A Bias-Correction Method for Closed-Loop Identification of Linear Parameter-Varying Systems ⋆

Manas Mejari [a], Dario Piga [b], Alberto Bemporad [a]

[a] IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy

[b] IDSIA Dalle Molle Institute for Artificial Intelligence, SUPSI-USI, 6928 Manno, Switzerland

Abstract

Due to safety constraints and unstable open-loop dynamics, system identification of many real-world processes often require gathering data from closed-loop experiments. In this paper, we present a bias-correction scheme for closed-loop identification of *Linear Parameter-Varying Input-Output* (LPV-IO) models, which aims at correcting the bias caused by the correlation between the input signal exciting the process and output noise. The proposed identification algorithm provides a consistent estimate of the open-loop model parameters when both the output signal and the scheduling variable are corrupted by measurement noise. The effectiveness of the proposed methodology is tested in two simulation case studies.

Key words: Closed-loop identification; Bias-corrected least squares; LPV systems.

1 Introduction

Many real world systems must be identified based on data collected from closed-loop experiments. This is typical for open-loop unstable plants, where a feedback controller is necessary to perform the experiments, and in many applications in which a controller is needed to keep the system at certain operating points. Safety, performance, and economic requirements are further motivations to operate in closed-loop.

From the system identification point of view, one of the main issues which makes identification from closed-loop experiments more challenging than in the open-loop setting is due to the correlation between the plant input and output noise. If such a correlation is not properly taken into account, approaches that work in open loop may fail when closed-loop data is used [14]. Several remedies have been proposed in the literature to overcome this problem, especially for the *Linear Time-Invariant* (LTI) case (see [8] and [9] for an overview). These approaches can be classified in: *direct methods*, which neglect the existence of the feedback loop and apply prediction error methods directly on the input-output data after properly parametrizing the noise model; *indirect methods*, where the closed-loop system is identified and the model of the open-loop plant is then extracted exploiting the knowledge of the controller and of the feedback structure; *joint input-output methods*, which treat the measured input and output signals as the outputs of an augmented multi-variable system driven by external disturbances. The model of the open-loop process is then extracted based on the estimate of different transfer functions of the augmented system. Unlike indirect methods, an exact knowledge of the controller is not needed.

Unfortunately, the extension of these approaches to the *Linear Parameter-Varying* (LPV) case is not straightforward, mainly because the classical theoretical tools which are commonly used in closed-loop LTI identification no longer hold in the LPV setting [20], such as transfer functions and commutative properties of operators. Therefore, only few contributions addressing identification of LPV systems from closed-loop data are available in the literature. A subspace method, which can be applied both for open- and closed-loop identification of LPV models was proposed in [22]. The idea of this method is to construct a matrix approximating the prod-

⋆ This paper was not presented at any IFAC meeting. Corresponding author Manas Mejari. This work was partially supported by the European Commission under project H2020-SPIRE-636834 “DISIRE - Distributed In-Situ Sensors Integrated into Raw Material and Energy Feedstock” (`http://spire2030.eu/disire/`).

Email addresses: `manas.mejari@imtlucca.it` (Manas Mejari), `dario.piga@supsi.ch` (Dario Piga), `alberto.bemporad@imtlucca.it` (Alberto Bemporad).

uct between the extended time-varying observability and controllability matrices, and later use an LPV extension of the predictor subspace approach originally proposed in [7]. As far as the identification of *LPV Input-Output* (LPV-IO) models is concerned, the closed-loop output error approach proposed in [12] in the LTI setting is extended in [5] to the identification of LPV-IO models, whose parameters are estimated recursively through a parameter adaptation algorithm. *Instrumental-Variable* (IV) based methods are proposed in [2,3,21]. The contribution in [2] is mainly focused on the identification of *quasi*-LPV systems, where the scheduling variable is a function of the output. The main idea in [2] is to recursively estimate the output signal (and thus the scheduling variable) through recursive least-squares and later use the estimated signals (instead of the measurements) to obtain a consistent estimate of the open-loop model parameters through IV methods. An indirect approach is used in [3], where IV methods are used to estimate a model of the closed-loop system based on pre-filtered external reference and output signals. The plant parameters are later extracted from the estimated closed-loop model using plant-controller separation methods. In [21], an iterative *Refined Instrumental Variable* (RIV) approach is proposed for closed-loop identification of LPV-IO models with Box-Jenkins noise structures. At each iteration of the IV algorithm, the signals are pre-filtered by stable LTI filters constructed using the parameters estimated at the previous iteration. The filtered signals are then used to build the instruments, which are used to recompute an (improved) estimate of the model parameters. Unlike the methods in [2,3], which are restricted to the case of LTI controllers, the approach in [21] can handle both LTI and LPV controllers.

This paper presents a bias-correction approach for closed-loop identification of LPV systems. The main idea underlying bias-correction methods is to eliminate the bias from ordinary *Least Squares* (LS) to obtain a consistent estimate of the model parameters. Bias-correction methods have been used in the past for the identification of LTI systems both in the open-loop [11,23] and closed-loop setting [24,10], as well as for open-loop identification of nonlinear [17] and LPV systems from noisy scheduling variable observations [16]. The main idea behind the closed-loop identification algorithm proposed in this paper is to quantify, based on the available measurements, the asymptotic bias due to the correlation between the plant input and the measurement noise. Recursive relations are derived to compute the asymptotic bias based on the knowledge of the controller and of the closed-loop structure of the system. Furthermore, in order to handle the more realistic scenario where not only the output signal, but also the scheduling variables are corrupted by a measurement noise, the proposed approach is combined with the ideas presented in [16], with the following improvements:

- an analytic expression, in terms of Hermite polynomials, is provided to compute the bias-correcting term used to handle the noise on the scheduling variable;
- as the bias-correcting term depends on the variance of the noise corrupting the scheduling variable, a bias-corrected cost function is introduced. This cost function serves as a tuning criterion to determine the value of the unknown noise variance via cross-validation.

Overall, the proposed closed-loop LPV identification approach offers a computationally low-demanding algorithm which: (i) provides a consistent estimate of the model parameters; (ii) can be applied under LTI or LPV controller structures; (iii) does not require to identify the closed-loop LPV system; (iv) can handle noisy observations of the scheduling signal.

The paper is organized as follows. The notation used throughout the paper is introduced in Section 2. The considered identification problem is formulated in Section 3. Section 4 describes the proposed closed-loop bias-correction approach that is extended in Section 5 to handle the case of identification from noisy measurements of the scheduling signal. Two case studies are reported in Section 6 to show the effectiveness of the presented method.

2 Notation

Let $\mathbb{R}^n$ be the set of real vectors of dimension n. The i-th element of a vector $x \in \mathbb{R}^n$ is denoted by x_i and $\|x\|^2 = x^\top x$ denotes the square of the 2-norm of x. For matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$, the Kronecker product between A and B is denoted by $A \otimes B \in \mathbb{R}^{mp\times nq}$. Given a matrix A, the symbol $[A]_{n\times m}$ means that A is a matrix of dimension $n \times m$. Let $\mathbb{I}_a^b$ be the sequence of successive integers $\{a, a+1, \cdots, b\}$, with $a < b$. The floor function is denoted by $\lfloor\cdot\rfloor$, where $\lfloor m \rfloor$ is the largest integer less than or equal to m. The expected value of a function f w.r.t. the random vector $x \in \mathbb{R}^n$ is denoted by $\mathbb{E}_{x_1,\cdots,x_n}\{f(x)\}$. The subscript $x_1, \cdots, x_n$ is dropped from $\mathbb{E}_{x_1,\cdots,x_n}$ when its meaning is clear from the context.

3 Problem Formulation

3.1 Data generating system

By referring to Figure 1, consider the LPV data-generating closed-loop system $\mathcal{S}_\text{o}$. We assume that the plant $\mathcal{G}_\text{o}$ is described by the LPV difference equations with output-error noise

$$\mathcal{G}_\text{o}: \begin{cases} \mathcal{A}_\text{o}(q^{-1}, p_\text{o}(k))x(k) = \mathcal{B}_\text{o}(q^{-1}, p_\text{o}(k))u(k), \\ \qquad\qquad\qquad\qquad y(k) = x(k) + e(k), \end{cases} \tag{1}$$

and that the controller $\mathcal{K}_o$ is a *known* LPV or LTI system described by

$$\mathcal{K}_o : \mathcal{C}_o(q^{-1}, p_o(k))u(k) = \mathcal{D}_o(q^{-1}, p_o(k))\,(r(k) - y(k)), \tag{2}$$

where $r(k)$ is a bounded reference signal of the closed-loop system $\mathcal{S}_o$; $u(k) \in \mathbb{R}$ and $y(k) \in \mathbb{R}$ are the measured input and output signals of the plant $\mathcal{G}_o$, respectively; $x(k)$ is noise-free output; $e(k) \sim \mathcal{N}(0, \sigma_e^2)$ is an additive zero-mean white Gaussian noise with variance σ_e^2 corrupting the output signal; $p_o(k) : \mathbb{N} \to \mathbb{P}$ is the measured (noise-free) scheduling signal and $\mathbb{P} \subseteq \mathbb{R}^{n_p}$ is a compact set where $p_o(k)$ is assumed to take values. In order not to make the notation too complex, from now on we assume that $p_o(k)$ is scalar (i.e., n_p=1). The operator q denotes the time shift (i.e., $q^{-i}x(k) = x(k-i)$), and $\mathcal{A}_o(q^{-1}, p_o(k))$, $\mathcal{B}_o(q^{-1}, p_o(k))$, $\mathcal{C}_o(q^{-1}, p_o(k))$ and $\mathcal{D}_o(q^{-1}, p_o(k))$ are polynomials in q^{-1} of degree n_a, n_b, n_c and $n_d - 1$, respectively, defined as follows:

$$\mathcal{A}_o(q^{-1}, p_o(k)) = 1 + \sum_{i=1}^{n_a} a_i^o(p_o(k))q^{-i},$$
$$\mathcal{B}_o(q^{-1}, p_o(k)) = \sum_{i=1}^{n_b} b_i^o(p_o(k))q^{-i},$$
$$\mathcal{C}_o(q^{-1}, p_o(k)) = 1 + \sum_{i=1}^{n_c} c_i^o(p_o(k))q^{-i},$$
$$\mathcal{D}_o(q^{-1}, p_o(k)) = \sum_{i=0}^{n_d-1} d_{i+1}^o(p_o(k))q^{-i},$$

where the coefficient functions $a_i^o, b_i^o, c_i^o, d_i^o$ are supposed to be polynomials in $p_o(k)$, i.e.,

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

$$b_i^o(p_o(k)) = \bar{b}_{i,0}^o + \sum_{s=1}^{n_g} \bar{b}_{i,s}^o p_o^s(k), \tag{3b}$$

$$c_i^o(p_o(k)) = \bar{c}_{i,0}^o + \sum_{s=1}^{n_g} \bar{c}_{i,s}^o p_o^s(k), \tag{3c}$$

$$d_i^o(p_o(k)) = \bar{d}_{i,0}^o + \sum_{s=1}^{n_g} \bar{d}_{i,s}^o p_o^s(k), \tag{3d}$$

with $\bar{a}_{i,s}^o \in \mathbb{R}$ and $\bar{b}_{i,s}^o \in \mathbb{R}$ being *unknown* real constants to be identified, while $\bar{c}_{i,s}^o \in \mathbb{R}$ and $\bar{d}_{i,s}^o \in \mathbb{R}$ are *known* coefficients characterizing the controller $\mathcal{K}_o$. In order not to burden the notation, the polynomials in (3) are assumed to have the same degree n_g.

The following assumptions are made for the closed-loop data generating system:

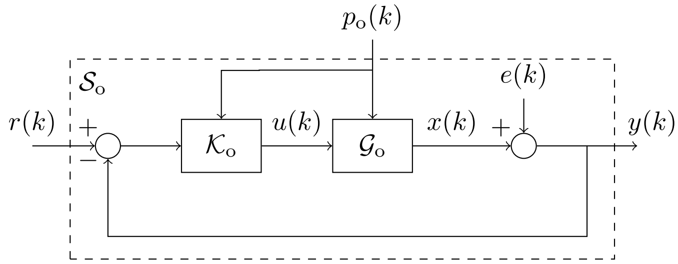


Fig. 1. Closed-loop LPV data-generating system

A1. the measurement noise $e(k)$ is uncorrelated with the scheduling signal $p_o(k)$ and with the external reference signal $r(k)$;
A2. to avoid algebraic loops, the open-loop plant is strictly causal, i.e., $b_0^o(p_o(k)) = 0$;
A3. the controller ensures closed-loop stability of the system $\mathcal{S}_o$ for any scheduling trajectory $p_o(k) \in \mathbb{P}$.

