

**Recursive identification of structured systems**  
**An instrumental-variable approach applied to mechanical systems**

Classens, Koen; González, Rodrigo A.; Oomen, Tom

**DOI**

[10.1016/j.ejcon.2025.101238](https://doi.org/10.1016/j.ejcon.2025.101238)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

European Journal of Control

**Citation (APA)**

Classens, K., González, R. A., & Oomen, T. (2025). Recursive identification of structured systems: An instrumental-variable approach applied to mechanical systems. *European Journal of Control*, 84, Article 101238. <https://doi.org/10.1016/j.ejcon.2025.101238>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



# Recursive identification of structured systems: An instrumental-variable approach applied to mechanical systems

Koen Classens<sup>a,\*,</sup>, Rodrigo A. González<sup>a</sup>, Tom Oomen<sup>a,b</sup>

<sup>a</sup> Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology, Eindhoven, The Netherlands

<sup>b</sup> Delft University of Technology, Department of Mechanical Engineering, Delft Center for Systems and Control, Delft, The Netherlands

## ARTICLE INFO

Recommended by T. Parisini

### Keywords:

Recursive identification  
Closed-loop system identification  
Refined instrumental variables  
Parameter estimation  
Parametric fault diagnosis  
Additive system

## ABSTRACT

Online system identification algorithms are widely used for monitoring, diagnostics and control by continuously adapting to time-varying dynamics. Typically, these algorithms consider a model structure that lacks parsimony and offers limited physical interpretability. The objective of this paper is to develop a real-time parameter estimation algorithm aimed at identifying time-varying dynamics within an interpretable model structure. An additive model structure is adopted for this purpose, which offers enhanced parsimony and is shown to be particularly suitable for mechanical systems. The proposed approach integrates the recursive simplified refined instrumental variable method with block-coordinate descent to minimize an exponentially-weighted output error cost function. This novel recursive identification method delivers parametric continuous-time additive models and is applicable in both open-loop and closed-loop controlled systems. Its efficacy is shown using numerical simulations and is further validated using experimental data to detect the time-varying resonance dynamics of a flexible beam system. These results demonstrate the effectiveness of the proposed approach for online and interpretable estimation for advanced monitoring and control applications.

## 1. Introduction

System identification involves obtaining mathematical models of complex phenomena from measured data, with wide-ranging applications across diverse fields (Ljung, 1999; Söderström & Stoica, 1989). The development of accurate models enables the simulation, analysis, and prediction of complex system behavior. In control engineering, models are the key to designing effective control strategies, while its predictive capabilities support decision-making and planning.

Online system identification plays a significant role in dynamic environments where systems are prone to change over time (Ljung & Söderström, 1983; Young, 2011). Unlike offline identification, which relies on a batch of data, online system identification continuously updates models in real-time as new data becomes available. Online system identification allows for quick adjustments to evolving conditions, which can enhance model accuracy, improve performance, and facilitate effective maintenance of various industrial process (Classens et al., 2023). Consequently, it is instrumental in maintaining optimal system performance, enabling timely fault detection and diagnosis, and supporting adaptive control in dynamic unpredictable settings.

Estimation of continuous-time (CT) dynamical models of industrial processes offers advantages over discrete-time (DT) models. Continuous-time models can provide a more accurate representation of the underlying dynamics of physical systems, since they allow for the direct

incorporation of *a priori* knowledge such as the relative degree of the physical systems they model (Garnier & Young, 2014). Incorporating prior knowledge in the form of interpretable physics has gained increasing importance in the fields of system identification, control, monitoring, and machine learning. In the context of linear system identification, incorporating relative degree information is particularly useful when the time evolution of the signals is naturally continuous, as seen in mechanical systems (Gawronski, 2004), where impulse response discontinuities are typically absent due to the double integration relationship between force and position.

Although a single non-additive transfer function is commonly employed in linear system identification, practical applications in flexible motion systems (Oomen, 2018; Voorhoeve et al., 2020), and vibration analysis are more effectively conceptualized as a sum of transfer functions with distinct denominators. For example, mechanical systems are interpreted more naturally in a modal form, where each submodel represents a different resonant mode (Gawronski, 2004). Additive model parametrizations bring benefits such as enhanced physical insight for fault diagnosis (Classens et al., 2022) and improved numerical conditioning for high-order systems (González, Classens et al., 2025). These parameterizations are also explored in statistics (Hastie & Tibshirani, 1986) and econometrics (Härdle et al., 2004; Opsomer

\* Corresponding author.

E-mail addresses: [k.h.j.classens@gmail.com](mailto:k.h.j.classens@gmail.com) (K. Classens), [r.a.gonzalez@tue.nl](mailto:r.a.gonzalez@tue.nl) (R.A. González), [t.a.e.oomen@tue.nl](mailto:t.a.e.oomen@tue.nl) (T. Oomen).

& Ruppert, 1999), offering increased model flexibility and the ability to decentralize analysis for optimization and control purposes. Despite contributions in nonlinear discrete-time finite-impulse response and generalized Hammerstein model estimation (Bai, 2005; Bai & Chan, 2008), the utilization of additive model structures in system identification has been limited.

The identification of additive models has recently gained increasing attention. An algorithm for offline estimation of continuous-time additive models in single-input single-output (SISO) systems, applicable to both open and closed-loop configurations, has been introduced (González, Classens et al., 2025). This approach is based on the simplified refined instrumental variable method, SRIVC (Young & Jakeman, 1980), which consists of iteratively computing instrumental variable estimates of the system parameters. Alternatively, a block coordinate algorithm is utilized for identifying these systems under this structure with a tailored block-coordinate descent version of the SRIVC method (González et al., 2023).

In contrast to these offline algorithms, many recursive algorithms have been developed by parameterizing the model as a single non-additive transfer function (Ljung & Söderström, 1983; Young, 2011). These online methods are typically based on a forgetting factor, or in state estimation mechanisms analogous to the Kalman filter. Most methods are categorized as modifications of offline identification methods, (non-)linear filtering methods, stochastic approximation methods, and pseudolinear regression-based methods (Ljung & Söderström, 1983). Methods for time-varying systems, such as Padilla et al. (2019, 2017) explore open and closed-loop approaches based on prefiltering the input and output derivatives with ad-hoc prefilters, similar to the state-variable filter and refined instrumental variable methods for linear and time-invariant systems (Garnier & Wang, 2008; Young & Jakeman, 1980).

Although current recursive methods can effectively track time-varying systems, these yield non-additive transfer function models with limited interpretability, as the estimated parameters may not directly correspond to physical system quantities. In this paper, additive linear time-varying systems are considered for which the parameter variation is slow compared to the system dynamics. To track these systems, a recursive counterpart is developed of the algorithm in González et al. (2023), where instrumental variable iterations are performed for each submodel until convergence, aligning with the philosophy of the block coordinate descent approach for non-convex optimization (Luenberger & Ye, 2008). A comprehensive identification method for modeling additive linear time-varying continuous-time systems is proposed for both open and closed-loop settings. The method is referred to as recursive SRIVC for additive systems, or abbreviated as additive RSRIVC.

In summary, the main contributions are summarized as follows.

- C1: A condition is provided under which it is preferable to estimate an additive model structure instead of an unfactored transfer function.
- C2: A recursive continuous-time system identification algorithm is derived for systems in additive form, for both open and closed-loop systems and its consistency is assessed.
- C3: The proposed method is evaluated through simulations and is experimentally validated on a flexible beam setup with a time-varying resonance mode.

The remainder of this paper is organized as follows. In Section 2 the system setup is presented and the problem is formulated. Section 3 elaborates on the benefits of the proposed approach, showcasing motivational examples and demonstrating where an additive decomposition is preferable to an unfactored transfer function. Section 4 contains the main contribution of this work, namely the derivation of a recursive identification method for continuous-time systems in additive form. The estimator is numerically evaluated in Section 5, and validated using a true mechanical system in Section 6. Lastly, Section 7 concludes this paper.

## 2. System setup and problem formulation

First, the system setup for additive systems in open-loop and closed-loop configuration is outlined, followed by the formulation of the estimation problem.

### 2.1. System setup

Consider the following continuous-time (CT), linear and time-variant (LTV), single-input single-output (SISO) system

$$x(t) = G^*(p, t)u(t), \quad (1)$$

where  $G^*(p, t) = B^*(p, t)/A^*(p, t)$ , with  $p$  being the Heaviside operator (i.e.,  $px(t) = \frac{dx(t)}{dt}$ ). The terms  $A^*(p, t)$  and  $B^*(p, t)$  are polynomials in  $p$  with time-varying coefficients of the form

$$A^*(p, t) = a_n^*(t)p^n + a_{n-1}^*(t)p^{n-1} + \dots + a_1^*(t)p + 1, \quad (2)$$

$$B^*(p, t) = b_m^*(t)p^m + b_{m-1}^*(t)p^{m-1} + \dots + b_1^*(t)p + b_0^*(t), \quad (3)$$

with  $a_n^* \neq 0$  and  $n \geq m$ . These polynomials are assumed to be coprime, i.e., for any fixed  $t$  along the parameter trajectory, they do not share roots in  $p$ .

