toward zero as N goes to infinity. This follows from the fact that $(\chi_o(k) - \chi(k))$ and $v_o(k)$ are zero mean and they are not correlated with the instrument $z(k)$ (see [3] for a rigorous proof). As far as the term (14d) is concerned, this is guaranteed to converge to zero only when **C1**, **C2** and both of the following conditions are fulfilled:

- the coefficient functions $a_i(p(k))$ and $b_j(p(k))$ depend affinely on the scheduling variable $p(k)$, i.e., $\mathbb{E}\{[\chi(k)\otimes(\mathbf{p}_o(k)-\mathbf{p}(k))]\} = 0$ since, if a_i and b_j are affine functions of $p(k)$, we have:

$$\begin{aligned}
\mathbb{E}\{\mathbf{p}_o(k)-\mathbf{p}(k)\} &= \mathbb{E}\left\{\begin{bmatrix}1\\ p_o(k)\end{bmatrix} - \begin{bmatrix}1\\ p(k)\end{bmatrix}\right\} = \\
&= \mathbb{E}\left\{\begin{bmatrix}0\\ -\eta_o(k)\end{bmatrix}\right\} = 0;
\end{aligned} \qquad (15)$$

- the instrument $z(k)$ is not correlated with the noise $\eta_o(k)$ affecting the scheduling signal measurements, i.e., $\mathbb{E}\{z(k)\eta_o(k)\} = 0$ for all k, or equivalently, provided that also the above condition and **C2** are satisfied, $\mathbb{E}\left\{z(k)[\chi(k)\otimes(\mathbf{p}_o(k)-\mathbf{p}(k))]^\top\right\} = 0$.

Thus, when at least one of the above conditions is not satisfied, the term (14d) introduces a bias in the estimate $\hat{\theta}_{\mathrm{IV}}$. Note that, in principle, the latter condition is not hard to fulfill since the instrument $z(k)$ is chosen by the user. However, in practice, in order to reduce the variance of the estimate of the system parameters θ_o, the instrument $z(k)$ should be correlated with the noise-free regressor $\phi_o(k)$ (given by $\chi_o(k)\otimes\mathbf{p}_o(k)$), as discussed in [9]. Although an approximation of the noise-free output can be obtained through iterative schemes based on simulation of a model of the system to be identified (see, e.g., [9]), approximating the noise-free scheduling signal $p_\mathrm{o}(k)$ is not an easy task. In order to overcome such a problem, two realizations of $p_\mathrm{o}(k)$ are used in [3].

In the following section, we propose an algorithm to eliminate the bias introduced by the term in (14d). This allows to compute a consistent estimate for LPV systems under the conditions that: (i) the scheduling signal observations $p(k)$ are corrupted by noise; (ii) the coefficient functions $a_i(p(k))$ and $b_j(p(k))$ depend polynomially on $p(k)$; (iii) only one realization of $p_\mathrm{o}(k)$ is available.

4 A bias-corrected IV estimate

4.1 Construction of a consistent estimate

Denote with $B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ the term in (14d), i.e., let

$$B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o}) = \Gamma_N^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)[\chi(k)\otimes(\mathbf{p}_\mathrm{o}(k)-\mathbf{p}(k))]^\top\theta_\mathrm{o}. \quad (16)$$

The term $B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ will be referred in the sequel as *structural bias*. Note that $B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ depends both on the true system parameters θ_o and on the noise-free observations $p_\mathrm{o}(k)$ of the scheduling signal. As a consequence, $B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ cannot be computed from the observations $\mathcal{D}_N = \{u(k), p(k), y(k)\}_{k=1}^N$ and thus it cannot directly be subtracted from the estimate $\hat{\theta}_\mathrm{IV}$.

Inspired by (14), the following corrected IV estimate is introduced:

$$\tilde{\theta}_\mathrm{CIV} = \hat{\theta}_\mathrm{IV} - B_\Delta(\tilde{\theta}_\mathrm{CIV}, \mathbf{p}, \mathbf{p}_\mathrm{o}), \quad (17)$$

where $B_\Delta(\tilde{\theta}_\mathrm{CIV}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ is (16) evaluated at $\tilde{\theta}_\mathrm{CIV}$ and the instruments $z(k)$ are chosen according to conditions **C1** and **C2**.

Algebraic manipulations of Eq. (17) lead to

$$\tilde{\theta}_\mathrm{CIV} = \underbrace{\Big(\frac{1}{N}\sum_{k=1}^{N} z(k)\left[\chi(k)\otimes\mathbf{p}_\mathrm{o}(k)\right]^\top\Big)^{-1}}_{R(\mathbf{p}_\mathrm{o})}\sum_{k=1}^{N}\frac{z(k)y(k)}{N}. \quad (18)$$

Property 1 *Let us assume that the following limit exists:*

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^{N} z(k)\left[\chi(k)\otimes\mathbf{p}_\mathrm{o}(k)\right]^\top\right)^{-1}. \quad (19)$$

Then, $\tilde{\theta}_\mathrm{CIV}$ is a consistent estimate of the true parameters θ_o, that is:

$$\lim_{N\to\infty}\tilde{\theta}_\mathrm{CIV} = \theta_\mathrm{o} \quad w.p.\ 1. \quad (20)$$

Proof: See Appendix 8.1. ■

Note that, unlike the structural bias $B_\Delta(\theta_\mathrm{o}, \mathbf{p}, \mathbf{p}_\mathrm{o})$, the correction term $B_\Delta(\tilde{\theta}_\mathrm{CIV}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ does not depend on the true parameter vector θ_o. However, $\tilde{\theta}_\mathrm{CIV}$ cannot be computed since $B_\Delta(\tilde{\theta}_\mathrm{CIV}, \mathbf{p}, \mathbf{p}_\mathrm{o})$ depends on the noise-free observations $p_\mathrm{o}(k)$ of the scheduling variable. In order to overcome such a problem, the estimate $\tilde{\theta}_\mathrm{CIV}$ (Eq. (18)) is modified by replacing each matrix $\Omega_k := z(k)\left[\chi(k)\otimes\mathbf{p}_\mathrm{o}(k)\right]^\top$ (for $k \in \mathbb{I}_1^N := \{1,\ldots,N\}$) with a matrix Ψ_k (constructed through the procedure described in Section 4.2) satisfying the following properties:

C3 for all $k \in \mathbb{I}_1^N$, the matrix Ψ_k does not depend on the noise-free scheduling signal $\mathbf{p}_\mathrm{o}(k)$, but only on the noisy observations $\mathbf{p}(k)$ and on the variance σ_η^2 of the noise corrupting these measurements.