In order to describe the plant $\mathcal{G}_o$ in a compact form, the following matrix notation is introduced:

$$\bar{a}_i^o = \begin{bmatrix}\bar{a}_{i,0}^o & \bar{a}_{i,1}^o \cdots & \bar{a}_{i,n_g}^o\end{bmatrix}^\top,$$
$$\bar{b}_j^o = \begin{bmatrix}\bar{b}_{j,0}^o & \bar{b}_{j,1}^o \cdots & \bar{b}_{j,n_g}^o\end{bmatrix}^\top,$$
$$\theta_o = \begin{bmatrix}(\bar{a}_1^o)^\top \cdots (\bar{a}_{n_a}^o)^\top & (\bar{b}_1^o)^\top \cdots (\bar{b}_{n_b}^o)^\top\end{bmatrix}^\top,$$

$$\mathbf{p}_o(k) = \begin{bmatrix}1 & p_o(k) & p_o^2(k) \cdots p_o^{n_g}(k)\end{bmatrix}^\top,$$
$$\chi_o(k) = [-x(k-1) \cdots - x(k-n_a) \;\; u(k-1) \cdots u(k-n_b)]^\top,$$
$$\phi_o(k) = \chi_o(k) \otimes \mathbf{p}_o(k). \tag{4}$$

Based on the above notation, the plant $\mathcal{G}_o$ in (1) can be rewritten as:

$$\mathcal{G}_o : y(k) = \phi_o^\top(k)\theta_o + e(k). \tag{5}$$

3.2 Model structure for identification

The following parametrized model structure $\mathcal{M}_\theta$ is considered to describe the true LPV plant $\mathcal{G}_o$ in (1):

$$\mathcal{M}_\theta : y(k) = -\sum_{i=1}^{n_a} a_i(p_o(k))y(k-i) + \sum_{j=1}^{n_b} b_j(p_o(k))u(k-j) + \epsilon(k), \tag{6}$$

where $\epsilon(k)$ is the residual term.

The functions $a_i : \mathbb{R} \to \mathbb{R}$ and $b_j : \mathbb{R} \to \mathbb{R}$ are parametrized as follows:

$$a_i(p_o(k)) = \bar{a}_{i,0} + \sum_{s=1}^{n_g} \bar{a}_{i,s} p_o^s(k) = \bar{a}_i^\top \mathbf{p}_o(k), \tag{7a}$$

$$b_j(p_o(k)) = \bar{b}_{j,0} + \sum_{s=1}^{n_g} \bar{b}_{j,s} p_o^s(k) = \bar{b}_i^\top \mathbf{p}_o(k). \tag{7b}$$

Note that, since the paper aims at presenting a consistent closed-loop identification algorithm, the problem of model structure selection is not addressed. Thus, we assume that both the true plant $\mathcal{G}_o$ and the model $\mathcal{M}_\theta$ share the same parameters n_a, n_b and n_g.

By using a similar matrix notation already introduced to describe the true plant $\mathcal{G}_o$ in (5), the LPV model $\mathcal{M}_\theta$ in (6) can be written in the linear regression form:

$$\mathcal{M}_\theta : y(k) = \phi^\top(k)\theta + \epsilon(k), \tag{8}$$

where

$$\theta = [\bar{a}_1^\top \cdots \bar{a}_{n_a}^\top \ \bar{b}_1^\top \cdots \bar{b}_{n_b}^\top]^\top, \tag{9}$$

is the vector of model parameters to be identified and $\phi(k)$ is the regressor with measured outputs and scheduling signals at time k, defined as

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

with

$$\chi(k) = [-y(k-1) \cdots - y(k-n_a), u(k-1), \cdots u(k-n_b)]^\top. \tag{11}$$

The identification problem addressed in this paper aims at computing a consistent estimate of the true system parameter vector θ_o, given the model orders n_a, n_b and n_g and an N-length observed sequence $\mathcal{D}_N = \{u(k), y(k), p_o(k), r(k)\}_{k=1}^N$ of data generated by the closed-loop system $\mathcal{S}_o$ in Figure 1. To this aim, a novel identification algorithm based on asymptotic bias-corrected least squares is described in the next sections.

4 Bias-corrected least squares

It is well known that ordinary least squares give an asymptotically biased estimate of the model parameters due to the feedback structure [19]. In this section we quantify this bias and show how to remove it to give a consistent estimate of the model parameter vector θ.

4.1 Bias in the least-squares estimate

Consider the LS estimate $\hat{\theta}_{LS}$ given by:

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

under the assumption that matrix Γ_N is invertible. In order to compute the difference between the LS estimate $\hat{\theta}_{LS}$ and true system parameters θ_o, the output signal (5) is rewritten as follows:

$$\begin{aligned} y(k) &= \phi_o^\top(k)\theta_o + e(k) \\ &= [\chi_o(k) \otimes \mathbf{p}_o(k)]^\top \theta_o + e(k) \\ &= [\chi(k) \otimes \mathbf{p}_o(k)]^\top \theta_o \\ &\quad + [(\chi_o(k) - \chi(k)) \otimes \mathbf{p}_o(k)]^\top \theta_o + e(k) \\ &= \phi^\top(k)\theta_o + \Delta\phi(k)\theta_o + e(k), \end{aligned} \tag{13}$$

with

$$\Delta\phi(k) = [(\chi_o(k) - \chi(k)) \otimes \mathbf{p}_o(k)]^\top = \phi_o^\top(k) - \phi^\top(k). \tag{14}$$

Based on the representation of $y(k)$ in (13), the difference between the least-square estimate $\hat{\theta}_{LS}$ and the true system parameter vector θ_o can be expressed as:

$$\begin{aligned} \hat{\theta}_{LS} - \theta_o &= \Gamma_N^{-1} \sum_{k=1}^N \frac{1}{N}\phi(k)y(k) - \theta_o \\ &= \Gamma_N^{-1} \frac{1}{N}\sum_{k=1}^N \phi(k)\left(\phi^\top(k)\theta_o + \Delta\phi(k)\theta_o + e(k)\right) \\ &\quad - \theta_o \\ &= \Gamma_N^{-1} \underbrace{\frac{1}{N}\sum_{k=1}^N \phi(k)\phi^\top(k)}_{\Gamma_N}\theta_o \\ &\quad + \Gamma_N^{-1}\frac{1}{N}\sum_{k=1}^N \phi(k)\Delta\phi(k)\theta_o \\ &\quad + \Gamma_N^{-1}\frac{1}{N}\sum_{k=1}^N \phi(k)e(k) - \theta_o \\ &= \underbrace{\Gamma_N^{-1}\frac{1}{N}\sum_{k=1}^N \phi(k)\Delta\phi(k)\theta_o}_{B_\Delta(\theta_o, \phi(k), \Delta\phi(k))} + \underbrace{\Gamma_N^{-1}\frac{1}{N}\sum_{k=1}^N \phi(k)e(k)}_{B_e} \end{aligned} \tag{15}$$

Because of strict causality of the plant $\mathcal{G}_o$ (see Assumption A2), the regressor $\phi(k)$ is uncorrelated with the current value of the noise $e(k)$. Thus, the term B_e in (15) asymptotically (as $N \to \infty$) converges to zero with probability 1 (*w.p.* 1). Therefore, asymptotically, the bias in the LS estimate $\hat{\theta}_{\mathrm{LS}}$ is only due to the term $B_{\Delta}(\theta_{\mathrm{o}}, \phi(k), \Delta\phi(k))$, i.e.,

$$\lim_{N\to\infty} \hat{\theta}_{\mathrm{LS}} - \theta_{\mathrm{o}} = \lim_{N\to\infty} B_{\Delta}(\theta_{\mathrm{o}}, \phi(k), \Delta\phi(k)).$$

Note that, since the bias term $B_{\Delta}(\theta_{\mathrm{o}}, \phi(k), \Delta\phi(k))$ depends on the true system parameters θ_{o} as well as on the noise-free regressor $\phi_{\mathrm{o}}(k)$, it cannot be computed and thus it cannot be simply removed from the LS estimate $\hat{\theta}_{\mathrm{LS}}$.

In order to overcome the first difficulty due to the dependence of $B_{\Delta}(\theta_{\mathrm{o}}, \phi(k), \Delta\phi(k))$ on θ_{o}, the following estimate, inspired by [16], is introduced:

$$\tilde{\theta}_{\mathrm{CLS}} = \hat{\theta}_{\mathrm{LS}} - B_{\Delta}(\tilde{\theta}_{\mathrm{CLS}}, \phi(k), \Delta\phi(k)), \tag{16}$$

with

$$B_{\Delta}(\tilde{\theta}_{\mathrm{CLS}}, \phi(k), \Delta\phi(k)) = \Gamma_N^{-1} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\Delta\phi(k)\tilde{\theta}_{\mathrm{CLS}}.$$

The main idea behind (16) is to correct the least-squares estimate $\hat{\theta}_{\mathrm{LS}}$ by removing the bias term B_{Δ}, which is evaluated at the parameter estimate $\tilde{\theta}_{\mathrm{CLS}}$ instead of at the unknown system parameters θ_{o}. Note that (16) provides an implicit expression for the estimate $\tilde{\theta}_{\mathrm{CLS}}$, as the term B_{Δ} depends on $\tilde{\theta}_{\mathrm{CLS}}$ itself. By simple algebraic manipulations, (16) can be rewritten as:

$$\begin{aligned}
\tilde{\theta}_{\mathrm{CLS}} =& \hat{\theta}_{\mathrm{LS}} - \Gamma_N^{-1} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\Delta\phi(k)\tilde{\theta}_{\mathrm{CLS}} \\
=& \hat{\theta}_{\mathrm{LS}} - \Gamma_N^{-1} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\phi_{\mathrm{o}}^{\top}(k)\tilde{\theta}_{\mathrm{CLS}} + \Gamma_N^{-1}\Gamma_N\tilde{\theta}_{\mathrm{CLS}} \\
=& \Gamma_N^{-1} \left(\frac{1}{N} \sum_{k=1}^{N} \phi(k)y(k) \right) + \\
& - \Gamma_N^{-1} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\phi_{\mathrm{o}}^{\top}(k)\tilde{\theta}_{\mathrm{CLS}} + \tilde{\theta}_{\mathrm{CLS}}.
\end{aligned}$$

Thus,

$$\tilde{\theta}_{\mathrm{CLS}} = \left(\frac{1}{N} \sum_{k=1}^{N} \phi(k)\phi_{\mathrm{o}}^{\top}(k) \right)^{-1} \frac{1}{N} \sum_{k=1}^{N} \phi(k)y(k). \tag{17}$$

Using the definition $\Delta\phi(k) = \phi_{\mathrm{o}}^{\top}(k) - \phi^{\top}(k)$, (17) can be written as

$$\tilde{\theta}_{\mathrm{CLS}} = \mathbf{R}_N^{-1} \left(\frac{1}{N} \sum_{k=1}^{N} \phi^{\top}(k)y(k) \right), \tag{18}$$

where

$$\mathbf{R}_N = \frac{1}{N} \left(\sum_{k=1}^{N} \phi(k)\phi^{\top}(k) + \sum_{k=1}^{N} \phi(k)\Delta\phi(k) \right).$$

Property 1 *Assuming that the following limit exists:*

$$\lim_{N\to\infty} \mathbf{R}_N^{-1},$$

then $\tilde{\theta}_{\mathrm{CLS}}$ is a consistent estimate of true system parameters θ_{o}, i.e.,

$$\lim_{N\to\infty} \tilde{\theta}_{\mathrm{CLS}} = \theta_{\mathrm{o}} \;\; w.p.\ 1. \tag{19}$$

Proof: By substituting (13) into (18), we obtain

$$\begin{aligned}
\tilde{\theta}_{\mathrm{CLS}} = & \mathbf{R}_N^{-1} \underbrace{\frac{1}{N} \left(\sum_{k=1}^{N} \phi(k)(\phi^{\top}(k) + \Delta\phi(k)) \right)}_{\mathbf{R}_N} \theta_{\mathrm{o}} \\
& + \mathbf{R}_N^{-1} \left(\frac{1}{N} \sum_{k=1}^{N} \phi(k)e(k) \right).
\end{aligned}$$

Since the regressor $\phi(k)$ is uncorrelated with the current value of the noise $e(k)$, the term $\frac{1}{N}\sum_{k=1}^{N} \phi(k)e(k)$ asymptotically converges to zero *w.p.* 1. Thus,

$$\lim_{N\to\infty} \tilde{\theta}_{\mathrm{CLS}} = \theta_{\mathrm{o}} \;\; w.p.\ 1. \qquad \blacksquare$$