The system in (1) is represented in its unfactored function form. Any non-additive transfer function that represents (1) can also be described as a sum of time-variant systems of lower order. Such decomposition is called an additive form of (1), which is described by

$$x_i(t) = G_i^*(p, t)u(t), \quad (4a)$$

$$x(t) = \sum_{i=1}^K x_i(t), \quad (4b)$$

where  $G_i^*(p, t) = B_i^*(p, t)/A_i^*(p, t)$ , and  $A_i^*(p, t)$  and  $B_i^*(p, t)$  have the same form as (2) and (3), but have orders  $n_i$  and  $m_i$  with  $a_{i,n_i}^* \neq 0$ ,  $n_i \geq m_i$ , and  $\sum_{i=1}^K n_i = n$ . That is,

$$A_i^*(p, t) = a_{i,n_i}^*(t)p^{n_i} + a_{i,n_i-1}^*(t)p^{n_i-1} + \dots + a_{i,1}^*(t)p + 1, \quad (5)$$

$$B_i^*(p, t) = b_{i,m_i}^*(t)p^{m_i} + b_{i,m_i-1}^*(t)p^{m_i-1} + \dots + b_{i,1}^*(t)p + b_{i,0}^*(t). \quad (6)$$

Without loss of generality, it is assumed that the polynomials  $A_i^*(p, t)$  are jointly coprime and anti-monic, i.e., their constant coefficient is fixed to one. In addition, for each separate submodel to be identifiable, the system is parameterized such that at most one subsystem  $G_i^*(p, t)$  is biproper. An approach to determine an additive decomposition of (1) is to calculate the partial fraction decomposition of  $B^*(p, t)/A^*(p, t)$ .

The polynomials  $A_i^*(p, t)$  and  $B_i^*(p, t)$  are jointly described by the parameter vector

$$\theta_i^*(t) = [a_{i,1}^*(t) \ a_{i,2}^*(t) \ \dots \ a_{i,n_i}^*(t) \ b_{i,0}^*(t) \ b_{i,1}^*(t) \ \dots \ b_{i,m_i}^*(t)]^T, \quad (7)$$

and

$$\rho^*(t) := [\theta_1^{*T}(t) \ \theta_2^{*T}(t) \ \dots \ \theta_K^{*T}(t)]^T. \quad (8)$$

A noisy output measurement is retrieved at equidistantly-spaced time instants  $t = t_k$ , see Fig. 1, i.e.,

$$y(t_k) = x(t_k) + v(t_k), \quad (9)$$

where  $v$  is a stochastic process of zero mean and finite variance  $\sigma^2$ .

If  $G^*(p, t)$  is in closed loop, see Fig. 2, it is assumed that the system is driven by a reference signal  $r(t_k)$  and the control input is determined through the control law

$$u(t_k) = C_d(q)(r(t_k) - y(t_k)), \quad (10)$$

where  $C_d(q)$  denotes the (discrete-time) controller transfer function, and  $q$  denotes the forward shift operator. It is assumed that the input sequence  $u(t_k)$  is interpolated via a zero-order hold prior to being used

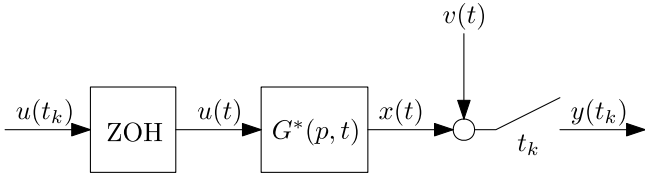


Fig. 1. Block diagram for the open-loop setting studied in this paper.

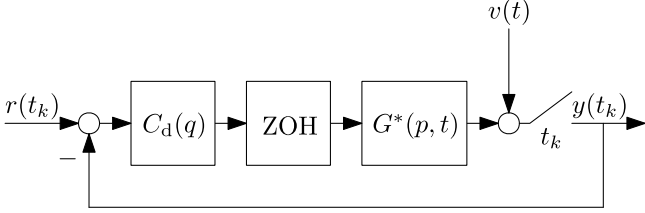


Fig. 2. Block diagram for the closed-loop setting studied in this paper.

to excite the system. Substitution of the output  $y(t_k) = G_d^*(q, t_k)u(t_k) + v(t_k)$  gives the input described as

$$u(t_k) = \tilde{r}(t_k) - \tilde{v}(t_k), \quad (11)$$

where

$$\tilde{r}(t_k) := S_{uo}^*(q, t_k)r(t_k), \quad \tilde{v}(t_k) := S_{uo}^*(q, t_k)v(t_k), \quad (12)$$

with  $S_{uo}^*(q, t_k)$  the sensitivity function  $S_{uo}^*(q, t_k) = C_d(q)/[1 + G_d^*(q, t_k)C_d(q)]$ , and  $G_d^*(q, t_k)$  being the zero-order-hold equivalent discrete-time system of  $G^*(p, t) = \sum_{i=1}^K G_i^*(p, t)$ . Explicit expressions for computing  $G_d^*(q, t_k)$  are obtained when changes in the parameters in  $A^*(p, t)$  and  $B^*(p, t)$  occur only at the sampling time instants  $t_k$ , see Tóth (2010, Chap. 6.3.1) for more details.

## 2.2. Problem formulation

Consider the CT LTV system with input  $u(t)$  and output  $y(t)$  in (4) and (9). The data-generating system is described by (9) in the open-loop setting. In case the system is closed-loop controlled, the data-generating system additionally includes (10) and the setpoint  $r(t)$ .

It is assumed that the plant is represented by the model set

$$x_i(t) = G_i(p, t, \theta_i)u(t), \quad (13a)$$

$$x(t) = \sum_{i=1}^K x_i(t), \quad (13b)$$

$$y(t_k) = x(t_k) + v(t_k). \quad (13c)$$

Here  $G_i(p, t, \theta_i) = B_i(p, t, \theta_i)/A_i(p, t, \theta_i)$ , where the polynomials  $A_i(p, t, \theta_i)$  and  $B_i(p, t, \theta_i)$  have the same structure as (5) and (6) with  $\theta_i(t) \in \mathbb{R}^{n_{\theta_i}}$ , and  $n_{\theta_i} = n_i + m_i + 1$ . The estimate of the total additive form is described by  $\beta(t) \in \mathbb{R}^{n_\beta}$ , which has the same structure as (8) and where  $n_\beta = \sum_{i=1}^K (n_i + m_i) + K$ .

The considered problem is to recursively estimate  $\beta(t)$  which characterizes the model structure given by (13), based on the input and output data  $\{u(t_k), y(t_k)\}_{k=1}^N$ . Here,  $N$  is the number of samples available at time  $t_N$  for the open-loop scenario, while  $\{r(t_k)\}_{k=1}^N$  is also known in the closed-loop setting.

Next, the primary benefits of additive parameterizations are presented, supported by multiple motivating examples.

## 3. Parsimony and interpretability in continuous-time system identification

The estimation in an additive form has several benefits including the following.

1. Direct continuous-time identification methods may suffer from a lack of parsimony when the sum of transfer functions of particular relative degrees is estimated (González et al., 2023).
2. The parameters in an additive structure may be more interpretable compared to an unfactored transfer function.
3. Typically, derivatives of the input and output data are required for recursive estimation. Considering an additive structure can significantly reduce the order of the derivatives to be estimated from sampled data.

First, the advantages with respect to parsimony and interpretability are described. Later, in Section 4.2, Remark 4.2, the additional advantage regarding the derivative order is highlighted.

Direct continuous-time identification approaches such as the SRIVC method may suffer from a lack of parsimony when identifying a system that consists of the sum of transfer functions. The following proposition quantifies the additional parameters being estimated when opting to estimate (1) with a model structure  $G(p) = B(p)/A(p)$  of relative degree  $r$  instead of separately estimating the parameters for each transfer function  $B_i(p)/A_i(p)$ .

**Proposition 1** (González et al., 2023). Consider the system in (1), and the model structure  $G(p) = \sum_{i=1}^K B_i(p)/A_i(p)$ . Opting for the model structure  $G(p) = B(p)/A(p)$  during identification (with the minimal relative degree  $r$  that encompasses the true system) results in a lack of parsimony for the latter model structure if and only if

$$\sum_{i=1}^K r_i - r > K - 1, \quad (14)$$

where  $r_i = n_i - m_i$  is the relative degree of  $B_i(p)/A_i(p)$ . The surplus in (14) is the number of additional parameters considered by the model structure  $G(p) = B(p)/A(p)$ .

An immediate implication derived from Proposition 1 is that when  $r = 0$  or  $r = 1$ , the model structure  $G(p) = B(p)/A(p)$  suffers from a lack of parsimony if a transfer function  $G_i^*(p)$  exists in its partial fraction expansion (1) with a relative degree greater than one. Two examples of this property are given below.