C4 $\displaystyle\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\underbrace{z(k)\left[\chi(k)\otimes\mathbf{p}_\mathrm{o}(k)\right]^\top}_{\Omega_k} = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\Psi_k$ w.p. 1.

4.2 Construction of the matrices Ψ_k

Under the assumption that the variance σ_η^2 of the noise corrupting the scheduling signal observations is known, the matrices Ψ_k satisfying conditions **C3** and **C4** can be constructed through the following procedure (inspired by [17]):

(i) For each $k \in \mathbb{I}_1^N$, compute the analytical expression of the matrix $\mathbb{E}\{\Omega_k \mid y_N\}$, which denotes the conditional expected value of the matrix $\Omega_k = z(k)\left[\chi(k)\otimes\mathbf{p}_\mathrm{o}(k)\right]^\top$ given the output observation sequence $y_N = \{y(k)\}_{k=1}^N$. Note that, since each component of $\mathbf{p}_\mathrm{o}(k)$ is a polynomial in $p_\mathrm{o}(k)$, the entries of $\mathbb{E}\{\Omega_k \mid y_N\}$ are described by an affine combination of the monomials $p_\mathrm{o}(k), p_\mathrm{o}^2(k), p_\mathrm{o}^3(k), \ldots$.

(ii) For each monomial $p_\mathrm{o}(k), p_\mathrm{o}^2(k), p_\mathrm{o}^3(k)\ldots$, compute the coefficients $\alpha_1(k),\ \alpha_2(k), \alpha_3(k), \ldots$ satisfying the following properties:

- $p_\mathrm{o}(k) = \mathbb{E}\{p(k) + \alpha_1(k)\}$
- $p_\mathrm{o}^2(k) = \mathbb{E}\left\{p^2(k) + \alpha_2(k)\right\}$
- $p_\mathrm{o}^3(k) = \mathbb{E}\left\{p^3(k) + \alpha_3(k)\right\}$
- $\vdots$

The coefficients $\alpha_1, \alpha_2, \alpha_3, \ldots$ are computed in a recursive way. Consider first $\alpha_1(k)$. Then,

$$\begin{aligned} p_\mathrm{o}(k) &= \mathbb{E}\{p(k) + \alpha_1(k)\} = \\ &= \mathbb{E}\{p_\mathrm{o}(k) + \eta_\mathrm{o}(k) + \alpha_1(k)\} = \\ &= p_\mathrm{o}(k) + \mathbb{E}\{\alpha_1(k)\}. \end{aligned} \quad (21)$$

Equation (21) implies that $\mathbb{E}\{\alpha_1(k)\} = 0$ and $p_\mathrm{o}(k) = \mathbb{E}\{p(k)\}$. Thus, a possible choice of $\alpha(k)$ is

$\alpha_1(k) = 0$ for all $k \in \mathbb{I}_1^N$.

For α_2, we have:

$$\begin{aligned} p_o^2(k) &= \mathbb{E}\left\{p^2(k) + \alpha_2(k)\right\} = \\ &= \mathbb{E}\left\{(p_o(k) + \eta_o(k))^2 + \alpha_2(k)\right\} = \\ &= p_o^2(k) + \sigma_\eta^2 + \mathbb{E}\left\{\alpha_2(k)\right\}. \end{aligned} \tag{22}$$

Therefore, a possible choice of $\alpha_2(k)$ is $\alpha_2(k) = -\sigma_\eta^2$, which provides $p_o^2(k) = \mathbb{E}\left\{p^2(k)\right\} - \sigma_\eta^2$ for all $k \in \mathbb{I}_1^N$.

In case $d > 2$, the values of $\alpha_d(k)$ can be recursively computed on the basis of the (previously computed) unbiased estimates of $p_o(k), p_o^2(k), \ldots, p_o^{d-1}(k)$. For instance, for $d = 3$, a possible choice of $\alpha_3(k)$ can be computed as follows:

$$\begin{aligned} p_o^3(k) &= \mathbb{E}\left\{p^3(k) + \alpha_3(k)\right\} = \\ &= \mathbb{E}\left\{(p_o(k) + \eta_o(k))^3 + \alpha_3(k)\right\} = \\ &= p_o^3(k) + 3p_o(k)\sigma_\eta^2 + \mathbb{E}\left\{\alpha_3(k)\right\}. \end{aligned} \tag{23}$$

Eq. (23) implies that $\alpha_3(k)$ should be such that:

$$\mathbb{E}\left\{\alpha_3(k)\right\} = -3p_o(k)\sigma_\eta^2. \tag{24}$$

Since, based on the previous computation, $p_o(k) = \mathbb{E}\left\{p(k)\right\}$, from Eq. (24) we get:

$$\mathbb{E}\left\{\alpha_3(k)\right\} = -3\mathbb{E}\left\{p(k)\right\}\sigma_\eta^2 = \mathbb{E}\left\{-3p(k)\sigma_\eta^2\right\}. \tag{25}$$

This means that a possible choice for $\alpha_3(k)$ is $\alpha_3(k) = -3p(k)\sigma_\eta^2$. Thus, $p_o^3(k) = \mathbb{E}\left\{p^3(k) - 3p(k)\sigma_\eta^2\right\}$ for all $k \in \mathbb{I}_1^N$.

(iii) The matrix Ψ_k is obtained by replacing, in the analytical expression of $\mathbb{E}\left\{\Omega_k \mid y_N\right\}$, the monomials $p_o(k), p_o^2(k), p_o^3(k), \ldots$ with $p(k) + \alpha_1(k), p^2(k) + \alpha_2(k), p^3(k) + \alpha_3(k), \ldots$.

An illustrative example of the construction of the matrices Ψ_k is reported in Appendix 8.4.