As $\Delta\phi(k)$ depends on the unknown noise-free regressors $\phi_{\mathrm{o}}(k)$ the estimate $\tilde{\theta}_{\mathrm{CLS}}$ in (18) cannot be computed. To overcome this problem, the term $\phi(k)\Delta\phi(k)$ is replaced by a bias-eliminating matrix Ψ_k, which is constructed (as explained in the following section) in such a way that it only depends on the available measurements $\mathcal{D}_N$ and it satisfies the following property:

$$\mathbf{C1}: \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\Delta\phi(k) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \Psi_k \;\; w.p.\ 1.$$

4.2 Construction of the bias-eliminating Ψ_k

A bias-eliminating matrix Ψ_k satisfying condition **C1** is constructed by evaluating the expected value of the matrix $\mathbb{E}\{\phi(k)\Delta\phi(k)\}$, as follows:

$$\begin{aligned}\Psi_k &= \mathbb{E}\{\phi(k)\Delta\phi(k)\}\\ &= \mathbb{E}\{(\chi(k)\otimes\mathbf{p}_\mathrm{o}(k))([(\chi_\mathrm{o}(k)-\chi(k))\otimes\mathbf{p}_\mathrm{o}(k)]^\top)\}\\ &= \mathbb{E}\{(\chi(k)\otimes\mathbf{p}_\mathrm{o}(k))([(\chi_\mathrm{o}(k)-\chi(k))^\top\otimes(\mathbf{p}_\mathrm{o}^\top(k))])\}\\ &= \mathbb{E}\{[\chi(k)(\chi_\mathrm{o}(k)-\chi(k))^\top]\otimes[\mathbf{p}_\mathrm{o}(k)\mathbf{p}_\mathrm{o}^\top(k)]\}\\ &= \mathbb{E}\{\Upsilon_k\otimes\mathbf{P}_\mathrm{o}(k)\}\\ &= \mathbb{E}\{\Upsilon_k\}\otimes\mathbf{P}_\mathrm{o}(k),\end{aligned} \tag{20}$$

with

$$\Upsilon_k = \chi(k)(\chi_\mathrm{o}(k)-\chi(k))^\top, \tag{21a}$$

$$\mathbf{P}_\mathrm{o}(k) = \mathbf{p}_\mathrm{o}(k)\mathbf{p}_\mathrm{o}^\top(k). \tag{21b}$$

The derivations reported above follow from the mixed-product property of the Kronecker product

$$(A\otimes B)(C\otimes D) = (AC)\otimes(BD). \tag{22}$$

Property 2 *The matrix $\mathbb{E}\{\Upsilon_k\}$ is given by*

$$[\mathbb{E}\{\Upsilon_k\}]_{(n_a+n_b)\times(n_a+n_b)} = \Lambda_k = \left[\begin{array}{c|c}(\Upsilon_k^y)_{n_a\times n_a} & \mathbf{0}_{n_a\times n_b},\\ \hline (\Upsilon_k^u)_{n_b\times n_a} & \mathbf{0}_{n_b\times n_b},\end{array}\right] \tag{23}$$

where Υ_k^y and Υ_k^u are upper triangular matrices,

$$\Upsilon_k^y = \begin{bmatrix} f_1(k-1) & f_2(k-2) & \cdots & f_{n_a}(k-n_a)\\ 0 & f_1(k-2) & \ddots & \vdots\\ \vdots & \ddots & \ddots & f_2(k-n_a)\\ 0 & \cdots & 0 & f_1(k-n_a)\end{bmatrix}, \tag{24a}$$

$$\Upsilon_k^u = \begin{bmatrix} g_1(k-1) & g_2(k-2) & \cdots & \cdots & g_{n_a}(k-n_a)\\ 0 & g_1(k-2) & \ddots & & \vdots\\ \vdots & \ddots & \ddots & & \vdots\\ 0 & \cdots & 0 & \ddots & g_{n_a-n_b+1}(n_a-n_b+1)\end{bmatrix}, \tag{24b}$$

and

$$f_m(k-j) = \mathbb{E}\{-y(k-j+m-1)e(k-j)\},$$
$$g_m(k-j) = \mathbb{E}\{u(k-j+m-1)e(k-j)\} \quad \forall m = \mathbb{I}_1^{n_a},$$

and

$$f_m(k) = g_m(k) = 0 \quad \text{for} \quad k \leq 0. \tag{25}$$

Proof: See Appendix A.1.

Property 3 *The relation between $f_m(k)$ and $g_m(k)$ can be expressed by the following recursion, initialized with $f_1(k) = -\sigma_e^2$ for all $k = 1,\ldots,N$,*

$$g_m(k) = -\sum_{i=1}^{\min(n_c,m-1)} c_i(p_\mathrm{o}(k+m-1))g_{m-i}(k) \tag{26a}$$

$$+\sum_{j=1}^{\min(n_d,m)} d_j(p_\mathrm{o}(k+m-1))f_{m-j+1}(k), \tag{26b}$$

$$f_m(k) = -\sum_{i=1}^{m-2} a_i^\mathrm{o}(p_\mathrm{o}(k+m-1))f_{m-i}(k) \tag{26c}$$

$$-\sum_{j=1}^{\min(n_b,m-1)} b_j^\mathrm{o}(p_\mathrm{o}(k+m-1))g_{m-j}(k). \tag{26d}$$

Proof: See Appendix A.2.

Remark 1 *In the case of open-loop data, the input signal is uncorrelated with the measurement noise affecting the output, i.e., $\mathbb{E}\{u(k-i)e(k-j)\} = 0, \forall i \neq j$. Moreover, as the measurement noise is assumed to be white, i.e., $\mathbb{E}\{y(k-i)e(k-j)\} = 0 \quad \forall i \neq j$, we have that*

1. *$\Upsilon_k^u = \mathbf{0}_{n_b\times n_a}$,*
2. *Υ_k^y is a diagonal matrix with the diagonal entries $[\Upsilon_k^y]_{i,i} = -\sigma_e^2$, and thus it does not depend on the true system parameter vector θ_o.*

The above matrices can be used to remove the bias in the identification of open-loop LPV models with an output-error type noise structure. ■

4.3 Bias corrected estimate

The matrix Ψ_k, which actually depends on the true system parameter θ_o, is constructed using Property 2 and Property 3 (namely, (20), (23) and (26)) using an estimated parameter vector $\hat{\theta}_\mathrm{CLS}$ instead of the unknown θ_o. Specifically, an implicit expression for the final bias-corrected estimate is given by:

$$\hat{\theta}_\mathrm{CLS} = \left(\frac{1}{N}\sum_{k=1}^{N}\Big(\phi(k)\phi^\top(k)+\Psi_k(\hat{\theta}_\mathrm{CLS})\Big)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)y(k)\right). \tag{27}$$

The main properties enjoyed by the estimate $\hat{\theta}_\mathrm{CLS}$ in (27) are reported in the following.

Property 4 *Assume that the following limit*

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^{N}\big(\phi(k)\phi^\top(k)+\Psi_k(\theta_\mathrm{o})\big)\right)^{-1} \tag{28}$$

exists. Then, asymptotically, the true system parameter vector θ_o *is a solution of* (27)*, namely, for* $\theta = \theta_\mathrm{o}$*,*

$$\theta = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^{N}\Big(\phi(k)\phi^\top(k)+\Psi_k(\theta)\Big)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)y(k)\right), \tag{29}$$

where the limit in (29) *holds w.p.* 1*. Thus, if* θ_o *is the unique solution of* (29)*, then the estimate* $\hat{\theta}_\mathrm{CLS}$ *in* (27) *is consistent, i.e.,*

$$\lim_{N\to\infty}\hat{\theta}_\mathrm{CLS} = \theta_\mathrm{o}. \tag{30}$$

Proof: By construction, $\mathbb{E}\{\Psi_k(\theta_\mathrm{o})\} = \mathbb{E}\{\phi(k)\Delta\phi(k)\}$, then Condition **C1** follows from Ninness' strong law of large numbers [15]. See [16, Appendix A2] for a detailed proof. By substituting

$$y(k) = \left(\phi^\top(k) + \Delta\phi(k)\right)\theta_\mathrm{o} + e(k)$$

into the right-hand side of (29), we obtain

$$\left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)\phi^\top(k)+\Psi_k(\theta_\mathrm{o})\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)\left(\phi^\top(k)+\Delta\phi(k)\right)\right)\theta_\mathrm{o} \tag{31a}$$

$$+\left(\frac{1}{N}\sum_{k=1}^{N}\Big(\phi(k)\phi^\top(k)+\Psi_k(\theta_\mathrm{o})\Big)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi(k)e(k)\right). \tag{31b}$$

As the regressor $\phi(k)$ is uncorrelated with the white noise $e(k)$, (31b) converges to zero *w.p.* 1 as $N\to\infty$. Furthermore, from condition **C1**, it follows that (31a) converges to θ_o as $N\to\infty$. Thus, (29) holds for $\theta=\theta_\mathrm{o}$. Furthermore, taking the limit of the left- and right-hand side of (27), (30) follows from (29) and uniqueness assumption. ∎

Note that (27) provides an implicit expression for the bias-corrected estimate $\hat{\theta}_\mathrm{CLS}$. In order to overcome this problem, (27) is solved iteratively as detailed in Algorithm 1. The main idea of Algorithm 1 is to compute, at each step τ, the bias-eliminating matrix Ψ_k using the estimate $\hat{\theta}_\mathrm{CLS}^{(\tau-1)}$ obtained at step $\tau-1$ and then to compute $\hat{\theta}_\mathrm{CLS}^{(\tau)}$ based on (27). Algorithm 1 can be initialized with a random vector $\hat{\theta}_\mathrm{CLS}^{(0)}$ or, for instance, with the LS estimates $\hat{\theta}_\mathrm{LS}$ in (12). Although convergence of Algorithm 1 is not theoretically proven, and its final solution may depend on the chosen initial condition, Algorithm 1 seems to be quite insensitive to initial conditions and its convergence has been empirically observed from numerical tests (cf. Section 6.1).

Algorithm 1 Iterative bias-correction algorithm

Input: noise variance σ_e^2; tolerance ϵ; maximum number τ^max of iterations; initial condition $\hat{\theta}_\mathrm{CLS}^{(0)}$.

1. **let** $\tau \leftarrow 0$;
2. **while:** $\tau \le \tau^\mathrm{max}$
   - 2.1. **let** $\tau \leftarrow \tau + 1$;
   - 2.2. **compute** $\Psi_k(\hat{\theta}_\mathrm{CLS}^{(\tau-1)})$ using Eqs. (20), (23) and (26);
   - 2.3. **calculate** the bias corrected estimates $\hat{\theta}_\mathrm{CLS}^{(\tau)}$ in (27);
   - 2.4. **if** $\left\|\hat{\theta}_\mathrm{CLS}^{(\tau)} - \hat{\theta}_\mathrm{CLS}^{(\tau-1)}\right\|_2 \le \epsilon$
     - 2.4.1. **exit** while;
   - 2.5. **end if**
3. **end while**

Output: Bias-corrected estimate $\hat{\theta}_\mathrm{CLS}$.

4.4 Estimate with unknown noise variance

In computing the bias-correcting matrix Ψ_k (and thus the bias-corrected estimate $\hat{\theta}_\mathrm{CLS}$ in (27)), the variance σ_e^2 of the noise corrupting the output signal measurements is assumed to be known. This is a restrictive assumption which may limit the applicability of the proposed identification approach. However, the unknown noise variance can be simply tuned via cross-validation. Specifically, the following cost can be minimized through a grid search over $\sigma_{e,i}^2$:

$$\mathcal{J}(\hat{\theta}_\mathrm{CLS}^i, \sigma_{e,i}^2) = \frac{1}{N_\mathrm{c}}\sum_{k=1}^{N_\mathrm{c}}\left(y(k) - \hat{x}^i(k)\right)^2, \tag{32}$$

where N_c is the length of the calibration set. The sequence $\hat{x}^i$ denotes the open-loop simulated output of the model with parameters $\hat{\theta}_\mathrm{CLS}^i$ estimated from Algorithm 1 using a given value of $\sigma_{e,i}^2$ as a guess for σ_e^2. The simulated output is defined as:

$$\hat{x}^i(k) = \hat{\phi}_\mathrm{cal}^\top(k)\hat{\theta}_\mathrm{CLS}^i,$$

where, the regressor $\hat{\phi}_\mathrm{cal}(k)$ (as defined in (4)) is given by

$$\hat{\chi}(k) = \left[-\hat{x}^i(k-1)\cdots-\hat{x}^i(k-n_a)\;\; u(k-1)\cdots u(k-n_b)\right]^\top,$$
$$\hat{\phi}_\mathrm{cal}(k) = \hat{\chi}(k)\otimes\mathbf{p}_\mathrm{o}(k).$$

It is worth stressing that the cost $\mathcal{J}$ in (32) is minimized only with respect to the scalar parameter σ_e. Specifically, once $\sigma_e^2 = \sigma_{e,i}^2$ is fixed, the corresponding $\hat{\theta}_\mathrm{CLS}^i$ (which depends on the chosen $\sigma_{e,i}^2$) is given by (27) and the corresponding cost $\mathcal{J}$ can be computed. Among the considered values of $\sigma_{e,i}^2$, the one minimizing $\mathcal{J}$ is taken.