**Example 3.1.** Consider the system

$$G^*(p) = \frac{b_{1,0}^*}{a_{1,1}^*p + 1} + \frac{b_{2,0}^*}{a_{2,2}^*p^2 + a_{2,1}^*p + 1},$$

where the  $a_{i,j}^*$  values are strictly positive, and the  $b_{i,0}^*$  are non-zero. Only 5 parameters must be estimated if the following model structure is used:

$$G(p) = \frac{b_{1,0}}{a_{1,1}p + 1} + \frac{b_{2,0}}{a_{2,2}p^2 + a_{2,1}p + 1}. \quad (15)$$

Alternatively, if the consideration of this partial fraction decomposition is disregarded and the decision is made to employ the model structure

$$G(p) = \frac{b_2p^2 + b_1p + b_0}{a_3p^3 + a_2p^2 + a_1p + 1}$$

with no constraints on the parameter values, then 6 parameters must be estimated. This model structure leads to a lack of parsimony compared to the one in (15).  $\Delta$

**Example 3.2.** Consider the system

$$G^*(p) = \frac{b_0^*}{a_2^*p^2 + a_1^*p + 1},$$

where  $a_1^*$  and  $a_2^*$  are strictly positive, and  $a_1^* > 2\sqrt{a_2^*}$ . Hence,  $G^*(p)$  has two real-valued poles. If the following partial fraction decomposition of  $B^*(p)/A^*(p)$  is considered

$$G(p) = \frac{b_{1,0}}{a_{1,1}p + 1} + \frac{b_{2,0}}{a_{2,1}p + 1},$$

then 4 parameters must be estimated instead of 3. This additive representation does not directly take into account the relationship between the numerator terms  $b_{1,0}$  and  $b_{2,0}$ , leading to a lack of parsimony.  $\Delta$

The following example shows a mechanical system in one of its canonical forms. In addition to parsimony for this type of system, the interpretability of the parameters is also examined.

**Example 3.3.** Mechanical systems are typically described as a sum of modes, i.e.,

$$G^*(p) = \sum_{i=1}^K \frac{b_{i,0}^*}{p^2/\omega_i^{*2} + 2(\zeta_i^*/\omega_i^*)p + 1}, \quad (16)$$

where  $\omega_i^* > 0$  and  $\zeta_i^* > 0$  denote the natural frequency and the damping coefficient of the  $i$ th subsystem (Gawronski, 2004). These systems are described by  $3K$  parameters if the following model structure is used:

$$G(p) = \sum_{i=1}^K \frac{b_{i,0}}{a_{i,2}p^2 + a_{i,1}p + 1}. \quad (17)$$

On the contrary, if the partial fraction decomposition is overlooked and the decision is made to adopt the widely utilized model structure of order  $n$  and relative degree two

$$G(p) = \frac{b_{n-2}p^{n-2} + b_{n-3}p^{n-3} + \dots + b_0}{a_n p^n + a_{n-1}p^{n-1} + \dots + a_1 p + 1}, \quad (18)$$

with  $n = 2K$  and no constraints on the parameter values, then a total of  $4K - 1$  parameters must be estimated. This model structure leads to a lack of parsimony compared to the one in (17) if  $K > 1$ , with a further degradation of parsimony for larger values of  $K$ . Furthermore, the parameters in (17) are physically interpretable, as they directly relate to the natural frequencies  $\omega_i^*$  and the damping coefficients  $\zeta_i^*$ . This interpretability is hidden and is lost mostly in (18).  $\Delta$

Interpretability of the parameters, as illustrated in the latter example, is highly favorable in many engineering applications. For example, in the field of fault diagnosis, this interpretability greatly simplifies the task of fault isolation, as faults often manifest in a specific submodel. Estimating the system in an unfactored form leads to variations in most parameters, while estimating it in an additive form results in variations solely in the directly related parameters.

#### 4. Recursive estimation of continuous-time systems: A block coordinate descent method

This section provides a recursive solution to estimate additive models. A key optimization tool that is considered in the approach is the block-descent method, which is examined in the next subsection. Then, the recursive estimators are presented for both open and closed-loop settings. In addition, a summary of the algorithm is presented, adaptations for marginally stable systems are given, and practical aspects are described.

##### 4.1. Block coordinate descent

Towards the goal of computing a recursive estimator for  $\beta^*$ , consider the minimization problem

$$\min_{\substack{\theta_i \in \Omega_i \\ i=1, \dots, K}} V_N(\beta), \quad (19)$$

where  $\Omega_i \subset \mathbb{R}^{n_{\theta_i}}$  is a compact set which contains the true parameters of the  $i$ th subsystem. The cost  $V_N$ , using the  $N$  available samples at  $t_N$ , is the weighted least squares output error cost function

$$V_N(\beta) = \sum_{k=1}^N \alpha_{N-k} \left[ y(t_k) - \sum_{i=1}^K G_i(p, t, \theta_i) u(t_k) \right]^2, \quad (20)$$

where the notation  $G_i(p, t, \theta_i)u(t_k)$  means that the input  $u(t)$  is filtered through the continuous-time transfer function  $G_i(p, t, \theta_i)$  and later evaluated at  $t = t_k$ . The function  $\alpha$  is a non-negative weighting function of the form

$$\alpha_k = \lambda^k, \quad \lambda \in (0, 1]. \quad (21)$$

Note that  $\alpha$  satisfies the multiplicative property  $\alpha_{k+1} = \lambda \alpha_k$ , with  $\alpha_0 = 1$ . The parameter  $\lambda$  is often referred to as the forgetting factor (Ljung, 1999).

The approach to solving (19) consists of iteratively minimizing the cost with respect to  $\theta_i$  while leaving the other variables fixed. Algorithm 1 describes the general procedure that is solved, typically after every incoming sample, known as *block coordinate descent* (Luenberger & Ye, 2008).

---

##### Algorithm 1 Block coordinate descent

---

- 1: Choose an initial parameter vector  $\theta_i^1$  for each  $i$
  - 2: **for**  $l = 1, 2, \dots$  **do**
  - 3:   **for**  $i = 1, \dots, K$  **do**
  - 4:      $\theta_i^{l+1} \leftarrow \arg \min_{\theta_i \in \Omega_i} V_N(\theta_1^{l+1}, \dots, \theta_{i-1}^{l+1}, \theta_i, \theta_{i+1}^l, \dots, \theta_K^l)$
  - 5:   **end for**
  - 6: **end for**
- 

Let  $\beta^l := [(\theta_1^l)^\top, \dots, (\theta_K^l)^\top]^\top \in \Omega \subset \mathbb{R}^{n_\beta}$ , with  $\Omega = \prod_{i=1}^K \Omega_i$  being the parameter space. Therefore, each iteration of the algorithm (in  $l$ ) can be written as  $\beta^{l+1} = \mathcal{A}(\beta^l)$ , where  $\mathcal{A}: \mathbb{R}^{n_\beta} \rightarrow \mathbb{R}^{n_\beta}$  is the composite mapping  $\mathcal{A} = \mathcal{T} \circ \mathcal{C}^K \circ \mathcal{T} \circ \mathcal{C}^{K-1} \circ \dots \circ \mathcal{T} \circ \mathcal{C}^1$ . This notation is adopted from (Luenberger & Ye, 2008), where the block coordinate descent iterations are referred to as  $\mathcal{A}$ -iterations. Here, the choice function is defined as  $\mathcal{C}^i(\beta) := (\beta, i)$ , and the optimization function  $\mathcal{T}(\beta, i) := (\theta_1, \dots, \theta_{i-1}, \bar{\theta}_i, \theta_{i+1}, \dots, \theta_K)$ , with  $\bar{\theta}_i$  being defined as  $\bar{\theta}_i = \arg \min_{\theta_i \in \Omega_i} V_N(\theta_1, \dots, \theta_i, \dots, \theta_K)$ . The following result concerns the convergence of Algorithm 1 to a stationary point of the cost (19).

**Theorem 1** (Global Convergence of Algorithm 1). *If the search along any coordinate direction  $\theta_i$  yields a unique minimum point of  $V_N$ , then the limit of any convergent subsequence of  $\{\beta^l\}$  obtained from  $\beta^{l+1} = \mathcal{A}(\beta^l)$  belongs to the set of fixed points  $\Gamma_N = \{\beta \in \Omega : \nabla V_N(\beta) = 0\}$ .*

**Proof.** See, e.g., Luenberger and Ye (2008, Section 8.9).  $\square$

**Remark 4.1.** The result in Theorem 1 still holds if the unique minimum assumption is replaced with the requirement that  $V_N$  decreases in each coordinate search. In this case, it is sufficient to find a parameter vector  $\bar{\theta}_i$  that reduces the cost  $V_N$  instead of minimizing it.