Remark 1 *Note that the structure of the matrices Ψ_k only depends on the monomials used to parameterize the p-dependent coefficient functions a_i and b_j in* (7). ■

Remark 2 *The assumption that the noise η_o corrupting the scheduling signal has a Gaussian distribution is only required for the computation of the α coefficients in the construction of the matrices Ψ_k. In case of other distributions, the statistical moments of η_o are required to construct Ψ_k. Therefore, the Gaussian distribution assumption can be relaxed by assuming that the moment-generating function of η_o is known or can be estimated.* ■

Remark 3 *By construction, the matrix Ψ_k satisfies the following condition for all $k = 1, \ldots, N$:*

$$\mathbb{E}\left\{\tfrac{1}{N}\Omega_k \mid y_N\right\} = \mathbb{E}\left\{\tfrac{1}{N}\Psi_k \mid y_N\right\} \tag{26}$$

■

Property 2 *The computed matrices Ψ_k satisfy conditions **C3** and **C4** under the assumption that the amplitude of the measured output and of the scheduling signal is bounded, i.e., there exists a $G \in \mathbb{R}$ such that: $\|y\|_\infty < G$ and $\|p_o\|_\infty < G$, where $\|\cdot\|_\infty$ denotes the ℓ_∞-norm of a signal.*

Proof: See Appendix 8.2. ■

4.3 A consistent data-dependent parameter estimate

The matrix Ψ_k represents the expectation of Ω_k and it is constructed by using the available observations of $p(k)$ and the knowledge of the variance σ_η^2 of the noise corrupting these observations. Thus, the new corrected IV estimate $\hat{\theta}_{\mathrm{CIV}}$ is given by replacing in (18) the term $\Omega_k = z(k)\left[\chi(k) \otimes \mathbf{p}_o(k)\right]^\top$ with its expected value Ψ_k, i.e.,

$$\hat{\theta}_{\mathrm{CIV}} = \left(\frac{1}{N}\sum_{k=1}^{N}\Psi_k\right)^{-1}\sum_{k=1}^{N}\frac{z(k)y(k)}{N}. \tag{27}$$

Property 3 *Let us assume that the limit in Eq.* (19) *exists. Then, $\hat{\theta}_{\mathrm{CIV}}$ is a consistent estimate of the true parameters θ_o, that is:*

$$\lim_{N\to\infty}\hat{\theta}_{\mathrm{CIV}} = \theta_o \quad w.p.\ 1. \tag{28}$$

Proof: See Appendix 8.3. ■

4.4 Choice of the instrument $z(k)$

As discussed in the previous section, the instrument $z(k)$ used to compute $\hat{\theta}_{\mathrm{CIV}}$ (Eq. (27)) has to satisfy condition **C2**, which means that $z(k)$ has to be chosen by the user so that it is independent of the output noise realization $v_o(k)$. It is worth pointing out that, in case the scheduling signal measurements are not affected by noise, the optimal instrument $z(k)$ minimizing the asymptotic covariance matrix of the estimated parameters is given by the noise-free regressor ([9]), i.e.,

$$z(k) = \phi_o(k) = \chi_o(k) \otimes \mathbf{p}_o(k). \tag{29}$$

Consequently, the instrument chosen to address the bias-corrected IV solution is inspired by the instrument proposed in the LPV identification setting with noise-free scheduling parameter. More precisely, the variable $z(k)$ is chosen to be maximally correlated with the noise-free part of the sample $\phi_o(k)$. Note that neither $\chi_o(k)$ nor $\mathbf{p}_o(k)$ are available in practice. Nevertheless, since $z(k)$

is not required to be uncorrelated with the noise $\eta_\mathrm{o}(k)$ affecting the scheduling signal measurements, the following instrument can be used to approximate the optimal instrument in (29):

$$z(k) = \hat{\chi}(k) \otimes \mathbf{p}(k), \tag{30}$$

where

$$\hat{\chi}(k) = [-\hat{y}(k-1)\ \dots\ -\hat{y}(k-n_\mathrm{a})\ \ u(k)\dots\ u(k-n_\mathrm{b})]^\top,$$

with $\hat{y}(k-1), \dots, \hat{y}(k-n_\mathrm{a})$ being an approximation of the noise-free output samples $y_\mathrm{o}(k-1), \dots, y_\mathrm{o}(k-n_\mathrm{a})$ obtained, for instance, by simulating an estimated model of the system. This choice of the instrument resembles the widely used IV solution for linear regression [21,10]. The following iterative algorithm can be implemented in order to mitigate the effect of the estimated noiseless signals on the IV scheme and hence "maximize" the accuracy of the IV solution by iteratively refining the instruments.

Algorithm 1 Recursive IV scheme

1: set $\tau \leftarrow 0$.
2: compute the LS parameter estimate $\hat{\theta}_\mathrm{LS}$ via (11), resulting in the biased model estimate $\mathcal{M}^{(0)}$.
3: **repeat**
4: set $\tau \leftarrow \tau + 1$
5: use $\mathcal{M}^{(\tau-1)}$ and the sequence $\{u(k), p(k)\}_{k=1}^N$ to generate, by simulation, $\{\hat{y}^{(\tau)}(k)\}_{k=1}^N$.
6: calculate $\{z(k)\}_{k=1}^N$ via (30) using $\{u(k), \hat{y}^{(\tau)}(k), p(k)\}_{k=1}^N$.
7: estimate $\hat{\theta}_\mathrm{CIV}^{(\tau)}$ via (27), resulting in the model estimate $\mathcal{M}^{(\tau)}$.
8: **until** $\hat{\theta}_\mathrm{CIV}^{(\tau)}$ has converged or the maximum number of iterations is reached.
9: **return** parameter estimate $\hat{\theta}_\mathrm{CIV}^{(\tau)}$.

5 Estimation with unknown noise variance

The application of the proposed identification approach is limited, in principle, to the case when the variance σ_η^2 of the noise $\eta_\mathrm{o}(k)$ affecting the scheduling signal measurements is known. In this section, we present an approach to extend the applicability of the developed identification procedure to the case when σ_η^2 is not *a-priori* available.