5 Bias-correction with noisy scheduling signal

So far we have assumed that noise-free measurements of the scheduling variable $p_{\mathrm{o}}(k)$ are available. However, in many real applications, this might not be a realistic assumption, as the scheduling signal is often related to a measured signal and thus inherently corrupted by measurement noise (e.g., velocity and lateral acceleration in vehicle lateral dynamics modelling [6], gate-source voltage of a transistor in the description of an electronic filter [13], air speed and flight altitude in aircraft control [4]). This noise induces a bias in the final parameter estimate $\hat{\theta}_{\mathrm{CLS}}$ (27). Starting from the results presented in Section 4 and in [16] (where open-loop LPV identification from noisy scheduling variable measurements is addressed), in this section we show how to compute an asymptotically bias-free estimate of the LPV model parameters from closed-loop data with noisy measurements of the scheduling signal.

In particular, we consider the closed-loop data-generating system $\mathcal{S}_{\mathrm{o}}$ in Figure 1, and we assume that the noise-free scheduling signal $p_{\mathrm{o}}(k)$ is corrupted by an additive zero-mean white Gaussian noise with variance σ_η^2, independent of the output noise $e(k)$, i.e.,

$$p(k) = p_{\mathrm{o}}(k) + \eta(k), \quad \mathbb{E}\left\{\eta(k)e(t)\right\} = 0, \quad \forall k, t.$$

Following the same ideas described in Section 4, we quantify the bias in the LS estimate stemming from the output noise $e(k)$ and from the scheduling signal noise $\eta(k)$.

5.1 Bias-corrected least squares

By defining the "observed" regressor vector as

$$\phi_p(k) = \chi(k) \otimes \mathbf{p}(k),$$

with $\chi(k)$ defined in (11) and

$$\mathbf{p}(k) = \left[1 \;\; p(k) \;\; p^2(k) \cdots p^{n_g}(k)\right]^\top, \tag{34}$$

the standard least-squares estimate is given by:

$$\hat{\theta}_{\mathrm{LS}}^p = \underbrace{\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p(k)\phi_p^\top(k)\right)^{-1}}_{\Gamma_N^p} \frac{1}{N}\sum_{k=1}^{N}\phi_p(k)y(k). \tag{35}$$

By similar algebraic manipulations used in (15), the asymptotic bias in the LS estimate (35) is expressed as

$$\begin{aligned}\lim_{N\to\infty}\hat{\theta}_{\mathrm{LS}}^p - \theta_{\mathrm{o}} &= \underbrace{\lim_{N\to\infty}(\Gamma_N^p)^{-1}\frac{1}{N}\sum_{k=1}^{N}\phi_p(k)\Delta\phi(k)\theta_{\mathrm{o}}}_{B_\Delta(\theta_{\mathrm{o}},\phi_p(k),\Delta\phi(k))} \\ &+ \underbrace{\lim_{N\to\infty}(\Gamma_N^p)^{-1}\frac{1}{N}\sum_{k=1}^{N}\phi_p(k)\Delta\phi_p(k)\theta_{\mathrm{o}}}_{B_p(\theta_{\mathrm{o}},\phi_p(k),\Delta\phi_p(k))},\end{aligned} \tag{36}$$

with $\Delta\phi(k)$ as defined in (14) and

$$\Delta\phi_p(k) = \left[\chi(k)\otimes(\mathbf{p}_{\mathrm{o}}(k)-\mathbf{p}(k))\right]^\top.$$

Following the same rationale used to define $\tilde{\theta}_{\mathrm{CLS}}$ in (16), let us introduce the bias-corrected estimate

$$\begin{aligned}\tilde{\theta}_{\mathrm{CLS}}^p = \hat{\theta}_{\mathrm{LS}}^p &- B_\Delta(\tilde{\theta}_{\mathrm{CLS}}^p, \phi_p(k), \Delta\phi(k)) \\ &- B_p(\tilde{\theta}_{\mathrm{CLS}}^p, \phi_p(k), \Delta\phi_p(k)).\end{aligned} \tag{37}$$

Remark 2 *In the case of noise-free scheduling signal observations (i.e.,* $\mathbf{p}_{\mathrm{o}}(k) = \mathbf{p}(k)$*)* $\phi_p(k) = \phi(k)$ *and* $B_p(\tilde{\theta}_{\mathrm{CLS}}^p, \phi_p(k), \Delta\phi_p(k)) = 0$*. Thus,* (37) *coincides with* (16)*.* ■

By algebraic manipulations, the estimate $\tilde{\theta}_{\mathrm{CLS}}^p$ in (37) can be rewritten explicitly as:

$$\tilde{\theta}_{\mathrm{CLS}}^p = \Bigg(\underbrace{\frac{\sum_{k=1}^N \phi_p(k)\phi_{\mathrm{o}}^\top(k)}{N}}_{R(\mathbf{p}_0)}\Bigg)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p(k)y(k)\right), \tag{38}$$

or equivalently as in (39) [1].

Then, a bias-corrected estimate $\hat{\theta}_{\mathrm{CLS}}^p$ can be obtained from (39d) as follows:

- replace the matrix $\chi(k)\chi^\top(k)\otimes\left[\mathbf{p}(k)\mathbf{p}_{\mathrm{o}}^\top(k)\right]$ by a matrix $\chi(k)\chi^\top(k)\otimes\Psi_k^p$ depending only on the available dataset $\mathcal{D}_N^p = \{u(k), y(k), p(k), r(k)\}_{k=1}^N$ and satisfying condition:

$$\mathbf{C2}: \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\mathbf{p}(k)\mathbf{p}_{\mathrm{o}}^\top(k) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\Psi_k^p \;\; w.p.\ 1. \tag{40}$$

[1] Eq. (39b) follows from (39a) and the Kronecker product property (22).

$$\tilde{\theta}_{\text{CLS}} = \left(\frac{\sum_{k=1}^{N}\left[\chi(k)\otimes\mathbf{p}(k)\right]\left[\chi_{\text{o}}(k)\otimes\mathbf{p}_{\text{o}}(k)\right]^\top}{N}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p^\top(k)y(k)\right) \tag{39a}$$

$$= \left(\frac{\sum_{k=1}^{N}\left[\chi(k)\chi_{\text{o}}^\top(k)\right]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]}{N}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p^\top(k)y(k)\right) \tag{39b}$$

$$= \left(\frac{\sum_{k=1}^{N}\left[\chi(k)(\chi_{\text{o}}(k)-\chi(k))^\top\right]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]+\left[\chi(k)\chi^\top(k)\right]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]}{N}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p^\top(k)y(k)\right) \tag{39c}$$

$$= \left(\frac{\sum_{k=1}^{N}\left[\chi(k)\Delta\chi(k)\right]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]+\left[\chi(k)\chi^\top(k)\right]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]}{N}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\phi_p^\top(k)y(k)\right). \tag{39d}$$

- replace the matrix $[\chi(k)\Delta\chi(k)]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]$ by a matrix Ω_k depending only on the available dataset $\mathcal{D}_N^p$ and satisfying condition:

$$\mathbf{C3}: \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}[\chi(k)\Delta\chi(k)]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right] = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\Omega_k \quad w.p.\ 1. \tag{41}$$

The procedure to construct the matrices Ψ_k^p and Ω_k satisfying conditions **C2** and **C3** is outlined in the following section.

5.2 Construction of the bias-eliminating matrices

5.2.1 Construction of Ψ_k^p

Inspired by [16], the bias-correction matrix Ψ_k^p satisfying condition **C2** in (40) is constructed as follows:

1. compute the analytic expression of $\mathbb{E}\{\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\}$. Note that, since $\mathbf{p}_{\text{o}}(k)$ and $\mathbf{p}(k)$ are polynomials in $p_{\text{o}}(k)$ and $p(k)$ (see (4) and (34)), the entries of $\mathbb{E}\{\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\}$ are polynomials in $p_{\text{o}}(k)$;
2. express the $n-$th order monomial $p_{\text{o}}^n(k)$ in terms of the expected value of the noise-corrupted observation $p^n(k)$ and noise variance σ_η^2 as the "probabilists" Hermite polynomial: [2]

$$p_{\text{o}}^n(k) = \mathbb{E}\left\{(n!)\sum_{m=0}^{\lfloor n/2\rfloor}\frac{(-1)^m\sigma_\eta^{2m}}{m!(n-2m)!}\frac{p^{n-2m}(k)}{2^m}\right\}; \tag{42}$$

3. compute the matrix Ψ_k^p by replacing each of the monomials $p_{\text{o}}(k)$, $p_{\text{o}}^2(k)$, $p_{\text{o}}^3(k), \ldots$ appearing in the analytic expression of $\mathbb{E}\{\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\}$, with the term inside the expectation operator in (42).

By construction, the matrix Ψ_k^p satisfies

$$\mathbb{E}\{\Psi_k^p\} = \mathbb{E}\left\{\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right\}. \tag{43}$$

Based on (43) and Ninness' strong law of large numbers [15], Ψ_k^p satisfies Condition **C2**. An example of construction of matrix Ψ_k^p is reported in Appendix A.3.

5.2.2 Construction of Ω_k

The matrix Ω_k satisfying condition **C3** can be constructed by properly combing the ideas used to construct the bias-eliminating matrices Ψ_k^p (see Section 5.2.1) and Υ_k^y and Υ_k^u (introduced in (24)). Specifically, matrix Ω_k is constructed in such a way that the following equality holds:

$$\mathbb{E}_{e,\eta}\{\Omega_k\} = \mathbb{E}_{e,\eta}\left\{[\chi(k)\Delta\chi(k)]\otimes\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]\right\}. \tag{44}$$

Since, $\chi(k)\Delta\chi(k)$ does not depend on the noise $\eta(k)$ and $\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)$ does not depend on the output noise e, and since the random variables $e(k)$ and $\eta(k)$ are independent, (44) is equivalent to

$$\mathbb{E}_{e,\eta}\{\Omega_k\} = \mathbb{E}_e\{[\chi(k)\Delta\chi(k)]\}\otimes\mathbb{E}_\eta\left\{\left[\mathbf{p}(k)\mathbf{p}_{\text{o}}^\top(k)\right]\right\}. \tag{45}$$

Note that $\chi(k)\Delta\chi(k)$ is equal to Υ_k as defined in (21a). Thus, $\mathbb{E}_e\{[\chi(k)\Delta\chi(k)]\}$ is equal to Λ_k (see (23)) and it can be constructed using the results in Property 2. However, Λ_k defined in (23) depends on the noise-free scheduling signal p_{o}, and thus its expression can be only derived analytically, but it cannot be constructed based on the available dataset $\mathcal{D}_N^p$. Nevertheless, as $\Lambda_k(\mathbf{p}_{\text{o}})$ does not depend on the random variable η, condition (45)

[2] The expression of $p_{\text{o}}^n(k)$ in terms of the expected value of the noise-corrupted observation $p^n(k)$ and noise variance σ_η^2 is not reported in [16] in terms of the Hermite polynomial (42), but in terms of recursive constructions which can be proved to have the compact expression in (42).

becomes

$$\mathbb{E}_{e,\eta}\left\{\Omega_k\right\} = \Lambda_k(\mathbf{p}_\text{o}) \otimes \mathbb{E}_\eta\left\{\left[\mathbf{p}(k)\mathbf{p}_\text{o}^\top(k)\right]\right\} \\ = \mathbb{E}_\eta\left\{\Lambda_k(\mathbf{p}_\text{o}) \otimes \left[\mathbf{p}(k)\mathbf{p}_\text{o}^\top(k)\right]\right\}. \quad (46)$$

Thus, Ω_k can be constructed based on the same procedure outlined in Section 5.2.1 to construct Ψ_k^p, replacing the term $\mathbb{E}\left\{\left[\mathbf{p}(k)\mathbf{p}_\text{o}^\top(k)\right]\right\}$ in Section 5.2.1 with the term $\mathbb{E}_\eta\left\{\Lambda_k(\mathbf{p}_\text{o}) \otimes \left[\mathbf{p}(k)\mathbf{p}_\text{o}^\top(k)\right]\right\}$.