The block coordinate descent algorithm described in Algorithm 1 requires a way to compute  $\mathcal{T}(\beta, i)$  at each iteration, for each  $i = 1, 2, \dots, K$ . That is, given the data up until  $t_N$ , compute

$$\begin{aligned} \theta_i^{l+1} = \arg \min_{\theta_i \in \Omega_i} \sum_{k=1}^N \alpha_{N-k} \left[ y(t_k) - \sum_{j=1}^{i-1} G_j(p, t, \theta_j^{l+1}) u(t_k) \right. \\ \left. - \sum_{j=i+1}^K G_j(p, t, \theta_j^l) u(t_k) - G_i(p, t, \theta_i) u(t_k) \right]^2 \end{aligned} \quad (22)$$

for  $i = 1, 2, \dots, K$ , for fixed values of  $\{\theta_j^{l+1}\}_{j=1}^{i-1}$  and  $\{\theta_j^l\}_{j=i+1}^K$ . For a sufficiently large compact parameter space  $\Omega_i$ , the optimization problem in (22) becomes a nonlinear least-squares problem that lends itself to iterative solution methods. To this end, define the residual output of each submodel as

$$\tilde{y}_i(t_k) := y(t_k) - \sum_{j=1}^{i-1} G_j(p, t, \theta_j^{l+1}) u(t_k) - \sum_{j=i+1}^K G_j(p, t, \theta_j^l) u(t_k).$$



Substitution of this residual output in (22) gives

$$\theta_i^{l+1} = \arg \min_{\theta_i \in \Omega_i} \sum_{k=1}^N \alpha_{N-k} \left[ \tilde{y}_i(t_k) - G_i(p, t, \theta_i) u(t_k) \right]^2. \quad (23)$$

Next is shown how this is solved through recursive SRIVC.

#### 4.2. Recursive SRIVC for additive open-loop systems

In the linear and time-invariant case, when solved iteratively, the optimization problem in (23) leads to a local minimum for the complete optimization problem (19) (González et al., 2023). Each optimization problem in (23) takes the form of minimizing the weighted sum of squares of the residual equal to  $\varepsilon_i(t_k, \theta_i) = \tilde{y}_i(t_k) - G_i(p, t, \theta_i) u(t_k)$ . This pseudolinear regression takes the general form

$$\tilde{y}_{i,f}(t_k, \theta_i) = \varphi_{i,f}^\top(t_k, \theta_i) \theta_i(t_k) + \varepsilon_i(t_k, \theta_i). \quad (24)$$

This generic form admits the refined instrumental variable iterations (Young & Jakeman, 1980). Essentially, the refined instrumental variable iterations provide a sequence which, under mild conditions, converges to a point  $\bar{\theta}_i$  that satisfies the first order optimality condition of the cost in (23), i.e.,

$$\sum_{k=1}^N \alpha_{N-k} \frac{\partial \varepsilon_i(t_k, \theta_i)}{\partial \theta_i} \Big|_{\theta_i = \bar{\theta}_i} \varepsilon_i(t_k, \bar{\theta}_i) = 0.$$

The instrumental variable refinements, from now denoted as SRIVC iterations, are denoted using the iterate counter  $s$ , as

$$\begin{aligned} \theta_{i,s}^{l+1} &= \left[ \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}) \right]^{-1} \\ &\times \left[ \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}) \right]. \end{aligned} \quad (25)$$

Appendix A shows how (23) fits the pseudolinear regression framework with (24) and how this leads to SRIVC iterations in (25). In (25), the subscripts  $i$  and  $l$  refer to the current block-coordinate descent iteration. The filtered residual output, regressor vector and instrument vector, indicated with subscript  $f$ , are respectively given by

$$\tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}) = \frac{1}{A_i(p, \theta_{i,s}^{l+1})} \tilde{y}_i(t_k), \quad (26)$$

$$\varphi_{i,f}(t_k, \theta_{i,s}^{l+1}) = \left[ \begin{array}{ccc} \frac{-p}{A_i(p, \theta_{i,s}^{l+1})} \tilde{y}_i(t_k) & \dots & \frac{-p^{n_i}}{A_i(p, \theta_{i,s}^{l+1})} \tilde{y}_i(t_k) \\ \frac{1}{A_i(p, \theta_{i,s}^{l+1})} u(t_k) & \dots & \frac{p^{m_i}}{A_i(p, \theta_{i,s}^{l+1})} u(t_k) \end{array} \right]^\top, \quad (27)$$

$$\hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) = \left[ \begin{array}{ccc} \frac{-p}{A_i(p, \theta_{i,s}^{l+1})} \hat{x}_i(t_k) & \dots & \frac{-p^{n_i}}{A_i(p, \theta_{i,s}^{l+1})} \hat{x}_i(t_k) \\ \frac{1}{A_i(p, \theta_{i,s}^{l+1})} u(t_k) & \dots & \frac{p^{m_i}}{A_i(p, \theta_{i,s}^{l+1})} u(t_k) \end{array} \right]^\top, \quad (28)$$

where  $\hat{x}_i(t_k)$  is the estimate of  $x_i(t_k)$ . This estimate is based on the latest parameter estimate of the previous time-sample, denoted by  $\bar{\theta}_i(t_{k-1})$ , i.e.,

$$\hat{x}_i(t_k) = \frac{B_i(p, \bar{\theta}_i(t_{k-1}))}{A_i(p, \bar{\theta}_i(t_{k-1}))} u(t_k).$$

To clarify, the refinement step refers to updating the denominator polynomial of the prefiltering step in (26) to (28). Hence, the filtered output, regressor vector, and instrument vector, are updated based on the estimates obtained from the previous iterate. The instrument vector is similar to the regressor vector but uses  $\hat{x}_i$  instead of  $\tilde{y}_i$ , i.e., an estimate of the output of the submodel instead of using the noisy measured variable.

**Remark 4.2.** The regressor (27) and the instrument (28) require the computation of derivatives up to order  $m_i$  for the input and  $n_i$  for the

residual output. Note that the order of the required derivatives is significantly lower compared to the case where a single unfactored transfer function would have been used. In this scenario, given that  $\sum_{i=1}^K n_i = n$ , obtaining the derivative of the  $n$ th order of the output is necessary for standard recursive continuous-time identification methods, along with a substantially higher derivative of the input.

**Remark 4.3.** Note that the expressions in (25) to (28) give the standard SRIVC iterations when  $\tilde{y}(t_k) = y(t_k)$  (i.e., when  $K = 1$  and all the modes of the composite transfer function are estimated jointly) and  $\alpha_{N-k} = 1$  for all integer  $k$  less than  $N$  (i.e., when there is no forgetting factor).

**Remark 4.4.** The expressions provided in (26), (27) and (28) require prefiltering the data, relying implicitly on the commutativity property given by

$$cG(p)x(t) = G(p)[cx(t)], \quad (29)$$

where  $c$  is a constant. It is not recommended to prefilter the data by transfer functions when the parameters exhibit rapid temporal variations, since the property (29) does not hold in general when  $c$  is time-varying (Tóth, 2010). However, it is important to note that (29) is a useful approximation in cases where the parameters exhibit slow temporal variations.

Instead of computing (25), an equivalent recursion is derived which does not involve the summation operator over all available samples and allows to achieve an update of the estimate after every incoming time sample. This recursion involves computing

$$K_{i,s+1}^{l+1}(t_N) = \frac{\bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})}{\lambda + \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})}, \quad (30)$$

$$\begin{aligned} \theta_{i,s+1}^{l+1}(t_N) &= \bar{\theta}_i(t_{N-1}) + K_{i,s+1}^{l+1}(t_N) \\ &\times \left( \tilde{y}_{i,f}(t_N, \theta_{i,s}^{l+1}) - \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{\theta}_i(t_{N-1}) \right), \end{aligned} \quad (31)$$

and

$$P_{i,s+1}^{l+1}(t_N) = \frac{1}{\lambda} \left( \bar{P}_i(t_{N-1}) - K_{i,s+1}^{l+1}(t_N) \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1}) \right), \quad (32)$$

where  $\bar{\theta}_i(t_{N-1})$ , and  $\bar{P}_i(t_{N-1})$  denote the last parameter estimate and inverse of the matrix being inverted in (25) at the previous time instance  $t_{N-1}$  after the  $\mathcal{A}$ -iterations over  $l$  and  $i$ , and after the SRIVC iterations using  $s$ . The derivation of this recursion is presented in Appendix B. Note that  $\bar{P}_i(t_{N-1})$  is interpreted as an approximate covariance matrix of the resulting estimate, see Young (2015) and Pan, Welsh et al. (2020) for more details. A summary of the algorithm is provided in Section 4.4.

#### 4.3. Consistency analysis with $\alpha_k = 1$

First, the conditions under which the additive RSIVC approach produces a consistent estimator are analyzed. To this end, consider the following assumptions to facilitate the analysis, theorem, and related remarks. These assumptions are standard in the SRIVC analysis literature, see, e.g., González et al. (2024), Pan, González et al. (2020) and González, Pan et al. (2025). The scenario with  $\alpha_k = 1$  allows to employ a similar reasoning as the non-recursive SRIVC approach as presented in Pan, González et al. (2020), although special caution is taken to address the time-varying nature of the models used for constructing the instrument and regressor vectors.

**Assumption 1 (BIBO Stability).** The true system and the estimated models are uniformly bounded-input, bounded-output (BIBO) stable, i.e., for any initial time instant  $t_0$  and bounded input signal, the output sequence  $\{y(t_k)\}_{k \geq 0}$  is bounded (Rugh, 1996).