In order to estimate the noise variance σ_η^2, an additional equation relating σ_η^2 and the system parameters θ_o is required. Inspired by the papers [29,8], where bias-eliminated least-squares algorithms for identification of LTI systems in the EIV framework are discussed, let us introduce the following augmented vectors:

$$\begin{aligned}
\underline{\theta_\mathrm{o}} &= \begin{bmatrix}\theta_\mathrm{o}^\top & 0\end{bmatrix}^\top, \\
\underline{z}(k) &= \begin{bmatrix}z^\top(k) & u(k-(n_\mathrm{b}+1))p(k)\end{bmatrix}^\top = \\
&= \begin{bmatrix}(\hat{\chi}(k)\otimes\mathbf{p}(k))^\top & u(k-(n_\mathrm{b}+1))p(k)\end{bmatrix}^\top,
\end{aligned}$$

and the augmented matrices $\underline{\Psi}_k$ (with $k \in \mathbb{I}_1^N$) which are constructed in order to satisfy the condition:

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N \underline{z}(k)\left[\,(\chi(k)\otimes\mathbf{p}_\mathrm{o}(k))^\top\ \ u(k-(n_\mathrm{b}+1))p(k)\right] \\
= \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N \underline{\Psi}_k \quad \text{w.p. } 1.
\end{aligned}$$

Note that $\underline{\Psi}_k$ can be constructed through the same procedure used to define the matrix Ψ_k and described in Section 4.2. Consider now the term $\frac{1}{N}\sum_{k=1}^N \underline{z}(k)y(k)$ which is guaranteed to satisfy the following property:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N \underline{z}(k)y(k) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N \underline{\Psi}_k\underline{\theta_\mathrm{o}}. \quad \text{w.p. } 1. \tag{31}$$

Eq. (31) follows from the fact that

$$\begin{aligned}
y(k) &= (\chi_\mathrm{o}(k)\otimes\mathbf{p}_\mathrm{o}(k))^\top\theta_\mathrm{o} + v_\mathrm{o}(k) \\
&= \left[(\chi_\mathrm{o}(k)\otimes\mathbf{p}_\mathrm{o}(k))^\top\ \ u(k-(n_\mathrm{b}+1))p_\mathrm{o}(k)\right]\underline{\theta_\mathrm{o}} + v_\mathrm{o}(k).
\end{aligned}$$

Eq. (31) provides an (asymptotic) additional relation between the noise variance σ_η^2 (which appears in the construction of $\underline{\Psi}_k$) and the true system parameters θ_o. Specifically, an estimate of the noise variance σ_η^2 and the system parameters θ_o can be computed by combing Eq. (27) and (31) for finite N. This leads to the set of bilinear equations in the variables (θ, σ^2):

$$\theta = \left(\frac{1}{N}\sum_{k=1}^N \Psi_k(\sigma^2)\right)^{-1}\sum_{k=1}^N \frac{z(k)y(k)}{N}, \tag{32a}$$

$$\left[\frac{1}{N}\sum_{k=1}^N \underline{z}(k)y(k)\right]_{n_\theta+1} = \left[\frac{1}{N}\sum_{k=1}^N \underline{\Psi}_k(\sigma^2)\begin{bmatrix}\theta\\0\end{bmatrix}\right]_{n_{\theta+1}}, \tag{32b}$$

where the symbol $\Psi_k(\sigma^2)$ has been used to highlight that the matrix Ψ_k depends on the chosen noise variance σ^2. In (32b), $[\,\centerdot\,]_n$ denotes the n-th component of a vector. This means that, for finite N, (32b) is the last equation of system of equations in (31). Indeed, as $N\to\infty$, the pair $(\theta_\mathrm{o}, \sigma_\eta^2)$ becomes a solution of the set of equations in (32). The solution of (32) can be easily computed through a grid over different noise variances, as described in the following:

(i) Generate a set $\{\sigma_i^2\}_{i=1}^M$ of M equally-spaced points in the interval $[0,\ \overline{\sigma}_\eta^2]$, with $\overline{\sigma}_\eta^2$ being an upper bound of the expected noise variance σ_η^2.

(ii) For all the points $\{\sigma_i^2\}_{i=1}^M$ generated at step (i), compute $\hat{\theta}^{(i)}$ through Eq. (32a) with $\sigma^2 = \sigma_i^2$, i.e.,

$$\hat{\theta}^{(i)} = \left(\frac{1}{N}\sum_{k=1}^{N}\Psi_k(\sigma_i^2)\right)^{-1}\sum_{k=1}^{N}\frac{z(k)y(k)}{N}.$$

(iii) Among the M parameter vectors $\hat{\theta}^{(i)}$ computed at step (ii), the estimate $\hat{\theta}_{\mathrm{CIV}}$ is given by the vector $\hat{\theta}^{(i)}$ which satisfies Eq. (32) with the least error.

Remark 4 *The identification approach developed in the paper can be extended, in principle, to the case of multi-dimensional scheduling variable $p_{\mathrm{o}}(k)$ (i.e., $n_{\mathrm{p}} > 1$), at the cost of increasing the computational complexity. As a matter of fact, increasing the dimension of the scheduling variable $p_{\mathrm{o}}(k)$ requires to consider a larger number of monomials to characterize the p-dependent coefficient functions $a_i(p(k))$ and $b_j(p(k))$ in* (7)*. As a consequence, the construction of the matrices Ψ_k (and $\underline{\Psi}_k$) requires to compute a larger number of coefficients $\alpha_i(k)$. Furthermore, in order to solve equations* (32)*, a gridding of an n_{p}-dimensional box should be performed. Thus, in case the variance of the noise corrupting the scheduling observations is unknown, the computational complexity of the developed algorithm increases exponentially with n_{p}.* ■

6 Simulation Example

The purpose of this section is to demonstrate the performance of the developed identification approach through a numerical example and to show how the noise on the measurements of the scheduling signal can deteriorate the estimate of the system parameters if the bias-correction scheme is not applied.

6.1 Data-generating system and model structure

The considered data-generating system $\mathcal{S}_{\mathrm{o}}$ is described by (1)-(3) with

$$a_1^{\mathrm{o}}(p_{\mathrm{o}}(k)) = 1 - 0.5p_{\mathrm{o}}(k) - 0.3p_{\mathrm{o}}^2(k), \quad (33a)$$

$$a_2^{\mathrm{o}}(p_{\mathrm{o}}(k)) = 0.5 - 0.7p_{\mathrm{o}}(k) - 0.5p_{\mathrm{o}}^2(k), \quad (33b)$$

$$b_1^{\mathrm{o}}(p_{\mathrm{o}}(k)) = 0.5 - 0.4p_{\mathrm{o}}(k) + 0.1p_{\mathrm{o}}^2(k), \quad (33c)$$

$$b_2^{\mathrm{o}}(p_{\mathrm{o}}(k)) = 0.2 - 0.3p_{\mathrm{o}}(k) - 0.2p_{\mathrm{o}}^2(k). \quad (33d)$$

In this example, we will focus on the comparison between the developed bias-correction approach and the IV based approach proposed in [3], hence the data-generating system is assumed to belong to the chosen model class. The following LPV model structure is used:

$$y(k) = -\sum_{i=1}^{2} a_i(p(k),\theta)y(k-i) + \sum_{j=1}^{2} b_j(p(k),\theta)u(k-j) + \varepsilon(k).$$