As the matrix $\Lambda_k(\mathbf{p}_\text{o})$ has a dynamic dependence on p_o (i.e., it is a function of $p_\text{o}(k)$, $p_\text{o}(k-1), \ldots$), the analytic expression of $\Lambda_k(\mathbf{p}_\text{o}) \otimes \left[\mathbf{p}(k)\mathbf{p}_\text{o}^\top(k)\right]$ has product terms such as $p_\text{o}^n(k)$, $p_\text{o}^n(k-1), \ldots$. Nevertheless, as the noise terms $\eta(k)$ and $\eta(k-t)$ are uncorrelated, $\forall t \neq 0$, we have that $\mathbb{E}_\eta\left\{p^n(k)p^n(k-1)\right\} = \mathbb{E}_\eta\left\{p^n(k)\right\}\mathbb{E}_\eta\left\{p^n(k-1)\right\}$, and the Hermite polynomial expression defined in (42) can be used to construct Ω_k.

5.3 Bias-corrected estimate

Based on (39d) and the the ideas introduced in the previous sections, the final bias-corrected estimate $\hat{\theta}_\text{CLS}^p$ is given by

$$\hat{\theta}_\text{CLS}^p = \left(\frac{1}{N}\sum_{k=1}^N \Omega_k(\hat{\theta}_\text{CLS}^p) + \left[\chi(k)\chi^\top(k)\right] \otimes \Psi_k^p\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^N \phi_p^\top(k)y(k)\right). \quad (47)$$

Note that, as in the case of noise-free scheduling signal, the matrix Ω_k depends on the true system parameter vector θ_o, and the estimate $\hat{\theta}_\text{CLS}^p$ should be computed based on an iterative approach similar to Algorithm 1.

Property 5 *Assume that the following limit*

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^N \Omega_k(\hat{\theta}_\text{CLS}^p) + \left[\chi(k)\chi^\top(k)\right] \otimes \Psi_k^p\right)^{-1}. \quad (48)$$

exists. Then, asymptotically, the true system parameter vector θ_o is a solution of (47), namely, for $\theta = \theta_\text{o}$,

$$\theta = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=1}^N \Omega_k(\theta) + \left[\chi(k)\chi^\top(k)\right] \otimes \Psi_k^p\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^N \phi_p^\top(k)y(k)\right), \quad (49)$$

where the limit in (49) holds w.p. 1. Thus, if θ_o is the unique solution of (49), then the estimated $\hat{\theta}_\text{CLS}^p$ in (47) is consistent, i.e.,

$$\lim_{N\to\infty}\hat{\theta}_\text{CLS}^p = \theta_\text{o}. \quad (50)$$

Proof: Property 5 follows from conditions **C2** and **C3** and from the same rationale used in the proof of Property 4. ■

5.4 Estimation with unknown variances σ_e^2 and σ_η^2

In computing the bias correcting matrices Ω_k and Ψ_k^p, the noise variances σ_e^2 and σ_η^2 are assumed to be known. In the case of noise-free scheduling variable, the open-loop simulation error was used in Section 4.4 as a performance criterion to tune σ_e^2 via cross validation. However, in the noisy p scenario, a cross-validation procedure will fail, as a model with the "true" system parameters θ_o will not provide the "true" output due the fact that the scheduling variable $p(k)$ used to simulate the output of the model is not the "true" one. In order to overcome this problem, we propose next a novel procedure based on a bias-free tuning criterion.

Let us introduce the simulated regressor

$$\hat{\chi}(k) = \left[-\hat{y}(k-1)\cdots-\hat{y}(k-n_a), u(k-1), \cdots u(k-n_b)\right]^\top, \quad (51)$$

where $\hat{y}(k)$ is the bias-corrected simulated model output at time k given by

$$\hat{y}(k) = \left[\hat{\chi}(k) \otimes \mathbf{p}^\text{C}(k)\right]^\top \hat{\theta}_\text{CLS}^p \quad (52)$$

and $\mathbf{p}^\text{C}(k)$ being the vector of bias-corrected monomials [3].

Given an estimate $\hat{\theta}_\text{CLS}^p$, computed through (47) for fixed values of σ_e and σ_η, and a calibration dataset of length N_c not used to compute $\hat{\theta}_\text{CLS}^p$, define the cost

$$\mathcal{J}_\text{BC}\left(\hat{\theta}_\text{CLS}^p(\sigma_e,\sigma_\eta)\right) = \Bigg\| \frac{1}{N_c}\sum_{k=1}^{N_c}\left[\chi(k)\hat{\chi}^\top(k) \otimes \Psi_k^p\right]\hat{\theta}_\text{CLS}^p(\sigma_e,\sigma_\eta) \\ - \frac{1}{N_c}\sum_{k=1}^{N_c}\left(\Omega_k(\hat{\theta}_\text{CLS}^p) + \left[\chi(k)\chi^\top(k)\right] \otimes \Psi_k^p\right)\hat{\theta}_\text{CLS}^p(\sigma_e,\sigma_\eta)\Bigg\|^2. \quad (53)$$

The cost $\mathcal{J}_\text{BC}$ will be referred to as bias-corrected cost and, as discussed in the following, it should be used as a criterion to tune the unknown noise variances σ_e^2 and σ_η^2.

Property 6 *The bias-corrected cost (53) asymptotically achieves its minimum at $\hat{\theta}_\text{CLS}^p = \theta_\text{o}$, i.e.,*

$$\theta_\text{o} = \arg\min_\theta \lim_{N_c\to\infty}\mathcal{J}_\text{BC}(\theta) \;\; \textit{w.p. } 1. \quad (54)$$

Proof: See Appendix A.4.

[3] The vector of bias-corrected monomials $\mathbf{p}^\text{C}(k)$ is such that it only depends on $\mathbf{p}(k)$ and σ_η and satisfies the condition $\mathbb{E}\left\{\mathbf{p}^\text{C}(k)\right\} = \mathbf{p}_\text{o}(k)$. Thus, it can constructed using the Hermite polynomial (42). For instance, when $\mathbf{p}_\text{o}(k) = \left[1 \;\; p_\text{o}(k) \;\; p_\text{o}^2(k)\right]^\top$, then $\mathbf{p}^\text{C}(k) = \left[1 \;\; p(k) \;\; p^2(k) - \sigma_\eta^2\right]^\top$.

Property 6 proves that, if $\mathcal{J}_{\text{BC}}(\theta)$ has asymptotically a unique minimizer, then its minimum is achieved at the true system parameter vector θ_{o}. Thus, $\mathcal{J}_{\text{BC}}(\theta)$ is an asympotically bias-free criterion which can be used to assess the quality of a given model parameter vector $\hat{\theta}^p_{\text{CLS}}$. Therefore, the hyper-parameters σ_e and σ_η can be tuned through a grid search using $\mathcal{J}_{\text{BC}}\left(\hat{\theta}^p_{\text{CLS}}(\sigma_e, \sigma_\eta)\right)$ as a performance metric on a calibration dataset.

6 Case studies

In order to show the effectiveness of the proposed identification method, we consider two examples. In the first example, we focus on the effect of the measurement noise on the final parameter estimate, hence the model structure of the true LPV data-generating system is assumed to be exactly known. As a more realistic case study, the second example addresses the identification of a nonlinear two-tank system. All the simulations are carried out on an i5 2.40-GHz Intel core processor with 4 GB of RAM running MATLAB R2015b.

The performance of the identified models is assessed on a noiseless validation dataset not used for training through the *Best Fit Rate* (BFR) index, defined as

$$\text{BFR} = \max\left\{1 - \sqrt{\frac{\sum_{k=1}^{N_{\text{val}}} (y(k) - \hat{y}(k))^2}{\sum_{k=1}^{N_{\text{val}}} (y(k) - \bar{y})^2}}, 0\right\}, \quad (55)$$

with N_{val} being the length of the validation set and $\hat{y}$ being the estimated model output and $\bar{y}$ the sample mean of the output signal.

6.1 Example 1

6.1.1 Data-generating system

The considered closed-loop data-generating system $\mathcal{S}_{\text{o}}$ is taken from [1], and it consists of an (unknown) LPV plant $\mathcal{G}_{\text{o}}$ described by (1), with

$$\mathcal{A}_{\text{o}}(q^{-1}, p_k) = 1 + a_1^{\text{o}}(p_{\text{o}}(k))q^{-1} + a_2^{\text{o}}(p_{\text{o}}(k))q^{-2}, \quad (56a)$$
$$\mathcal{B}_{\text{o}}(q^{-1}, p_k) = b_1^{\text{o}}(p_{\text{o}}(k))q^{-1} + b_2^{\text{o}}(p_{\text{o}}(k))q^{-2}, \quad (56b)$$

where,

$$a_1^{\text{o}}(p_{\text{o}}(k)) = 1.0 - 0.5p_{\text{o}}(k) - 0.1p_{\text{o}}^2(k), \quad (57a)$$
$$a_2^{\text{o}}(p_{\text{o}}(k)) = 0.5 - 0.7p_{\text{o}}(k) - 0.1p_{\text{o}}^2(k), \quad (57b)$$
$$b_1^{\text{o}}(p_{\text{o}}(k)) = 0.5 - 0.4p_{\text{o}}(k) + 0.01p_{\text{o}}^2(k), \quad (57c)$$
$$b_2^{\text{o}}(p_{\text{o}}(k)) = 0.2 - 0.3p_{\text{o}}(k) - 0.02p_{\text{o}}^2(k). \quad (57d)$$

The noise term $e(k)$ corrupting the output observations is a white Gaussian noise with standard deviation $\sigma_e =$ 0.05. This corresponds to a *Signal-to-Noise Ratio* (SNR) of 12.5 dB, where the SNR on the output channel is defined as

$$\text{SNR}_{\text{y}} = 10 \log \frac{\sum_{k=1}^{N} (x(k) - \bar{x})^2}{\sum_{k=1}^{N} e^2(k)}, \quad (58)$$

with $\bar{x}$ denoting the mean of the noise free output.

The controller $\mathcal{K}_{\text{o}}$ is LTI and known, and it is described by (2) with

$$\mathcal{C}_{\text{o}}(q^{-1}, p_k) = 1 + c_1^{\text{o}}(p_{\text{o}}(k))q^{-1} + c_2^{\text{o}}(p_{\text{o}}(k))q^{-1},$$
$$\mathcal{D}_{\text{o}}(q^{-1}, p_k) = d_1^{\text{o}}(p_{\text{o}}(k)) + d_2^{\text{o}}(p_{\text{o}}(k))q^{-1} + d_3^{\text{o}}(p_{\text{o}}(k))q^{-2},$$

with

$$c_1^{\text{o}}(p_{\text{o}}(k)) = -0.28, \;\; c_2^{\text{o}}(p_{\text{o}}(k)) = 0.5,$$
$$d_1^{\text{o}}(p_{\text{o}}(k)) = 0.35, \;\; d_2^{\text{o}}(p_{\text{o}}(k)) = -0.28, \;\; d_3^{\text{o}}(p_{\text{o}}(k)) = 0.1.$$

The scheduling signal trajectory is described by:

$$p_{\text{o}}(k) = 1.1(0.5\sin(0.35\pi k) + 0.05).$$

The reference $r(k)$ is a white noise signal with uniform distribution in the interval $[-1 \;\; 1]$. A training data set $\mathcal{D}_N$ of length $N = 20,000$ is used to estimate the plant $\mathcal{G}_{\text{o}}$ and, in order to assess the statistical properties of the proposed identification approach, a Monte-Carlo study with 100 runs is performed. At each Monte-Carlo run, a new data set of inputs $u(k)$, scheduling variables $p_{\text{o}}(k)$, reference signal $r(k)$ and noise $e(k)$ is generated.

6.1.2 Model structure

As a model structure for the plant $\mathcal{G}_{\text{o}}$, we consider the second-order LPV model

$$y(k) = -\sum_{i=1}^{2} a_i(p_{\text{o}}(k))y(k-i) + \sum_{j=1}^{2} b_j(p_{\text{o}}(k))u(k-j),$$

where the coefficient functions $a_i(p_{\text{o}}(k))$ and $b_j(p_{\text{o}}(k))$ are parametrized as second order-polynomials:

$$a_1(p_{\text{o}}(k)) = a_{1,0} + a_{1,1}p_{\text{o}}(k) + a_{1,2}p_{\text{o}}^2(k),$$
$$a_2(p_{\text{o}}(k)) = a_{2,0} + a_{2,1}p_{\text{o}}(k) + a_{2,2}p_{\text{o}}^2(k),$$
$$b_1(p_{\text{o}}(k)) = b_{1,0} + b_{1,1}p_{\text{o}}(k) + b_{1,2}p_{\text{o}}^2(k),$$
$$b_2(p_{\text{o}}(k)) = b_{2,0} + b_{2,1}p_{\text{o}}(k) + b_{2,2}p_{\text{o}}^2(k).$$

6.1.3 Identification from noise-free scheduling signal

First, we assume that the observations of the scheduling variable $p_{\text{o}}(k)$ are not corrupted by a measurement noise. The following two cases are considered:

1. the variance σ_e^2 of the noise $e(k)$ on the output signal $y(k)$ is known;
2. σ_e^2 is unknown.