**Assumption 2 (Persistence of Excitation).** The input sequence is persistently exciting of order no less than  $2n$ .

**Assumption 3 (Coprimeness and Sampling Frequency).** The additive subsystems pertaining to the true system and the models obtained through the descent iterations do not have zero-pole cancellations. Furthermore, the sampling frequency is more than twice the largest positive imaginary part of the zeros of  $\prod_{i=1}^K A_i(p, \theta_{i,s}^l) A_i^*(p)$ . The subscript  $s$  and superscript  $l$  relate to the SRIVC iterations and block-coordinate descent iterations, respectively, defined in Section 4.

**Assumption 4 (Independence and Stationarity).** The input sequence  $\{u(t_k)\}$  is quasi-stationary (Ljung, 1999, p. 34), the disturbance  $\{v(t_l)\}$  is stationary, and they are mutually independent for all integers  $k$  and  $l$ .

**Assumption 5 (Parameter Trajectories).** The changes in the parameters occur only at the sampling instants  $t_k$ , i.e., the parameter trajectories are piecewise constant.

**Theorem 2.** Consider the proposed estimator for the open-loop setting in Fig. 1 with  $\alpha_k = 1$  for all  $k$ , and suppose the following assumptions, Assumptions 1 to 5 hold. In addition, assume that the SRIVC iterations of the estimator for all  $N$  sufficiently large converge, where the converged parameter is denoted by  $\theta_{i,\infty}^{l+1}$ . Then, if both the proposed model and system converge as  $N \rightarrow \infty$  to linear and time-invariant transfer functions, then their limits are equal with probability 1 and generically consistent<sup>1</sup> with respect to the limiting system and model parameters.

The proof is provided in Appendix C.

**Remark 4.5.** Note that consistency is only considered for  $\alpha_k = 1$  for all  $k$ . This is a special case of (21). In case (21) is used with a different constant-rate-forgetting factor  $\lambda < 1$ , the effect of noise does not disappear for  $N \rightarrow \infty$ . Therefore, considering more data does not reduce the influence of measurement noise making the estimator not consistent. Variable-rate-forgetting factors may be considered, ensuring consistent estimators (see Bruce et al., 2020 for details); however, these considerations are beyond the scope of this work.

#### 4.4. Summary iterative recursive estimation algorithm

The proposed additive RSRIVC algorithm involves nested iterations in  $l$ ,  $i$ , and  $s$ . To facilitate implementation, the nested iterations are discussed in this subsection.

First of all, the recursive nature allows to update the estimate after every incoming sample. Therefore, these iterations will be performed at the discrete times  $t_k$ , where  $k$  denotes the count in the discrete-time domain. At each of these points in time, the  $\mathcal{A}$ -iterations are performed  $M_l$  times, see Section 4.1, i.e., the block-coordinate method runs through each of  $K$  additive models for  $M_l$  times. The current additive model is denoted by  $i$  and the  $\mathcal{A}$ -iteration is denoted by  $l$ . Finally, within each of these parameter updates,  $M_s$  SRIVC iterations are performed, with the current SRIVC iterate being indicated by  $s$ .

To summarize, compute the recursive formulas in (30), (31) and (32) at each time step and for every additive model  $G_i$  to solve the additive model identification problem in (23). This way, the proposed method solves (23) recursively. The algorithm is presented in Algorithm 2. Note that the regressor vector has an explicit dependence on the parameter vector of the other models being estimated; this dependence is implicit in the equations via the expression for the residual  $\tilde{y}_i$ . A block-diagram of the block-coordinate descent RSRIVC algorithm for additive systems is shown in Fig. 3. Two distinct blocks are highlighted in this diagram, which refer to the SRIVC and the block-coordinate descent step. These are executed after every incoming sample. During

**Algorithm 2** Recursive estimator for open-loop continuous-time systems in additive form

```

1: Initialization: Set  $M_l$ ,  $M_s$ ,  $\lambda$ . Initialize  $\bar{\theta}_i(t_0)$ , and  $\bar{P}_i(t_0)$  for
   each  $i = 1, \dots, K$ , by using the first part of the input-
   output data  $\{u(t_k), y(t_k)\}_{k=-N_{\text{init}}}^0$  with any linear and time-invariant
   continuous-time system identification method. With this  $\bar{\theta}_i(t_0)$ ,
   compute  $\varphi_{i,f}$ ,  $\hat{\varphi}_{i,f}$ , and  $\tilde{y}_{i,f}$  at  $t_0$  by filtering the input-output
   data  $\{u(t_k), y(t_k)\}_{k=-N_{\text{init}}}^0$  through (26)-(28) using the denominator
    $A_i(p, \bar{\theta}_i(t_0))$ .
2: for  $k = 1, \dots$  do
3:    $\theta_i^1(t_k) \leftarrow \bar{\theta}_i(t_{k-1})$  for all  $i$ 
4:   for  $l = 1, \dots, M_l$  do
5:     for  $i = 1, \dots, K$  do
6:        $\theta_{i,0}^{l+1}(t_k) \leftarrow \theta_i^l(t_k)$ 
7:       for  $s = 0, 1, \dots, M_s$  do
8:         Parameter  $\theta_{i,s+1}^{l+1}(t_k)$  update:

$$K_{i,s+1}^{l+1}(t_k) \leftarrow \frac{\bar{P}_i(t_{k-1}) \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1})}{\lambda + \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}) \bar{P}_i(t_{k-1}) \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1})},$$


$$\theta_{i,s+1}^{l+1}(t_k) \leftarrow \bar{\theta}_i(t_{k-1}) + K_{i,s+1}^{l+1}(t_k) \left( \tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}) - \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}) \bar{\theta}_i(t_{k-1}) \right),$$


$$P_{i,s+1}^{l+1}(t_k) \leftarrow \frac{1}{\lambda} \bar{P}_i(t_{k-1}) - \frac{1}{\lambda} K_{i,s+1}^{l+1}(t_k) \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}) \bar{P}_i(t_{k-1}).$$

9:       end for
10:       $\theta_i^{l+1}(t_k) \leftarrow \theta_{i,M_s+1}^{l+1}(t_k)$ 
11:       $P_i^{l+1}(t_k) \leftarrow P_{i,M_s+1}^{l+1}(t_k)$ 
12:    end for
13:  end for
14:   $\bar{\theta}_i(t_k) = \theta_i^{M_l+1}(t_k)$  for all  $i$ 
15:   $\bar{P}_i(t_k) = P_i^{M_l+1}(t_k)$  for all  $i$ 
16: end for
17: Output: the parameter estimates  $\{\bar{\theta}_i(t_k)\}_{k \geq 1}$ .

```

these iterations, the quantities  $\tilde{y}_{i,f}$ ,  $\varphi_{i,f}$ ,  $\hat{\varphi}_{i,f}$  are updated, which only involves filtering between  $t_{k-1}$  and  $t_k$ . Optionally, a stability test is included which is further examined in Section 4.7.

**Remark 4.6.** Note that the SRIVC refinements in  $s$  are optional. In case the refinements are omitted, the optimization problem (23) is not being solved until convergence at every time-step, but rather a quasi-Newton descent algorithm of the form (25) is being implemented (Young, 2015). Despite this, the iterations stem from the fact that solving (23) and iterating over the additive models is a viable way to derive a model that fully minimizes the cost in (19).

**Remark 4.7.** While not inherently restrictive to the method, it is practically advantageous if parameters change gradually compared to the sampling rate. In general, slower parameter variation improves the effectiveness of the methods as it allows reducing the number of iterations  $M_l$  and  $M_s$ , which in turn reduces computational burden and improves parameter adaptation efficiency and robustness.

#### 4.5. Closed-loop case

The identical recursion as in (30), (31) and (32) applies to the closed-loop scenario, albeit with a different instrument vector. Employing an output error cost for system identification is known to introduce asymptotic bias in the presence of output measurement noise (Van den Hof, 1998). Consequently, instrumental variable methods are used in the closed-loop setting to mitigate this bias (Gilson et al., 2011; González, Pan et al., 2025). Instead of (23), the instrumental variable

<sup>1</sup> A statement  $s$ , which depends on the elements  $\rho$  of some open set  $\Omega \subseteq \mathbb{R}^n$ , is *generically true with respect to*  $\Omega$  (Söderström & Stoica, 1983) if the set  $\{\rho \in \Omega | s(\rho) \text{ is not true}\}$  has Lebesgue measure zero in  $\Omega$ .