The functions a_i and b_j, are parameterized as follows:

$$a_1(p(k),\theta) = a_{1,0} + a_{1,1}p(k) + a_{1,2}p^2(k), \quad (34a)$$

$$a_2(p(k),\theta) = a_{2,0} + a_{2,1}p(k) + a_{2,2}p^2(k), \quad (34b)$$

$$b_1(p(k),\theta) = b_{1,0} + b_{1,1}p(k) + b_{1,2}p^2(k), \quad (34c)$$

$$b_2(p(k),\theta) = b_{2,0} + b_{2,1}p(k) + b_{2,2}p^2(k). \quad (34d)$$

The input $u(k)$ is taken as a white-noise sequence with uniform distribution $\mathcal{U}(0,1)$ and length $N = 4000$, while the noise-free scheduling signal $p_{\mathrm{o}}(k)$ is given by:

$$p_{\mathrm{o}}(k) = 0.2 + 0.4\sin(0.035\pi k) + \delta(k), \quad (35)$$

with $\delta(k)$ being a random variable with uniform distribution $\mathcal{U}(0,0.3)$. The noises $v_{\mathrm{o}}(k)$ and $\eta_{\mathrm{o}}(k)$ corrupting the output and the scheduling signal measurements, respectively, are white Gaussian noise processes with standard deviation $\sigma_v = 0.08$ and $\sigma_\eta = 0.10$. This corresponds to the following signal-to-noise-ratios:

$$\mathrm{SNR}_y = 10\log\frac{\sum_{k=1}^{N}(y_{\mathrm{o}}(k)-\bar{y}_{\mathrm{o}})^2}{\sum_{k=1}^{N}(v_{\mathrm{o}}^2(k))} = 17\mathrm{dB},$$

$$\mathrm{SNR}_p = 10\log\frac{\sum_{k=1}^{N}(p_{\mathrm{o}}(k)-\bar{p}_{\mathrm{o}})^2}{\sum_{k=1}^{N}(\eta_{\mathrm{o}}^2(k))} = 21\mathrm{dB},$$

with $\overline{y}_{\mathrm{o}}$ and $\overline{p}_{\mathrm{o}}$ denoting the mean value of the noise-free output and scheduling signal, respectively. Note that, since $v_{\mathrm{o}}(k)$ is white, the data-generating system (1) is an LPV system with an output-error type of model structure. In order to empirically study the statistical properties of the developed bias-correction scheme, a Monte Carlo study with $N_{\mathrm{MC}} = 100$ runs with new noise and input realizations in each run is carried out.

6.2 Obtained results

The model parameters θ are computed with the following three approaches:

- IV based approach ([3]);
- Bias-corrected IV identification approach with known noise variance σ_η^2.
- Bias-corrected IV identification approach with unknown noise variance σ_η^2. The parameter estimate $\hat{\theta}_{\mathrm{CIV}}$ is then computed through the procedure discussed at the end of Section 5, by setting the upper bound $\bar{\sigma}_\mu^2$ on the noise variance equal to 0.2, and by gridding the interval $[0\ \ \bar{\sigma}_\mu^2]$ with $M = 100$ equidistant points.

The obtained results are reported in Table 1, which shows the average of the estimated parameters and their

Table 2
Validation results of the estimated models. Average and standard deviation over the Monte Carlo simulation of the best fit rate (BFR) and of the mean squared error (MSE).

	BFR (mean)	BFR (std)	MSE (mean)	MSE (std)
IV estimate	45%	21%	$1.78{\cdot}10^{-2}$	$7.10{\cdot}10^{-3}$
Bias-corrected IV	96%	2%	$1.04{\cdot}10^{-4}$	$9.04{\cdot}10^{-6}$

standard deviation over the 100 Monte Carlo runs. Table 1 shows that the estimate of the system parameters is very sensitive to the noise on the scheduling signal measurements. As a matter of fact, although the SNR on the scheduling signal measurements is relatively moderate (21 dB), the IV based identification approach in [3] provides a biased parameter estimate, with a bias which, in some cases, has the same magnitude as the true value of the parameters (see, e.g., the estimate of the parameters $a_{1,1}$, $a_{2,2}$ and $b_{1,2}$). On the other hand, in line with the theory, the bias-corrected IV method proposed in the paper provides an unbiased estimate of the true system parameters $\theta_{\rm o}$, also when the noise variance is unknown.

The performance of the IV based approach [3] and the bias-corrected IV identification approach with unknown noise variance σ_η^2 is also tested on a noiseless validation data sequence of length $N_{\rm val} = 200$. The true output $y_{\rm o}$ and the simulated output sequences $\hat{y}$ of the estimated models are plotted in Fig. 1a-b, while the error between the true output $y_{\rm o}(k)$ and $\hat{y}(k)$ is plotted in Fig. 1c-d. The BFR and the *mean squared error* (MSE), defined as

$$\text{MSE} = \frac{1}{N}\sum_{k=1}^{N}\left(y_{\rm o}(k) - \hat{y}(k)\right)^2, \tag{36}$$

$$\text{BFR} = \max\left\{1 - \frac{\|y_{\rm o}(k) - \hat{y}(k)\|_2}{\|y_{\rm o}(k) - \overline{y}_{\rm o}\|_2}, 0\right\} \cdot 100\%, \tag{37}$$

and computed on the simulated response $\hat{y}$ of the estimated models, are reported in Table 2. The obtained results show that bias-corrected algorithm significantly outperforms the standard IV approach.

7 Conclusion

In this paper, a solution has been proposed for the challenging problem of LPV system identification under the realistic assumption of noise-corrupted measurements of scheduling variable. The solution relies on an innovative combination of IV-based methods and bias correction. While the IV optimization ensures unbiased estimates under the sole condition that the noise corrupting the output observations is zero mean, the bias correction enables the relaxation of the usual assumption of noise-free scheduling variables. The analysis has been driven in a context where the scheduling noise is a white Gaussian process and the parametrization of the scheduling dependent coefficient functions is based on polynomial basis. These assumptions are realistic in many real applications where the scheduling noise comes from measurements and the nonlinearities characterizing the scheduling dependencies are smooth enough to be well approximated by polynomial functions. Furthermore, the assumption that the noise corrupting the scheduling signal observations is Gaussian distributed can be relaxed by assuming that the moment-generating function of the noise is known. Hence, this paper presents one of the first solution towards EIV problems in LPV settings. The variance of the estimates can be reduced by extending the developed method to *output-error* and *Box-Jenkins* noise models via the results of [9]. Finally, in trying to come closer to more realistic assumptions, this work is intended to be extended towards scheduling noise distribution uncertainties and non-parametric estimation of nonlinearities.