Furthermore, since Algorithm 1 depends on the initial guess $\hat{\theta}_{\text{CLS}}^{(0)}$ used to iteratively compute the bias-correcting matrix $\Psi_k(\hat{\theta}_{\text{CLS}}^{(\tau-1)})$ (see Step 2.2), we test its sensitivity w.r.t. different initial conditions $\hat{\theta}_{\text{CLS}}^{(0)}$.

Identification with known variance σ_e^2

The identification results obtained through standard least-squares and the closed-loop bias-correction approach presented in Algorithm 1 are compared in Table 1, which shows the averages and the standard deviations of the estimated model parameters over 100 Monte-Carlo runs. The average CPU time for computing the estimate for a given value of noise variance is 2.5 sec.

In order to further assess the performance of the developed identification scheme, we also compute the BFR on a noise-free validation data set of length $N_{\text{val}} = 10,000$, which is reported in Table 2. The obtained results shows that, unlike the least squares, the proposed approach provides a consistent estimate of the system parameters. This leads to a higher BFR (namely, better reconstruction of the output signal on the validation set) w.r.t. least squares.

In order to analyze the sensitivity of Algorithm 1 w.r.t. the initial condition $\hat{\theta}_{\text{CLS}}^{(0)}$, we initialize Algorithm 1 with 100 different random values of $\hat{\theta}_{\text{CLS}}^{(0)}$. The initial values of each component of $\hat{\theta}_{\text{CLS}}^{(0)}$ are chosen randomly from a uniform distribution in the interval $[0 \;\; 1]$. The iterative algorithm is stopped when no change in the final estimate is observed or when a maximum number of iterations $\tau^{\max} = 50$ is reached. The same training data-set is used in all runs. We observe that the algorithm is insensitive to the initial conditions and it provides the same model estimate, resulting in an equal BFR for all the 100 different initial conditions $\hat{\theta}_{\text{CLS}}^{(0)}$ (see Figure 2).

The proposed method is also compared with a *prediction-error method* (PEM). In the prediction-error identification framework, the unknown plant parameters θ are obtained by minimizing the one-step ahead prediction-error: $\epsilon_\theta(k) = y(k) - \hat{y}(k \mid k-1) = \hat{\phi}^\top(k)\theta$, resulting in the minimization of the following non-convex loss function:

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

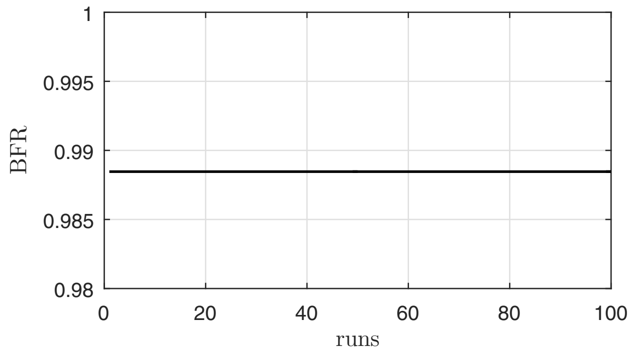


Fig. 2. Example 1. Best Fit Rate for different initial conditions $\hat{\theta}_{\text{CLS}}^{(0)}$ of Algorithm 1.

where, the regressor $\hat{\phi}(k)$ (as defined in (4)) is given by

$$\hat{\chi}(k) = [-\hat{y}(k-1) \cdots - \hat{y}(k-n_a) \;\; u(k-1) \cdots u(k-n_b)]^\top,$$
$$\hat{\phi}(k) = \hat{\chi}(k) \otimes \mathbf{p}_o(k).$$

The average CPU time taken by the PEM to find the estimate is 2.5 sec. The estimated model parameters and the achieved BFR are reported in Table 1 and 2, respectively. Similar results are obtained by the bias-correction approach and PEM. However, unlike PEM, the proposed bias-correction approach leads to a consistent parameter estimate also in the case of noisy scheduling variable observations (as shown in the results reported in Section 6.1.4).

Identification with unknown noise variance σ_e^2

We now consider the case where the variance σ_e^2 of the noise corrupting the output signal is not known a priori, but recovered through the cross-validation procedure described in Section 4.4. Figure 3 shows the cost function $\mathcal{J}(\sigma_e)$ (multiplied by N_c for a better visualization) defined in (32) against different values of the hyper-parameter σ_e. Note that the minimum of $\mathcal{J}$ is achieved exactly at the true value of the noise standard deviation (i.e., at $\sigma_e = 0.05$). Thus, since the true value of σ_e^2 is exactly recovered, the estimated model parameters coincide with the ones obtained in the case of known variance σ_e^2 (and already provided in Table 1).

6.1.4 *Identification from noisy scheduling signal*

In this paragraph, the proposed closed-loop identification algorithm is tested for the case of noisy measurements of the scheduling signal. To this aim, the scheduling variable observations are corrupted by an additive zero-mean white Gaussian noise $\eta_o(k)$ with standard deviation $\sigma_\eta = 0.12$. This corresponds to a *Signal-To-Noise Ratio* SNR_p equal to 10 dB [4].

[4] The *Signal-To-Noise Ratio* SNR_p on scheduling variable observations is defined similarly to (58).

Table 1
Example 1. Identification from noise-free scheduling signal measurements: means and standard deviations (over 100 Monte-Carlo runs) of the estimated parameters using Least Squares , the proposed closed-loop bias-correction method and prediction-error method (PEM).

True Value	Least Squares	Bias-Correction	PEM
1	0.7542 ± 0.0085	0.9992 ± 0.0138	0.9976 ± 0.0082
-0.5	-0.2117 ± 0.0194	-0.4785 ± 0.0425	-0.4999 ± 0.0188
-0.1	-0.9288 ± 0.0525	-0.1245 ± 0.1229	-0.0873 ± 0.0606
0.5	0.3449 ± 0.0057	0.5000 ± 0.0088	0.4986 ± 0.0057
-0.7	-0.7288 ± 0.0099	-0.6961 ± 0.0181	-0.7016 ± 0.0088
-0.1	-0.1685 ± 0.0295	-0.0994 ± 0.0609	-0.0898 ± 0.0403
0.5	0.5001 ± 0.0037	0.5008 ± 0.0041	0.4996 ± 0.0023
-0.4	-0.4007 ± 0.0070	-0.3997 ± 0.0081	-0.4007 ± 0.0027
0.01	-0.0266 ± 0.0194	0.0063 ± 0.0235	0.0109 ± 0.0108
0.2	0.0671 ± 0.0058	0.1995 ± 0.0082	0.1986 ± 0.0043
-0.3	-0.0788 ± 0.0136	-0.2887 ± 0.0267	-0.2999 ± 0.0118
-0.02	-0.4697 ± 0.0352	-0.0337 ± 0.0680	-0.0147 ±0.0328

Table 2
Example 1. Identification from noise-free scheduling signal measurements: *Best Fit Rates* (BFRs) over (noise-free) validation data.

Method	BFR
Least-squares	0.8202
Bias-correction	0.9964
PEM	0.9984

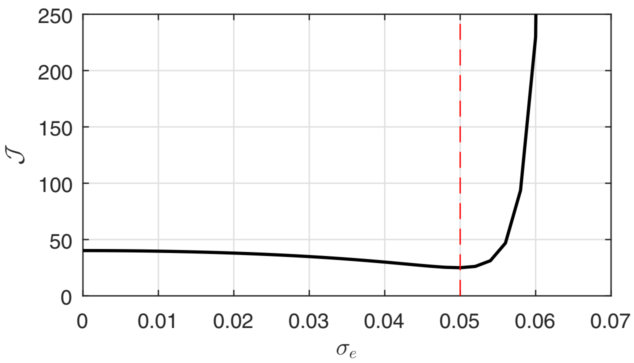


Fig. 3. Example 1. Bias-corrected cost $\mathcal{J}$ (defined in (32)) vs noise standard deviation σ_e.

The unknown model parameters are computed through the following three approaches:

1. Least Squares;
2. Bias Correction 1: closed-loop bias-correction without handling the bias due to the noise on p. The model parameters are estimated using Algorithm 1, correcting only the bias due to the output noise e.
3. Bias Correction 2: closed-loop bias-correction correcting both the bias due to the noise on the scheduling signal observations and the bias due to the feedback structure.

First, we consider the case when the noise variances σ_e^2 and σ_η^2 are known. The estimated model parameters are provided in Table 3. The norm $\left\|\theta_{\rm o} - \hat{\theta}\right\|_2$ of the difference between the true system parameters $\theta_{\rm o}$ and the estimate parameters $\hat{\theta}$ is reported in Table 4, along with the BFRs on validation data. The obtained results show that correcting the bias due to the noise on the scheduling signal observations further improves the final model parameter estimate.

Finally, we present the results of the proposed method when no information is available a priori about the variance of the noise corrupting the output and the scheduling signal measurements. As detailed in Section 5.4, the standard deviations of the noise signals is estimated by cross-validation using the bias corrected cost function $\mathcal{J}_{\rm BC}$ in (53) as a performance criterion. Figure 4 shows the bias-corrected cost $\mathcal{J}_{\rm BC}$ plotted against range of values of σ_e and σ_η. For clarity, we have shown the 2-D plot of $\mathcal{J}_{\rm BC}$ versus σ_e for different values of σ_η in Figure 4b. The cost $\mathcal{J}_{\rm BC}$ as a function of σ_η for fixed value of σ_e at which the minimum is achieved (i.e., at $\sigma_e = 0.05$) is plotted in Figure 4c. It can be seen from the figure that the minimum is achieved at the true values of σ_η and σ_e (i.e., $\sigma_e = 0.05$ and $\sigma_\eta = 0.12$).

6.2 *LPV identification of a nonlinear two-tank system*

As a second case study, we consider the identification of the nonlinear two-tank system reported in [18]. The physical system consists of two tanks, placed one above

Table 3
Example 1. Identification with noisy scheduling signal measurements: means and standard deviations (over 100 Monte-Carlo runs) of the estimated parameters using Least Squares, Bias Correction 1 and Bias Correction 2. For the sake of simplicity, the coefficients multiplying the quadratic terms in (57) are set to 0.

True Value	Least Squares	Bias Correction 1	Bias Correction 2
1	0.6908 ± 0.0070	1.0161 ± 0.0087	0.9965 ± 0.0087
-0.5	-0.3337 ± 0.0160	-0.4524 ± 0.0311	-0.4970 ± 0.0358
0.5	0.3297 ± 0.0044	0.5003 ± 0.0049	0.4948 ± 0.0051
-0.7	-0.6123 ± 0.0082	-0.6274 ± 0.0152	-0.6906 ± 0.0169
0.5	0.4970 ± 0.0021	0.5002 ± 0.0023	0.5155 ± 0.0023
-0.4	-0.3769 ± 0.0062	-0.3662 ± 0.0067	-0.4221 ± 0.0075
0.2	0.0357 ± 0.0043	0.2083 ± 0.0055	0.2058 ± 0.0057
-0.3	-0.1727 ± 0.0105	-0.2765 ± 0.0176	-0.3095 ± 0.0204

Table 4
Example 1. Identification with noisy scheduling signal observations: Best Fit Rates (BFRs) over validation data achieved by: Least-squares; Bias Correction 1 and Bias Correction 2.

Method	$\|\theta_o - \hat{\theta}\|_2$	BFR
Least Squares	0.4513	0.7784
Bias Correction 1	0.0977	0.9641
Bias Correction 2	0.0314	0.9710

the other. The upper tank receives the liquid inflow through a pump. The voltage applied to the pump is the input $u(t)$, which controls the inflow of the liquid in the upper tank. The lower tank gets the liquid inflow via a small hole at the bottom of the upper tank. The output $y(t)$ is the liquid level of the lower tank. The following nonlinear equations are used to simulate the behaviour of the system:

$$\dot{x_1}(t) = (1/A_1)(ku(t) - a_1\sqrt{2gx_1(t)}), \tag{60a}$$

$$\dot{x_2}(t) = (1/A_2)(a_1\sqrt{2gx_1(t)} - a_2\sqrt{2gx_2(t)}), \tag{60b}$$

$$y(t) = x_2(t), \tag{60c}$$

where $A_1 = 0.5$ m^2, $A_2 = 0.25$ m^2 are the cross-section areas of tank 1 and 2, respectively, $a_1 = 0.02$ m^2, $a_2 = 0.015$ m^2 are the cross-section areas of the holes in the two tanks, $g = 9.8$ m/s^2 is the acceleration due to gravity, $x_1(t)$ and $x_2(t)$ are the liquid levels in tank 1 and tank 2, respectively. The reader is referred to [18] for a more detailed description of the considered two-tank system.