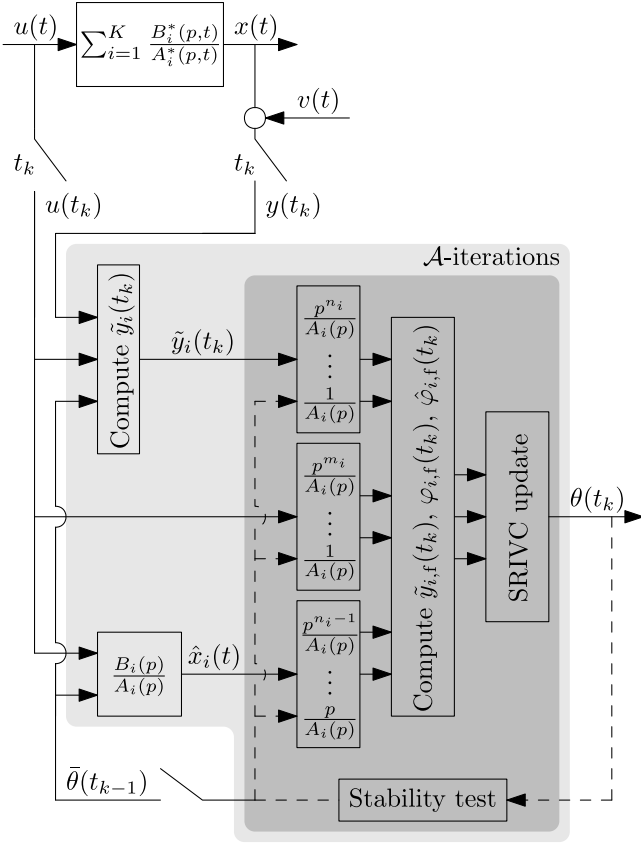


Fig. 3. Block diagram of the block-coordinate descent RSRIVC for open-loop systems, summarized by Algorithm 2. The section featuring SRIVC iterations is highlighted (■), see (25) to (28), alongside the block containing the  $\mathcal{A}$ -iterations (□), described in Section 4.1.

solution is employed which is given by

$$\theta_i^{l+1} \in \underset{\theta_i \in \Omega_i}{\text{sol}} \left\{ \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_i) \left( \tilde{y}_{i,f}(t_k, \theta_i) - \varphi_{i,f}^T(t_k, \theta_i) \theta_i(t_k) \right) = 0 \right\},$$

with  $\hat{\varphi}_{i,f}$  being the filtered instrument vector related to additive model  $i$ . Users have the flexibility to choose from various instrument vectors (Söderström & Stoica, 1983). In this context, the instrument is constructed as a noiseless version of the regressor, a choice recognized in similar identification contexts for yielding minimal asymptotic covariance among all instrumental variable estimators (Boeren et al., 2018; Gilson et al., 2008; González, Classens et al., 2025; Mooren et al., 2023). To this end, the instrument (28) is rewritten as

$$\hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) = \begin{bmatrix} \frac{-p}{A_i(p, \theta_{i,s}^{l+1})} \frac{B_i(p, \bar{\theta}_i(t_{k-1}))}{A_i(p, \bar{\theta}_i(t_{k-1}))} & \dots & \frac{-p^{n_i}}{A_i(p, \theta_{i,s}^{l+1})} \frac{B_i(p, \bar{\theta}_i(t_{k-1}))}{A_i(p, \bar{\theta}_i(t_{k-1}))} \\ \frac{1}{A_i(p, \theta_{i,s}^{l+1})} & \dots & \frac{p^{m_i}}{A_i(p, \theta_{i,s}^{l+1})} \end{bmatrix}^T z(t_k), \quad (33)$$

however, instead of  $z(t_k) = u(t_k)$ , the noiseless input is estimated as  $z(t_k) = S_{uo}(q, t_k)r(t_k)$ , where the sensitivity function  $S_{uo}(q, t_k) = C_d(q)/[1 + C_d(q)G_d(q, t_k)]^{-1}$ . Hence, the instrument now depends on an estimate of the control input, computed by means of a model of the sensitivity and the known setpoint  $r(t_k)$ .

#### 4.6. Marginally stable systems

Mechatronic position systems, frequently encountered in engineering applications, lend themselves to a decomposition comprising both rigid body and flexible modes. Notably, the rigid body modes are considered as double integrators, leading to eigenvalues precisely situated at the origin. The denominator polynomial of such marginally stable system cannot be represented in an anti-monic form as in (5) and (6). To this end, the proposed parametrization is slightly adapted to accommodate for the estimation of marginally stable systems as follows.

Consider the system (4), but where the first submodel has  $\ell$  poles at the origin. In other words, let  $G_1^*(p, t) = B_1^*(p, t)/[p^\ell A_1^*(p, t)]$ , where  $A_1^*(p, t)$  and  $B_1^*(p, t)$  are co-prime, and have the same form as in (2) and (3). Let the remaining submodels be parameterized as originally defined in Section 2.2. For either open or closed-loop variants of the proposed approach, the computation of the gradient of each submodel with respect to their respective parameter vector is required. For  $G_1(p)$ , this computation results in the modified instrument vector

$$\hat{\varphi}_{1,f}(t_k, \theta_{1,s}^{l+1}) = \begin{bmatrix} \frac{-p}{p^\ell A_1(p, \theta_{1,s}^{l+1})} \frac{B_1(p, \bar{\theta}_1(t_{k-1}))}{A_1(p, \bar{\theta}_1(t_{k-1}))} & \dots & \frac{-p^{n_1}}{p^\ell A_1(p, \theta_{1,s}^{l+1})} \frac{B_1(p, \bar{\theta}_1(t_{k-1}))}{A_1(p, \bar{\theta}_1(t_{k-1}))} \\ \frac{1}{p^\ell A_1(p, \theta_{1,s}^{l+1})} & \dots & \frac{p^{m_1}}{p^\ell A_1(p, \theta_{1,s}^{l+1})} \end{bmatrix}^T z(t_k), \quad (34)$$

where  $z(t_k) = u(t_k)$  for the open-loop algorithm, and  $z(t_k) = S_{uo}(q, t_k)r(t_k)$  for the closed-loop variant. On the other hand, the model residual retains the same form as in 4.1, with the filtered residual output given by (24), but with the regressor vector now expressed as

$$\varphi_{1,f}(t_k, \theta_{1,s}^{l+1}) = \begin{bmatrix} \frac{-p}{A_1(p, \theta_{1,s}^{l+1})} \tilde{y}_1(t_k) & \dots & \frac{-p^{n_1}}{A_1(p, \theta_{1,s}^{l+1})} \tilde{y}_1(t_k) \\ \frac{1}{p^\ell A_1(p, \theta_{1,s}^{l+1})} u(t_k) & \dots & \frac{p^{m_1}}{p^\ell A_1(p, \theta_{1,s}^{l+1})} u(t_k) \end{bmatrix}^T. \quad (35)$$

By computing the iterations in (25) with the first instrument and regressor vectors given by (34) and (35) respectively, a direct extension of the proposed method for identifying marginally stable systems in additive form is obtained. Aside from these modifications to the first submodel, the algorithm remains unchanged.

**Remark 4.8.** Assumption 1 is exclusively employed in the proof for generic consistency with uniformly stable systems. This excludes marginally stable systems, yet it does not constrain the proposed approach.

#### 4.7. Practical aspects for implementation

Next, several practical aspects of the additive RSRIVC implementation are investigated. First, the initialization process of the algorithm is discussed, followed by a discussion of the robustness of the algorithm. Subsequently, to alleviate the computational load associated with the approach, output subsampling is examined.

A proper initialization of the algorithm is important, particularly the initial estimate  $\bar{\theta}_i(t_0)$ , which should be selected carefully, as it influences the initial values of  $\varphi_{i,f}$ ,  $\hat{\varphi}_{i,f}$ , and  $\tilde{y}_{i,f}$  at  $t_0$ . Inadequate initialization can lead to divergence. The parameters  $\bar{\theta}_i(t_0)$  and covariance  $\bar{P}_i(t_0)$  can be obtained through any offline, non-recursive method using the input-output data  $\{u(t_k), y(t_k)\}_{k=N_{\text{init}}-1}^0$ . For example, approaches such as those described in González et al. (2023) or (González, Classens et al., 2025) may be utilized.

In the offline SRIVC method, ensuring model stability is vital, especially in challenging conditions like low Signal-to-Noise Ratio (SNR) or mismatched model orders (Ha & Welsh, 2014). Typically, stability



is monitored and enforced throughout iterations (Garnier, 2015). In RSRIVC, where both prefilter and auxiliary model depend on preceding estimates, stability verification is more intricate. For slowly-varying systems, checking characteristic equation poles and reflecting right-half plane poles to the left suffices for confirming bounded-input, bounded-output stability under specific conditions (Padilla et al., 2019). Notably, in time-varying parameter models, uniform exponential stability is assured by assuming slow parameter changes and keeping system matrix eigenvalues on the left half plane. RSRIVC incorporates a projection algorithm to guarantee the stability of the prefilter and auxiliary model, reflecting unstable poles into the left half plane while preserving the magnitude characteristics. This approach addresses potential stability issues that may arise, particularly at the beginning of the data record.