8 Appendix

8.1 Proof of Property 1

In order to prove Property 1, the following necessary results coming from the application of the results presented in [3] are first reported:

$$\lim_{N\to\infty}\sum_{k=1}^{N}\frac{1}{N}z(k)[(\chi_{\rm o}(k)-\chi(k))\otimes\mathbf{p}_{\rm o}(k)]^\top = 0, \tag{38a}$$

$$\lim_{N\to\infty}\sum_{k=1}^{N}\frac{1}{N}z(k)v_{\rm o}(k) = 0, \tag{38b}$$

with probability 1. Eq. (38a) and (38b) hold since $\chi_{\rm o}(k)$–$\chi(k)$ and $v_{\rm o}(k)$ are zero mean and both of them are not correlated with the instrument $z(k)$.

Property 1 is now proven. Substitution of Eq. (13) in (18) leads to:

$$\tilde{\theta}_{\rm CIV} = R(\mathbf{p}_{\rm o})^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)[(\chi_{\rm o}(k)-\chi(k))\otimes\mathbf{p}_{\rm o}(k)]^\top\theta_{\rm o} \tag{39a}$$

$$+ R(\mathbf{p}_{\rm o})^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)v_{\rm o}(k) \tag{39b}$$

$$+ R(\mathbf{p}_{\rm o})^{-1}\underbrace{\sum_{k=1}^{N}\frac{1}{N}z(k)\left[\chi(k)\otimes\mathbf{p}_{\rm o}(k)\right]^\top}_{R(\mathbf{p}_{\rm o})}\theta_{\rm o}. \tag{39c}$$

From (38a) and (38b), Eq. (39a) and (39b) converge to zero as $N \to \infty$. This implies (20).

8.2 Proof of Property 2

The matrices Ψ_k satisfy condition **C3** by construction. In order to prove that Ψ_k satisfy also condition **C4**, the

Table 1
Mean and standard deviation of the estimates of the model parameters over the 100 Monte Carlo runs.

		True value	IV-estimate [3]	bias-corrected estimate $\hat{\theta}_{\mathrm{CIV}}$ (σ_η^2 known)	bias-corrected estimate $\hat{\theta}_{\mathrm{CIV}}$ (σ_η^2 unknown)
$a_{1,0}$	mean	1	0.8904	0.9987	1.0006
	std	–	0.0163	0.0171	0.0182
$a_{1,1}$	mean	−0.5	−1.1307	−0.4936	−0.4922
	std	–	0.1302	0.1461	0.1493
$a_{1,2}$	mean	−0.3	−0.4929	−0.3127	−0.3143
	std	–	0.1494	0.1657	0.1699
$a_{2,0}$	mean	0.5	0.4276	0.4997	0.5005
	std	–	0.0065	0.0080	0.0083
$a_{2,1}$	mean	−0.7	−1.0638	−0.7015	−0.6882
	std	–	0.0803	0.0943	0.0985
$a_{2,2}$	mean	−0.5	−0.0261	−0.5134	−0.5169
	std	–	0.1098	0.1134	0.1148
$b_{1,0}$	mean	0.5	0.4906	0.5006	0.5006
	std	–	0.0035	0.0028	0.0031
$b_{1,1}$	mean	−0.4	−0.3316	−0.4012	−0.4038
	std	–	0.0240	0.0267	0.0271
$b_{1,2}$	mean	0.1	0.0423	0.0995	0.1004
	std	–	0.0259	0.0316	0.0378
$b_{2,0}$	mean	0.2	0.1482	0.2002	0.1997
	std	–	0.0084	0.0069	0.0070
$b_{2,1}$	mean	−0.3	−0.5648	−0.2948	−0.2879
	std	–	0.0382	0.0432	0.0476
$b_{2,2}$	mean	−0.2	−0.2610	−0.2035	−0.2071
	std	–	0.0513	0.0643	0.0672

following necessary lemma coming from a direct application of the Ninness's strong law of large numbers [15] is first presented.

Lemma 1 ***Ninness' strong law of large numbers [15].*** *Let $\{\nu(t)\}$ be a sequence of random variables with arbitrary correlation structure (not necessarily stationary) that is characterized by the existence of a finite value C such that*

$$\sum_{t=1}^{N}\sum_{s=1}^{N}\mathbb{E}\{\nu(t)\nu(s)\} < CN. \tag{40}$$

Then,

$$\frac{1}{N}\sum_{t=1}^{N}\nu(t) \xrightarrow{a.s.} 0 \quad as \quad N\to\infty. \tag{41}$$

■

Property 2 will now be proven. Let us denote with $[\,\centerdot\,]_{i,j}$ the (i,j)-th entry of a matrix. Let us define the variable $\nu_{i,j}(k)$ as follows:

$$\left[\Psi_k - z(k)\left(\chi(k)\otimes \mathbf{p}_{\mathrm{o}}(k)\right)^\top\right]_{i,j} = \nu_{i,j}(k). \tag{42}$$

By construction of the matrix Ψ_k (see Remark 3), we have:

$$\mathbb{E}\{\nu_{i,j}(k) \mid y_N\} = 0 \quad \text{for all } k = 1,\ldots,N. \tag{43}$$

Moreover, since the noise process η_{o} corrupting the scheduling signal observations is white, we have:

$$\mathbb{E}\{\nu_{i,j}(k)\nu_{i,j}(t) \mid y_N\} = 0 \quad \text{for all } k,t\geq 0, k\neq t. \tag{44}$$

Note also that, since the output signal $y(k)$ and the scheduling variable $p_{\mathrm{o}}(k)$ are assumed to be bounded, then $\mathbb{E}\{\nu_{i,j}(k)\nu_{i,j}(k) \mid y_N\}$ is bounded for any index-pair $k>0$, i.e., there exists a positive constant $G_{i,j}$ such that

$$\mathbb{E}\{\nu_{i,j}(k)\nu_{i,j}(k) \mid y_N\} < G_{i,j} \quad \text{for all } k>0. \tag{45}$$