The plant is controlled by a proportional controller $u = Kx_2(t)$, with $K = 1$, and the output $y(t)$ is measured with a sampling time of 0.3 s. To gather data, the closed-loop system is excited with a discrete-time zero-mean white noise reference signal $r(k)$ uniformly distributed in the interval [2 15] followed by a zero-order hold block. The measured output $y(k)$ is corrupted by a white Gaussian noise $\mathcal{N}(0, \sigma_e^2)$ with $\sigma_e = 0.01$, which corresponds to an SNR of 20 dB.

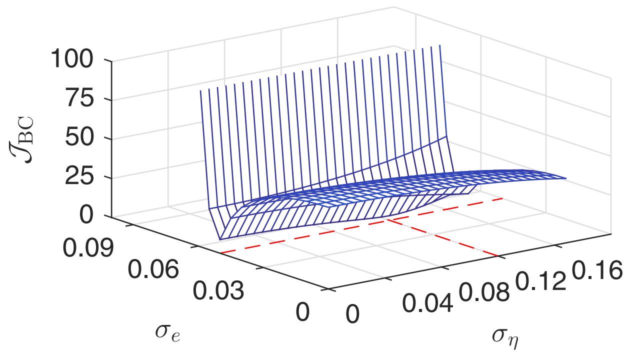
Table 5
Example 2. Best Fit Rates over validation data achieved by Least-Squares and closed-loop bias-correction.

Method	BFR
Least Squares	0.3517
Bias-correction	0.7748

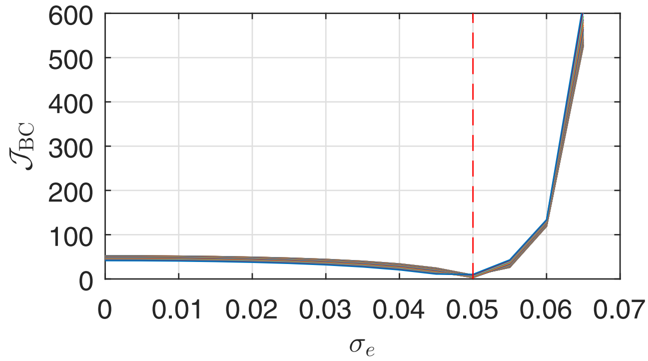
To estimate the plant, we consider the LPV model structure $\mathcal{M}_\theta$ described in (6) and (7), with $n_a = 2$, $n_b = 1$ and polynomial degree $n_g = 2$. The input $u(k-1)$ is used as a scheduling signal $p(k)$. Thus, the considered model is actually *quasi*-LPV. $N = 20,000$ and $N_{\text{val}} = 5,000$ samples are used for training and validation, respectively. The actual and simulated outputs of the models estimated through standard least-squares and the proposed bias-correction method are plotted in Figure 5. For the sake of visualization, only a subset of validation data is plotted. Furthermore, the BFRs of the estimated models are reported in Table 5. Note that, although the true system (60) does not belong to the model class $\mathcal{M}_\theta$, the proposed bias-correction approach outperforms standard least squares in estimating the dynamics of the two-tank system.

7 Conclusions

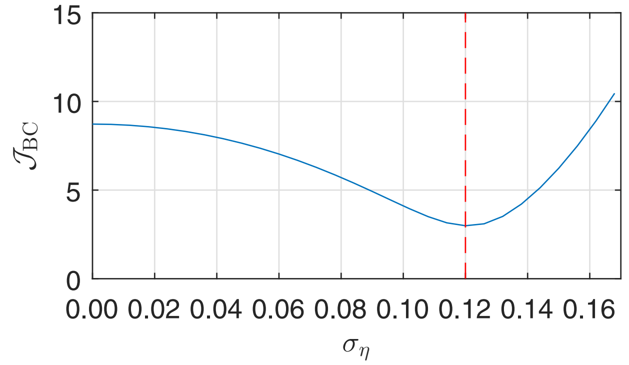
This paper has introduced a novel bias-correction approach for closed-loop identification of LPV systems. Starting from a least-square estimate, the proposed method exploits the knowledge of the controller to recursively compute an estimate of the asymptotic bias in the model parameters due to the feedback loop. This bias is then eliminated in order to obtain a consistent estimate of the open-loop plant. Based on a similar rationale, the bias caused by the noise corrupting the scheduling variable observations is also corrected, thus extending the applicability of the approach to realistic scenarios where not only the output signal, but also the scheduling signal observations are affected by a measurement noise. The computation of the bias strongly depends



(a) Bias corrected cost $\mathcal{J}_{\rm BC}$ vs σ_e and σ_η.



(b) $\mathcal{J}_{\rm BC}$ vs σ_e for differerent values of σ_η.



(c) $\mathcal{J}_{\rm BC}$ vs σ_η for fixed value of $\sigma_e = 0.05$.

Fig. 4. Example 1. Bias Corrected Cost $\mathcal{J}_{\rm BC}$ (defined in (53)) vs hyper-parameters σ_e and σ_η.

on the noise variance. In case this is not available or it cannot be estimated through dedicated experiments, a bias-corrected cost serves as a performance criterion for tuning the noise variance. The reported examples point out that the proposed method outperforms least-squares in terms of achieving a consistent estimate of the open-loop model parameters, provided that the true system belongs to the chosen model class. Although the latter assumption is barely achieved in practice, correcting the bias due to the measurement noise also leads to a significant improvement in the final model

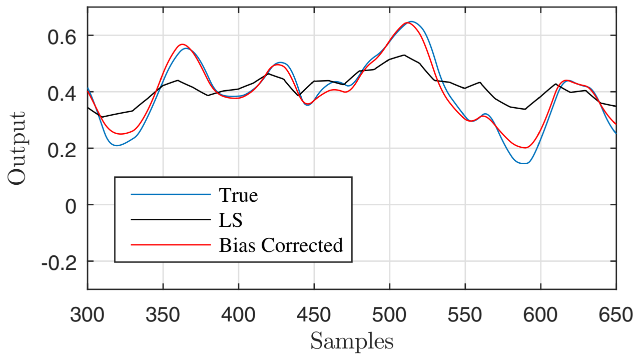


Fig. 5. Example 2. Validation dataset: true output, simulated output of the LS model, simulated output of the bias–corrected model.

estimate when an under-parametrized model structure is considered, as shown in the second case study. Future activities will be devoted to the extension of the presented approach under more general controller structures, like linear model-predictive controllers, which are characterized by piecewise-affine state-feedback control laws. Furthermore, conditions to guarantee convergence of the iterative Algorithm 1 will be sought.

A Appendix

A.1 Proof of Property 2

The a-priori known controller $\mathcal{K}_{\rm o}$ and the closed-loop structure $\mathcal{S}_{\rm o}$ in Figure 1 is exploited to construct the matrix Ψ_k, taking into account that the input signals depend on the measurement noise $e(k)$ due to the presence of feedback. Property 2 can be proved as follows. According to (20),

$$\Psi_k = \mathbb{E}\{\phi(k)\Delta\phi(k)\} = \mathbb{E}\left\{\Upsilon_k\right\} \otimes \mathbf{P}_{\rm o}(k),$$

with

$$\mathbb{E}\left\{\Upsilon_k\right\} = \mathbb{E}\left\{\chi(k)(\chi_{\rm o}(k) - \chi(k))^\top\right\}. \tag{A.1}$$

By definition of $\chi(k)$ and $\chi_{\rm o}(k)$, we have:

$$\chi_{\rm o}(k) - \chi(k) = \begin{bmatrix} e(k-1) & \cdots & e(k-n_a) & \mathbf{0}_{1\times n_b}\end{bmatrix}^\top.$$

Then,

$$\mathbb{E}\left\{\Upsilon_k\right\} = \mathbb{E}\left\{\chi(k)(\chi_{\rm o}(k) - \chi(k))^\top\right\} =$$

$$\mathbb{E}\left\{\left[\begin{array}{ccc|c} -y(k-1)e(k-1) & \cdots & -y(k-1)e(k-n_a) & \\ \vdots & -y(k-i)e(k-i) & \vdots & \mathbf{0}_{n_a\times n_b} \\ -y(k-n_a)e(k-1) & \cdots & -y(k-n_a)e(k-n_a) & \\ \hline u(k-1)e(k-1) & \cdots & u(k-1)e(k-n_a) & \\ \vdots & \ddots & \vdots & \mathbf{0}_{n_b\times n_b} \\ u(k-n_b)e(k-1) & \cdots & u(k-n_b)e(k-n_a) & \end{array}\right]\right\} \tag{A.2}$$

The following observations are made to compute $\mathbb{E}\{\Upsilon_k\}$ explicitly. The value of input and output at time k does not depend on the future values of the measurement noise e, i.e.,

$$\mathbb{E}\{y(k-i)e(k-j)\} = 0,$$
$$\mathbb{E}\{u(k-i)e(k-j)\} = 0 \quad \forall i > j.$$

This implies that the matrices Υ_k^y and Υ_k^u are upper triangular as in (24a) and (24b).

A.2 Proof of Property 3

The recurrence relations in Property 3 can be proved with the following observations:

1. Due to the strict causality of the plant $\mathcal{G}_{\mathrm{o}}$ and since e is white, the noise-free output $x(k)$ does not depend on the current and future values of the measurement noise, i.e.,

$$\mathbb{E}\{x(k-i)e(k-j)\} = 0 \quad \forall i \geq j.$$

Thus, for $i = j$,

$$\begin{aligned} &-\mathbb{E}\{y(k-i)e(k-i)\} \\ =&-\mathbb{E}\{(x(k-i)+e(k-i))e(k-i)\} \\ =&-\mathbb{E}\{x(k-i)e(k-i)\} - \mathbb{E}\{e(k-i)e(k-i)\} \\ =&0 - \sigma_e^2 = -\sigma_e^2 = f_1(k-i). \end{aligned} \tag{A.3}$$

2. The terms $f_m(k)$ and $g_m(k)$ can be computed in a recursive manner as described in the following.
Let us first consider the term $f_m(k)$. By definition:

$$f_m(k) = -\mathbb{E}\{y(k+m-1)e(k)\}.$$

By writing $y(k)$ as $x(k)+e(k)$, we have

$$\begin{aligned} f_m(k) &= -\mathbb{E}\{x(k+m-1)e(k)+e(k+m-1)e(k)\} \\ &= -\mathbb{E}\{x(k+m-1)e(k)\} \\ &= -\mathbb{E}\Bigg\{-\sum_{i=1}^{n_a} a_i^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))x(k+m-1-i)e(k) \\ &\qquad + \sum_{i=1}^{n_b} b_j^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))u(k-m-1-j)e(k)\Bigg\} \\ &= -\sum_{i=1}^{n_a} a_i^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))(-\mathbb{E}\{x(k+m-1-i)e(k)\}) \\ &\qquad - \sum_{i=1}^{n_b} b_j^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))(\mathbb{E}\{u(k-m-1-j)e(k)\}) \end{aligned}$$

with $f_1(k) = -\sigma_e^2$ (see (A.3)). Note that,

$$-\mathbb{E}\{x(k+m-1-i)e(k)\} = f_{m-i}(k),$$
$$\mathbb{E}\{u(k-m-1-j)e(k)\} = g_{m-j}(k).$$

Thus,

$$\begin{aligned} f_m(k) = &-\sum_{i=1}^{n_a} a_i^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))f_{m-i}(k) \\ &-\sum_{i=1}^{n_b} b_j^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))g_{m-j}(k). \end{aligned}$$

Since $f_m = 0$ and $g_m = 0$, for $m \leq 0$, we have

$$\begin{aligned} f_m(k) = &-\sum_{i=1}^{m-2} a_i^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))f_{m-i}(k) \\ &-\sum_{j=1}^{\min(n_b,m-1)} b_j^{\mathrm{o}}(p_{\mathrm{o}}(k+m-1))g_{m-j}(k). \end{aligned}$$

Consider now the term $g_m(k)$. Since the reference signal $r(k)$ is uncorrelated with the measurement noise $e(k)$, i.e., $\mathbb{E}(r(k)e(k')) = 0, \ \forall k, k'$, the terms $g_m(k)$ (for $m = 1, \cdots, n_a$) can be computed as

$$\begin{aligned} g_m(k) &= \mathbb{E}\{u(k+m-1)e(k)\} \\ &= \mathbb{E}\Bigg\{-\sum_{i=1}^{n_c} c_i(p_{\mathrm{o}}(k+m-1))u(k+m-1-i)e(k) \\ &\qquad + \sum_{j=0}^{n_d-1} d_{j+1}(p_{\mathrm{o}}(k+m-1))(-x(k+m-1-j)e(k))\Bigg\} \\ &= -\sum_{i=1}^{n_c} c_i(p_{\mathrm{o}}(k+m-1))(\mathbb{E}\{u(k+m-1-i)e(k)\}) \\ &\qquad + \sum_{j=0}^{n_d-1} d_{j+1}(p_{\mathrm{o}}(k+m-1))(-\mathbb{E}\{x(k+m-1-j)e(k)\}) \\ &= -\sum_{i=1}^{n_c} c_i(p_{\mathrm{o}}(k+m-1))g_{m-i}(k) \\ &\qquad + \sum_{j=0}^{n_d-1} d_{j+1}(p_{\mathrm{o}}(k+m-1))f_{m-j}(k). \end{aligned}$$

Since $f_m = 0$ and $g_m = 0$, for $m \leq 0$, it follows that

$$\begin{aligned} g_m(k) = &-\sum_{i=1}^{\min(n_c,m-1)} c_i(p_{\mathrm{o}}(k+m-1))g_{m-i}(k) \\ &+\sum_{j=1}^{\min(n_d,m)} d_j(p_{\mathrm{o}}(k+m-1))f_{m-j+1}(k). \end{aligned}$$

Thus, the recurrence relations in Property 3 are proved.