The algorithm presented is theoretically sound, however, the recursion in (32) may not be numerically reliable (Ljung & Söderström, 1983, Chapter 11). Since  $P(t_k)$  is essentially computed by successive subtractions, round-off errors can accumulate and make  $P(t_k)$  ill-conditioned. Other continuous-time recursive methods based on instrumental variables (Padilla et al., 2017; Pan et al., 2021) incorporate a covariance matrix  $Q(t_k)$  that is added to  $P(t_k)$  at each iteration. The proposed algorithm does not include this matrix since it is not compatible with the coordinate descent interpretation of the iterations in Algorithm 2. However, from a practical perspective, including this matrix can help with the convergence of the parameters. To achieve smoother estimates, a constant prefilter can replace the adaptive filter in the proposed approach, though this may prevent SRIVC iterations from converging to a local minimum, affecting the global convergence of the block coordinate descent.

Subsampling data may be valuable to managing computational complexity without compromising significant accuracy. This facilitates the implementation of recursive algorithms in resource-constrained environments. Subsampling should be performed cautiously to ensure that the subsampled data still captures the smallest timescale for a reliable estimate. Therefore, it is primarily applicable when the timescale of the system dynamics is considerably shorter than the sampling time. When subsampling the output, it is advised to take all input samples into account for the estimation of the filtered input and filtered auxiliary variables. Hence, the block-coordinate descent algorithm and additive RSRIVC iterations may be performed at this lower sampling rate. Furthermore, as hinted in Remark 4.6, it is not always necessary to perform multiple iterations. In case the parameters of the system vary slowly with respect to the sampling period, typically no refinements are necessary and a single  $\mathcal{A}$ -iteration can achieve accurate estimates.

## 5. Numerical simulation

In this section, the proposed method is illustrated using the following additive system

$$G^*(p, t) = \sum_{i=1}^3 \frac{b_{i,0}^*(t)}{a_{i,2}^*(t)p^2 + a_{i,1}^*(t)p + 1},$$

with initial DC-gains  $\begin{bmatrix} b_{1,0}^*(0) & b_{2,0}^*(0) & b_{3,0}^*(0) \end{bmatrix} = \begin{bmatrix} 3 & 0.5 & 1 \end{bmatrix}$ , and associated denominator parameters  $\begin{bmatrix} a_{1,2}^*(0) & a_{2,2}^*(0) & a_{3,2}^*(0) \end{bmatrix} = \begin{bmatrix} 0.25 & 0.04 & 0.0025 \end{bmatrix}$ , and  $\begin{bmatrix} a_{1,1}^*(0) & a_{2,1}^*(0) & a_{3,1}^*(0) \end{bmatrix} = \begin{bmatrix} 0.25 & 0.01 & 0.01 \end{bmatrix}$ . The output  $y(t_k)$  is measured at a sampling frequency of 20 [Hz] and is contaminated by a zero-mean Gaussian white noise with a variance of 0.01. The system is operated in closed-loop with feedback controller

$$K(q) = \frac{0.02329q^2 - 0.2058q + 0.00454}{q^2 - q},$$

and the setpoint equals  $r(t) = 3 \sin(0.005t) + \sin(2t) + \sin(5t) + \sin(17.5t)$ . The algorithm is initialized with covariance matrices  $\bar{P}_i(t_{N_1}) = \text{diag}(10^{-3}, 10^{-3}, 10^{-5})$ , and forgetting factor  $\lambda = 0.999$ . The estimator is initialized at a random system with parameters  $\hat{\theta}_i(t_{N_1}) = \theta_i^*(0)(1 +$

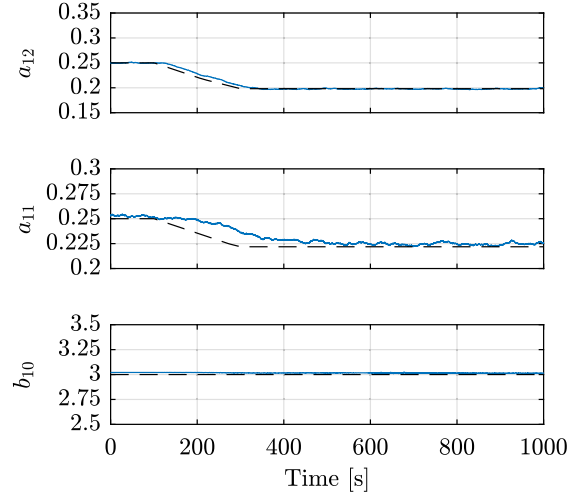


Fig. 4. Parameter estimates (—) and true parameters (---) related to the first submodel with decreasing denominator parameters.

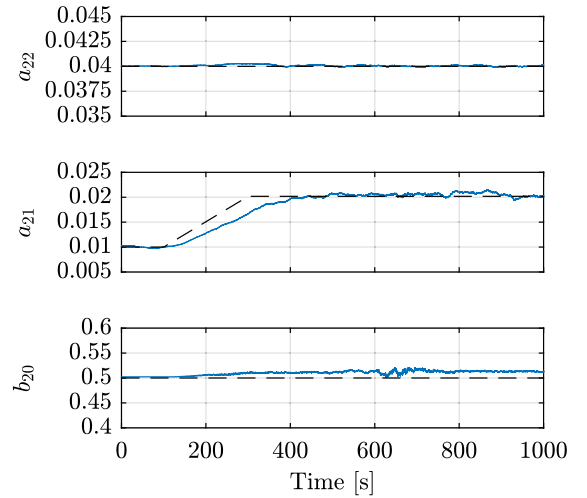


Fig. 5. Parameter estimates (—) and true parameters (---) related to the second submodel with decreasing damping coefficient  $\zeta_2$ , which corresponds to an increasing  $a_{21}$ .

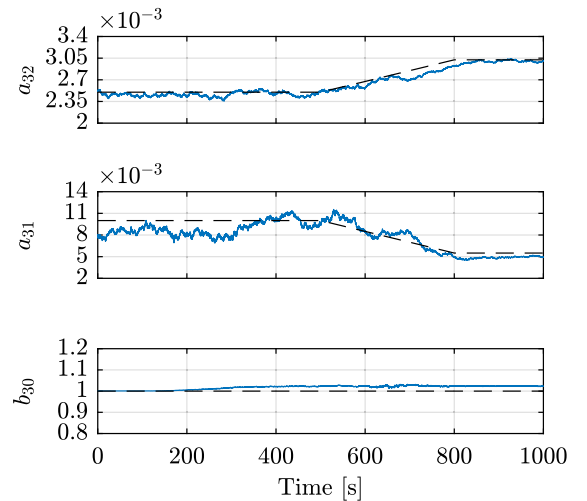


Fig. 6. Parameter estimates (—) and true parameters (---) related to the third submodel with decreasing natural frequency  $\omega_3$ , which corresponds to an increasing  $a_{32}$  and decreasing  $a_{31}$ .

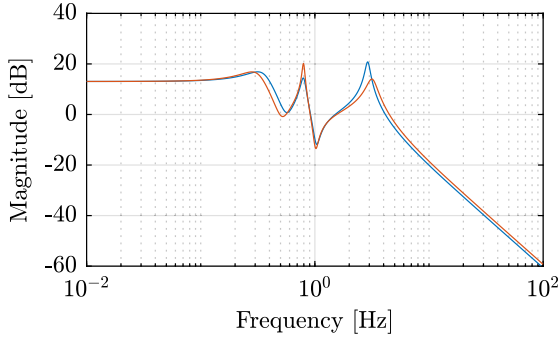


Fig. 7. Plant estimate at  $t = 50$  [s] (—) and at  $t = 950$  [s] (—). Clearly, all resonances changed. Note that the natural frequency of the first mode decreased and the damping of the second mode decreased. A combination of the two occurred at the third mode.

$\mathcal{U}(-0.02, 0.02)$ ), where  $\mathcal{U}(a, b)$  denotes a uniform distribution with lower limit  $a$  and upper limit  $b$ .

First, in the interval  $t \in [100, 300]$  [s], the parameters  $a_{1,1}$ ,  $a_{1,2}$ , and  $a_{2,1}$ , are varied in such a way that the corresponding natural frequency  $\omega_1$  and damping coefficient  $\zeta_2$  decrease, c.f., (16). Subsequently, in the interval  $t \in [500, 800]$  [s], the denominator parameters of the third model, i.e.,  $a_{3,2}$  and  $a_{3,1}$  are varied. Algorithm 2 is deployed to estimate the time-varying parameters, where no SRIVC iterations are performed and only a single coordinate descent iteration is performed at every incoming sample, i.e., the two loops in the algorithm are omitted as  $s = 0$  and  $M_l = 1$ . Since the system is operating in closed loop, the instrument is chosen as (33). The obtained estimates and the true parameters are depicted for each of the submodels in Figs. 4 to 6. The estimated model at  $t = 50$  [s] and  $t = 950$  [s] are depicted in Fig. 7. From the figures is concluded that no further refinements and gradient descent steps are required to achieve an accurate tracking of the parameters.

## 6. Experimental validation on an overactuated and oversensed flexible beam setup

In this section, the method is tested on an experimental setup. To this end, consider the setup shown in Fig. 8. The system consists of a thin flexible steel beam of  $500 \times 20 \times 2$  [mm]. It is equipped with five contactless fiberoptic sensors and three voice-coil actuators and is suspended by wire flexures, leaving one rotational and one translational direction unconstrained. The system is operating at a sampling frequency of 4096 [Hz]. The second and fourth sensor are not used for the conducted experiments.