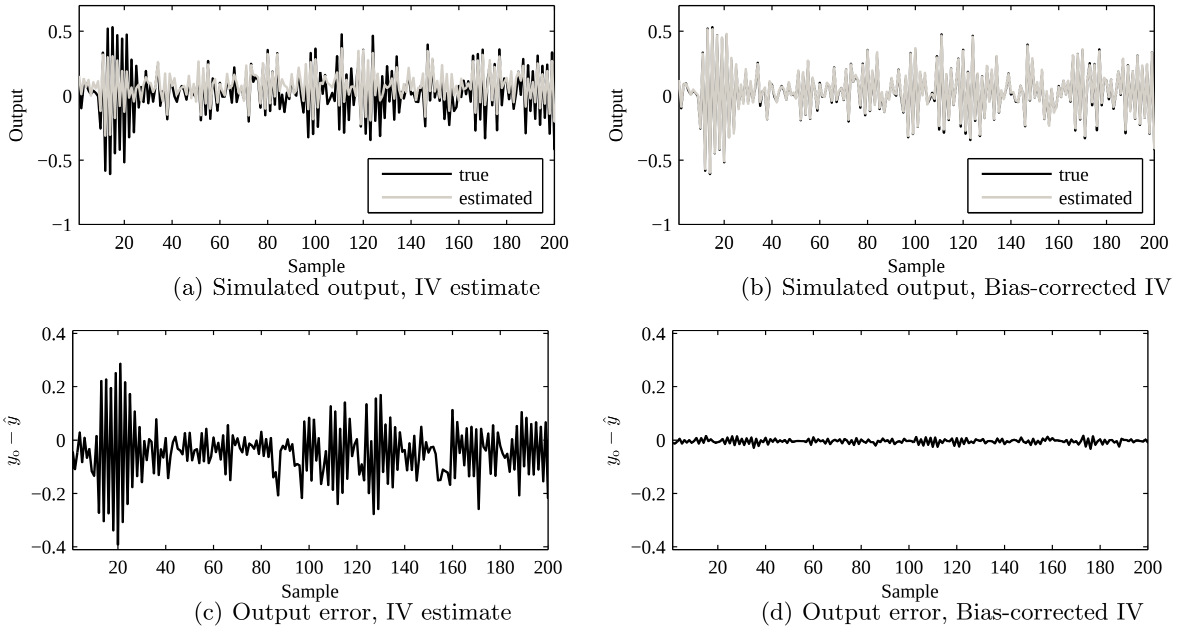

Fig. 1. Validation data based comparison of the true output $y_{\rm o}(k)$ (black line) with the simulated output sequence $\hat{y}(k)$ (gray line) of one model estimated from the Monte Carlo simulation by the IV-based approach and the bias-corrected IV approach.

Based on the above considerations, we have:

$$\sum_{k=1}^{N}\sum_{t=1}^{N}\mathbb{E}\{\nu_{i,j}(k)\nu_{i,j}(t)\mid y_N\} = \\ = \sum_{k=1}^{N}\mathbb{E}\left\{\nu_{i,j}(k)\nu_{i,j}(k)\mid y_N\right\} < G_{i,j}N. \quad (46)$$

Therefore, from Lemma 1, it follows

$$\frac{1}{N}\sum_{t=1}^{N}\nu_{i,j}(k) \xrightarrow{a.s.} 0 \quad \text{as } N\to\infty,$$

or equivalently, (see Eq. (42))

$$\lim_{N\to\infty}\frac{1}{N}\left[\sum_{k=1}^{N}\Psi_k\right]_{i,j} = \\ = \lim_{N\to\infty}\frac{1}{N}\left[\sum_{k=1}^{N}z(k)\left(\chi(k)\otimes\mathbf{p}_{\rm o}(k)\right)^\top\right]_{i,j} \quad \text{w.p. 1.}$$

8.3 Proof of Property 3

Property 3 is proven similarly to Property 1. More specifically, let us decompose the estimate $\hat{\theta}_{\rm CIV}$ as follows

$$\hat{\theta}_{\rm CIV} = \left(\sum_{k=1}^{N}\frac{1}{N}\Psi_k\right)^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)[(\chi_{\rm o}(k)-\chi(k))\otimes\mathbf{p}_{\rm o}(k)]^\top\theta_{\rm o} \quad (47a)$$

$$+\left(\sum_{k=1}^{N}\frac{1}{N}\Psi_k\right)^{-1}\sum_{k=1}^{N}\frac{1}{N}z(k)v_{\rm o}(k) \quad (47b)$$

$$+\left(\sum_{k=1}^{N}\frac{1}{N}\Psi_k\right)^{-1}\sum_{k=1}^{N}\frac{1}{N}\underbrace{z(k)\left[\chi(k)\otimes\mathbf{p}_{\rm o}(k)\right]^\top}_{\Omega_k}\theta_{\rm o}. \quad (47c)$$

From (38a) and (38b), Eq. (47a) and (47b) converge to zero as $N\to\infty$. Furthermore, since the matrices Ψ_k are constructed in order to satisfy condition **C4**, the term in (47c) converges to $\theta_{\rm o}$ as $N\to\infty$. This implies (28).

8.4 Construction of the matrices Ψ_k: an illustrative example

Consider an LPV data-generating system $\mathcal{S}_{\rm o}$ of the form:

$$\begin{aligned} y_{\rm o}(k) =& -a_{1,2}p_{\rm o}^2(k)y_{\rm o}(k-1)+b_{0,2}p_{\rm o}^2(k)u(k),\\ y(k) =& y_{\rm o}(k)+v_{\rm o}(k),\\ p(k) =& p_{\rm o}(k)+\eta_{\rm o}(k).\end{aligned}$$

Then, according to the definitions used in the paper, we have:

$$\chi(k) = \left[-y(k-1) \;\; u(k)\right]^\top,$$
$$\mathbf{p}(k) = p^2(k), \;\; \mathbf{p}_\text{o}(k) = p_\text{o}^2(k),$$

$$\chi(k) \otimes \mathbf{p}_\text{o}(k) = \left[-y(k-1)p_\text{o}^2(k) \quad u(k)p_\text{o}^2(k)\right]^\top,$$

and

$$z(k) = \left[-\hat{y}(k-1)p^2(k) \quad u(k)p^2(k)\right]^\top,$$

where $\hat{y}(k-1)$ is the instrument and it is chosen to be uncorrelated with the noise process $v_\text{o}(k)$.