A.3 Construction of Ψ_k^p

For clarity of exposition, the procedure outlined in Section 5.2 to construct Ψ_k^p is shown via the following example.

Consider the following vector of monomials

$$\mathbf{p}_\mathrm{o}(k) = [1 \;\; p_\mathrm{o}(k) \;\; p_\mathrm{o}^2(k)]^\top, \quad \mathbf{p}(k) = [1 \;\; p(k) \;\; p^2(k)]^\top.$$

Then,

$$\mathbf{p}(k)\mathbf{p}_\mathrm{o}^\top(k) = \begin{bmatrix} 1 & p_\mathrm{o}(k) & p_\mathrm{o}^2(k) \\ p(k) & p(k)p_\mathrm{o}(k) & p(k)p_\mathrm{o}^2(k) \\ p^2(k) & p^2(k)p_\mathrm{o}(k) & p^2(k)p_\mathrm{o}^2(k) \end{bmatrix}$$

By writing $p(k)$ as $p_\mathrm{o}(k) + \eta(k)$ and taking the expectation of $\mathbf{p}(k)\mathbf{p}_\mathrm{o}^\top(k)$ w.r.t. the random variable $\eta(k)$, we get

$$\mathbb{E}[\mathbf{p}(k)\mathbf{p}_\mathrm{o}^\top(k)] = \begin{bmatrix} 1 & p_\mathrm{o}(k) & p_\mathrm{o}^2(k) \\ p_\mathrm{o}(k) & p_\mathrm{o}^2(k) & p_\mathrm{o}^3(k) \\ p_\mathrm{o}^2(k) + \sigma_\eta^2 & p_\mathrm{o}^3(k) + \sigma_\eta^2 p_\mathrm{o}(k) & p_\mathrm{o}^4(k) + \sigma_\eta^2 p_\mathrm{o}^2(k) \end{bmatrix}.$$

Then, matrix Ψ_k^p is constructed by replacing each monomial $p_\mathrm{o}^n(k)$ with the Hermite polynomial in (42), that is

$$\Psi_k^p = \begin{bmatrix} 1 & p(k) & p^2(k) - \sigma_\eta^2 \\ p(k) & p^2(k) - \sigma_\eta^2 & p^3(k) - 3\sigma_\eta^2 p(k) \\ p^2(k) & p^3(k) - 2\sigma_\eta^2 p(k) & p^4(k) - 5\sigma_\eta^2 p^2(k) + 2\sigma_\eta^4 \end{bmatrix}.$$

A.4 Proof of Property 6

Because of conditions **C2** and **C3**, we have:

$$\lim_{N_c\to\infty} \frac{1}{N_c}\sum_{k=1}^{N_c} \Omega_k(\theta_\mathrm{o}) + \left[\chi(k)\chi^\top(k)\right] \otimes \Psi_k^p = \lim_{N_c\to\infty} \frac{1}{N_c}\sum_{k=1}^{N_c} \phi_p(k)\phi_\mathrm{o}^\top(k) \;\; w.p.\ 1. \tag{A.4}$$

Let us now focus on the term $\chi(k)\hat{\chi}^\top(k) \otimes \Psi_k^p$ appearing in the bias-corrected cost (53).

For the sake of simplicity, let us assume that the initial condition $\hat{\chi}(1)$ used to simulate the bias-corrected output $\hat{y}(1)$ is known, i.e., $\hat{\chi}(1) = \chi_\mathrm{o}(1)$. This means:

$$\mathbb{E}_\eta\left\{\hat{y}(i)\right\} = y_\mathrm{o}(i) \;\; \forall i = -n_\mathrm{a}+1, \ldots, 0. \tag{A.5}$$

Let us now prove, by induction, that for $\hat{\theta}_\mathrm{CLS}^p = \theta_\mathrm{o}$,

$$\mathbb{E}_\eta\left\{\hat{y}(k)\right\} = y_\mathrm{o}(k) \;\; \forall\, k > 0. \tag{A.6}$$

Suppose that the above equation holds for $k-n_\mathrm{a}, \ldots, k-1$, i.e.,

$$\mathbb{E}_\eta\left\{\hat{y}(k-i)\right\} = y_\mathrm{o}(k-i) \;\; \forall i = 1, \ldots, n_\mathrm{a}. \tag{A.7}$$

Note that, for $\hat{\theta}_\mathrm{CLS}^p = \theta_\mathrm{o}$,

$$\begin{aligned}
\mathbb{E}_\eta\left\{\hat{y}(k)\right\} &= \theta_\mathrm{o}^\top\left(\mathbb{E}_\eta\left\{\hat{\chi}(k)\otimes\mathbf{p}^\mathrm{C}(k)\right\}\right) && \text{(A.8a)}\\
&= \theta_\mathrm{o}^\top\left(\mathbb{E}_\eta\left\{\hat{\chi}(k)\right\}\otimes\mathbb{E}_\eta\left\{\mathbf{p}^\mathrm{C}(k)\right\}\right) && \text{(A.8b)}\\
&= \theta_\mathrm{o}^\top\left(\mathbb{E}_\eta\left\{\hat{\chi}(k)\right\}\otimes\mathbf{p}_\mathrm{o}(k)\right) && \text{(A.8c)}\\
&= \theta_\mathrm{o}^\top\left[\chi_\mathrm{o}(k)\otimes\mathbf{p}_\mathrm{o}(k)\right] && \text{(A.8d)}\\
&= y_\mathrm{o}(k), && \text{(A.8e)}
\end{aligned}$$

where (A.8b) follows from white noise assumption on η and (A.8d) follows from (A.7) and construction of the bias-corrected monomials $\mathbf{p}^\mathrm{C}(k)$. Thus, from (A.5), (A.7) and (A.8), (A.6) follows by induction.

Eq. (A.6) also implies that

$$\mathbb{E}_\eta\left\{\hat{\chi}(k)\right\} = \chi_\mathrm{o}(k) \;\; \forall\, k > 0. \tag{A.9}$$

Thus,

$$\begin{aligned}
&\mathbb{E}_\eta\left\{\chi(k)\hat{\chi}^\top(k)\otimes\Psi_k^p\right\} && \text{(A.10a)}\\
=&\mathbb{E}_\eta\left\{\chi(k)\hat{\chi}^\top(k)\right\}\otimes\mathbb{E}_\eta\left\{\Psi_k^p\right\} && \text{(A.10b)}\\
=&\chi(k)\chi_\mathrm{o}^\top(k)\otimes\mathbb{E}_\eta\left\{\mathbf{p}(k)\mathbf{p}_\mathrm{o}^\top(k)\right\} && \text{(A.10c)}\\
=&\mathbb{E}_\eta\left\{\chi(k)\chi_\mathrm{o}^\top(k)\otimes\mathbf{p}(k)\mathbf{p}_\mathrm{o}^\top(k)\right\} && \text{(A.10d)}\\
=&\mathbb{E}_\eta\left\{\left(\chi(k)\otimes\mathbf{p}(k)\right)\left(\chi_\mathrm{o}(k)\otimes\mathbf{p}_\mathrm{o}(k)\right)^\top\right\} && \text{(A.10e)}\\
=&\mathbb{E}_\eta\left\{\phi_p(k)\phi_\mathrm{o}^\top(k)\right\}, && \text{(A.10f)}
\end{aligned}$$

where, (A.10c) follows from (A.9) and (43). Then, because of (A.10) and Ninness' strong law of large numbers [15], at $\hat{\theta}_\mathrm{CLS}^p = \theta_\mathrm{o}$, we have

$$\lim_{N_c\to\infty} \frac{1}{N_c}\sum_{k=1}^{N_c} \chi(k)\hat{\chi}^\top(k)\otimes\Psi_k^p = \lim_{N_c\to\infty} \frac{1}{N_c}\sum_{k=1}^{N_c} \phi_p(k)\phi_\mathrm{o}^\top(k) \;\; w.p.\ 1. \tag{A.11}$$

Thus, from (A.4) and (A.11), we obtain:

$$\lim_{N_c\to\infty} \mathcal{J}_\mathrm{BC}\left(\theta_\mathrm{o}\right) = 0 \;\; w.p.\ 1. \tag{A.12}$$

Property 6 follows from (A.12) and because of nonnegativity of the cost $\mathcal{J}_\mathrm{BC}$. This completes the proof.

Note that, even if the initial conditions are not exactly known (i.e., assumption (A.5) is not satisfied), Property 6 still holds since the error due to the mismatch between the true initial conditions and the ones used to simulate the output $\hat{y}$ vanishes asymptotically, under the assumption that the system is asymptotically stable.

References

[1] H. Abbas, M. Ali, and H. Werner. Linear recurrent neural network for open- and closed-loop consistent identification of LPV models. In *Proc. of the 49th IEEE Conference on Decision and Control*, Atlanta, USA, 2010.

[2] 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, 2009.

[3] A. Ali, M. Ali, and H. Werner. Indirect closed-loop identification of input-output LPV models: A pre-flitering approach. In *Proc. of the 18th IFAC World Congress*, pages 2606–2611, Milano, Italy, 2011.

[4] P. Apkarian, P. Gahinet, and G. Becker. Self-scheduled h_∞ control of linear parameter-varying systems: a design example. *Automatica*, 31(9):1251–1261, 1995.

[5] S. Boonto and H. Werner. Closed-loop system identification of LPV input-output models - application to an arm-driven inverted pendulum. In *Proc. of the 47th IEEE Conference on Decision and Control*, pages 2606–2611, Cancun, 2008.

[6] V. Cerone, D. Piga, and D. Regruto. Set-membership LPV model identification of vehicle lateral dynamics. *Automatica*, 47(8):1794–1799, 2011.

[7] A. Chiuso. The role of vector autoregressive modeling in predictor-based subspace identification. *Automatica*, 43(6):1034–1048, 2007.

[8] P. Van den Hof. Closed-loop issues in system identification. *Annual reviews in control*, 22:173–186, 1998.

[9] U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, 35(7):1215–1241, 1999.

[10] M. Gilson and P. Van den Hof. On the relation between a bias-eliminated least-squares (BELS) and an IV estimator in closed-loop identification. *Automatica*, 37(10):1593–1600, 2001.

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

[12] I. D. Landau and A. Karimi. Recursive algorithms for identification in closed-loop: A unified approach and evaluation. *Automatica*, 33(8):1499–1523, 1997.

[13] J. Lataire, E. Louarroudi, R. Pintelon, and Y. Rolain. Benchmark data on a linear time-and parameter-varying system. In *17th IFAC Symposium on System Identification*, pages 1477–1482, Beijing, China.

[14] L. Ljung. *System Identification: Theory for the user*. Prentice Hall, 1999.

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

[16] D. Piga, P. Cox, R. Tóth, and V. Laurain. LPV system identification under noise corrupted scheduling and output signal observations. *Automatica*, 53:329–338, 2015.

[17] D. Piga and R. Tóth. A bias-corrected estimator for nonlinear systems with output-error type model structures. *Automatica*, 50(9):2373–2380, 2014.

[18] R. S. Smith and J. Doyle. The two tank experiment: A benchmark control problem. In *Proc. American Control Conference*, pages 2026–2031, June 1988.

[19] T. Söderström and P. Stoica. *System Identification*. Prentice Hall, 1989.

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

[21] R. Tóth, V. Laurain, M. Gilson, and H. Garnier. Instrumental variable scheme for closed-loop LPV model identification. *Automatica*, 48(9):2314 –2320, 2012.

[22] 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.

[23] W. X. Zheng. A bias correction method for identification of linear dynamic errors-in-variables models. *IEEE Transactions on Automatic Control*, 47(7):1142–1147, 2002.

[24] W. X. Zheng and C. Feng. A bias-correction method for indirect identification of closed-loop systems. *Automatica*, 33(8):1499–1523, 1997.