Then, the matrix $z(k)\left(\chi(k) \otimes \mathbf{p}_\text{o}(k)\right)^\top$ is given by

$$\begin{pmatrix} \hat{y}(k-1)y(k-1)p^2(k)p_\text{o}^2(k) & -\hat{y}(k-1)u(k)p^2(k)p_\text{o}^2(k) \\ -u(k)y(k-1)p^2(k)p_\text{o}^2(k) & u^2(k)p^2(k)p_\text{o}^2(k) \end{pmatrix}.$$

and $\mathbb{E}\left\{z(k)\left(\chi(k) \otimes \mathbf{p}_\text{o}(k)\right)^\top \mid y_N\right\}$ is given by

$$\left(p_\text{o}^4(k) + p_\text{o}^2(k)\sigma_\eta^2\right)\begin{pmatrix} \hat{y}(k-1)y(k-1) & -\hat{y}(k-1)u(k) \\ -u(k)y(k-1) & u^2(k) \end{pmatrix}.$$

Since

$$p_\text{o}^2(k) = \mathbb{E}\left\{p^2(k) - \sigma_\eta^2\right\},$$

and

$$p_\text{o}^4(k) = \mathbb{E}\left\{p^4(k) + 3\sigma_\eta^4 - 6\sigma_\eta^2 p^2(k)\right\},$$

then, the matrix Ψ_k, with $k = 1, \ldots, N$, is given by

$$\left(p^4(k) - 5p^2(k)\sigma_\eta^2 + 2\sigma_\eta^4\right)\begin{pmatrix} \hat{y}(k-1)y(k-1) & -\hat{y}(k-1)u(k) \\ -u(k)y(k-1) & u^2(k) \end{pmatrix}.$$

References

[1] H. Abbas and H. Werner. An instrumental variable technique for open-loop and closed-loop identification of input-output LPV models. In *Proc. of the European control conference*, pages 2646–2651, Budapest, Hungary, 2009.

[2] B. A. Bamieh and L. Giarré. Identification of linear parameter-varying models. *International Journal of Robust Nonlinear Control*, 12(9):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, South Korea, 2008.

[4] V. Cerone, D. Piga, and D. Regruto. A convex relaxation approach to set-membership identication of LPV systems. *Automatica*, 49(9):2853–2859, 2013.

[5] M. Corno, M. Tanelli, S. M. Savaresi, and L. Fabbri. Design and validation of a gain-scheduled controller for the electronic throttle body in ride-by-wire racing motorcycles. *IEEE Trans. on Control Systems Technology*, 19(1):18–30, 2011.

[6] F. Felici, J. W. Van Wingerden, and M. Verhaegen. Subspace identification of MIMO LPV systems using a periodic scheduling sequence. *Automatica*, 43:1684–1697, 2007.

[7] S. M. Hashemi, H.S. Abbas H. S., and H. Werner. Low-complexity linear parameter-varying modeling and control of a robotic manipulator. *Control Engineering Practice*, 20(3):248–257, 2012.

[8] M. Hong, T. Söderström, and W. X. Zheng. Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables systems. *Automatica*, 43(9):1590–1596, 2007.

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

[10] L. Ljung. *System Identification, Theory for the User.* Prentince Hall, Upper Saddle River, 1999.

[11] P. Lopes dos Santos, J. A. Ramos, and J. L. Martins de Carvalho. Subspace identification of linear parameter-varying systems with innovation-type noise models driven by general inputs and a measurable white noise time-varying parameter vector. *Int. J. of Systems Science*, 39(9):897–911, 2008.

[12] B. Lu, F. Wu, and S.W. Kim. Switching LPV control of an F-16 aircraft via controller state reset. *IEEE Trans. on Control Systems Technology*, 14(2):267–277, 2006.

[13] A. Marcos and G.J. Balas. Development of linear-parameter-varying models for aircraft. *Journal of Guidance Control and Dynamics*, 27(2):218–228, 2004.

[14] G. Mercère, H. Palsson, and T. Poinot. Continuous-time linear parameter-varying identification of a cross flow heat exchanger: A local approach. *IEEE Trans. on Control Systems Technology*, 19(1):64–76, 2011.

[15] B. Ninness. Strong laws of large numbers under weak assumptions with application. *IEEE Transactions on Automatic Control*, 45(11):2117–2122, 2000.

[16] C. Novara, F. Ruiz, and M. Milanese. Direct identification of optimal SM-LPV filters and application to vehicle yaw rate estimation. *IEEE Trans. on Control Systems Technology*, 19(1):5–17, 2011.

[17] D. Piga and R. Tóth. Bias-corrected estimators for nonlinear systems with output-error typemodel structures. *Automatica*, In press. DOI: 10.1016/j.automatica.2014.07.021.

[18] E. Prempain, I. Postlethwaite, and A Benchaib. A Linear Parameter Variant $\mathcal{H}_\infty$ control design for an induction motor. *Control Engineering Practice*, 10(6):633–644, 2002.

[19] J. S. Shamma and M. Athans. Analysis of gain scheduled control for nonlinear plants. *IEEE Transactions on Automatic Control*, 35(8):898–907, 1990.

[20] T. Söderström. Errors-in-variables methods in system identification. *Automatica*, 43(6):939–958, 2007.

[21] T. Söderström and P. Stoica. *Instrumental variable methods for system identification*, volume 161. Springer-Verlag Berlin, 1983.

[22] M. Sznaier and C. Mazzarro. An LMI approach to control-oriented identification and model (in)validation of LPV systems. *IEEE Transaction on Automatic Control*, 48(9):1619–1624, 2003.

[23] R. Tóth. *Modeling and identification of linear parameter-varying systems.* Lecture Notes in Control and Information Sciences, Vol. 403, Springer, Heidelberg, 2010.

[24] R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof. Asymptotically optimal orthonormal basis functions for LPV system identification. *Automatica*, 45(6):1359–1370, 2010.

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

[26] V. Verdult and M. Verhaegen. Kernel methods for subspace identification of multivariable LPV and bilinear systems. *Automatica*, 41:1557–1565, 2005.

[27] S. K. Wattamwar, S. Weiland, and T. Backx. Identification of low-order parameter-varying models for large-scale systems. *Journal of Process Control*, 20(2):158–172, 2010.

[28] X. Wei and L. del Re. Gain scheduled H_∞ control for air path systems of diesel engines using LPV techniques. *IEEE Trans. on Control Systems Technology*, 15(3):406–415, 2007.

[29] W. X. Zheng. Transfer function estimation from noisy input and output data. *International Journal of Adaptive Control and Signal Processing*, 12(4):365–380, 1998.