The main aim is to estimate the SISO system  $G^*$  between the input  $u$ , exciting the system at the outer ends of the beam equally, see Figs. 8 and 9, and the output  $y$  which is the average deflection of the outer ends. This average deflection is controlled by a feedback controller and the setpoint is a square wave with a frequency of 2 [Hz]. By means of a second internal control loop, the stiffness of the beam is artificially manipulated such that the observed behavior is time-varying, see Classens et al. (2021, 2022) for details.

Lightly-damped systems such as the one under study are well-suited to be described in a modal representation, see Example 3.3, i.e.,

$$G^*(p) = \sum_{i=1}^K \frac{b_{i,0}^*}{p^2/\omega_i^2 + 2(\zeta_i/\omega_i)p + 1},$$

where  $\zeta_i$  and  $\omega_i$  represent the damping coefficients and natural frequencies of the flexible modes respectively, see Gawronski (2004). The flexible beam fits the modal description, and the system is approximated with  $K = 2$  modes. The first mode is the suspension mode of the system and the second mode is the first internal flexible mode which is artificially manipulated. Fig. 12 shows the model of the system in

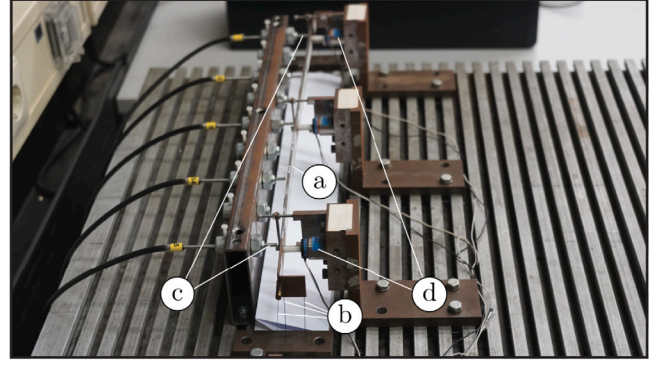


Fig. 8. Prototype experimental flexible beam setup. The moving part is indicated by (a) and is suspended by wire flexures (b). The deflection is measured with five contactless fiber optic sensors, of which two are used (c) and the setup is actuated with three current-driven voice coils of which the outer two are used (d).

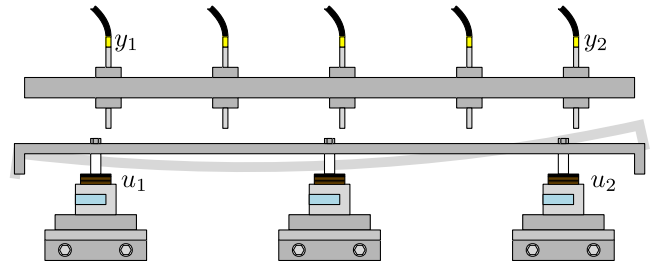


Fig. 9. Schematic illustration of the experimental setup. The three actuators are visible on the bottom part of the image and the five sensors at the top. During operation, the beam translates, rotates and exhibits internal flexible behavior. The inputs and outputs are transformed as  $u_1 = u_2 = \frac{1}{2}u$  and  $y = \frac{1}{2}(y_1 + y_2)$ , respectively.

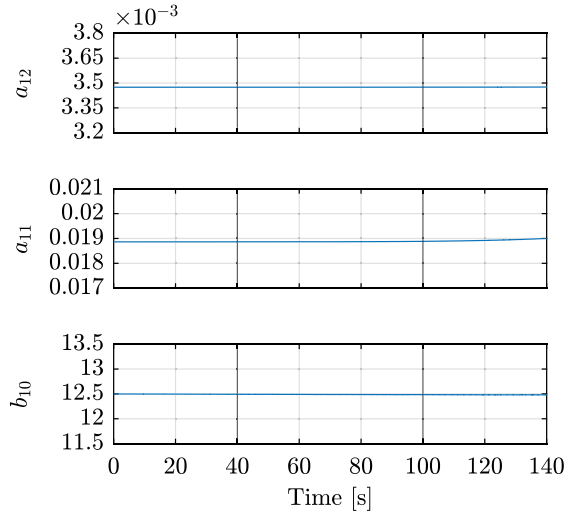


Fig. 10. Parameter estimate related to the first submodel. This first mode relates to the suspension of the flexible beam setup which remains unaltered throughout the experiment explaining the graphs.

the frequency domain, including the neglected higher order modes, and the model at  $t_0$  after initializing the system. Two distinct forgetting factors are used tailored to the distinct modes, namely  $\lambda_1 = 0.9999$  and  $\lambda_2 = 0.999$ , and the approximate covariance matrices are initialized at  $\bar{P}_1(t_0) = \bar{P}_2(t_0) = \text{diag}(10^{-11}, 10^{-6}, 10^{-1})$ .

An experiment is conducted where the internal flexible mode is actively manipulated from  $t \in [40, 100]$  [s]. Algorithm 2 is deployed

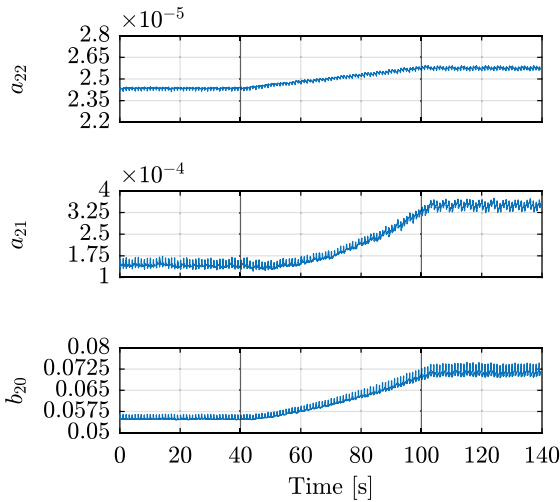


Fig. 11. Parameter estimate related to the second submodel. This mode is artificially manipulated between  $t = 40$  [s] and  $t = 100$  [s], which explains the changing parameters.

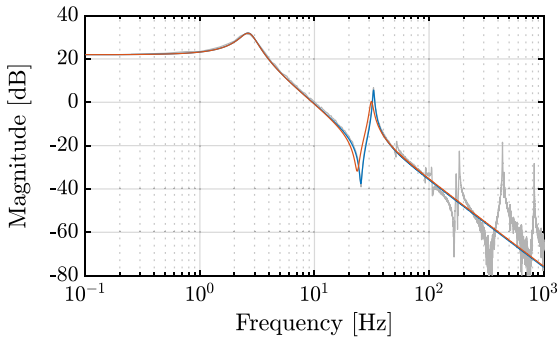


Fig. 12. Frequency response function of the nonmanipulated system (—), the plant estimate at  $t = t_{N_1} = 0$  [s] (—), and at the plant estimate at  $t = 135$  [s] (—).

to estimate the time-varying behavior, where no SRIVC iterations are performed and only a single coordinate descent iteration is performed at every incoming sample, i.e.,  $s = 0$  and  $M_l = 1$ . Since the system is operating in closed loop, the instrument is chosen as (33). The obtained estimates are depicted for each of the two submodels in Figs. 10 and 11, and the estimated model at  $t = 0$  [s] and  $t = 135$  [s] is depicted in Fig. 12. From the figures is concluded that the shifting resonance mode is effectively observed, and the first mode does not exhibit parameter variations. The small fluctuations in the estimate of the second mode are at the frequency of the setpoint.

## 7. Conclusions

This article addresses a recursive method to identify additive systems for both open and closed-loop setups. The proposed algorithms, based on a block-coordinate descent with refined instrumental variables, are capable to track time-varying parameters of continuous-time systems in real-time. This allows tracking of more parsimonious and physically-relevant model representations, particularly for mechanical systems. Practical aspects are provided and finally the effectiveness of the proposed approach is illustrated through numerical validation and application on a real-time setup.

## CRediT authorship contribution statement

**Koen Classens:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Rodrigo A. González:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Tom Oomen:** Writing – review & editing, Supervision, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This work is supported by ASML, Veldhoven, the Netherlands.

## Appendix A. Derivation of the refined instrumental variable iterations

In this appendix, it is shown that a stationary point of the following optimization problem is obtained from the refined instrumental variable method. Recall the optimization problem

$$\theta_i^{l+1}(t_N) = \arg \min_{\theta_i \in \Omega_i} \sum_{k=1}^N \alpha_{N-k} \left[ \tilde{y}_i(t_k) - G_i(p, t, \theta_i) u(t_k) \right]^2, \quad (\text{A.1})$$

with

$$\tilde{y}_i(t_k) := y(t_k) - \sum_{j=1}^{i-1} G_j(p, t, \theta_j^{l+1}) u(t_k) - \sum_{j=i+1}^K G_j(p, t, \theta_j^l) u(t_k).$$

and define

$$\varepsilon_i(t_k, \theta_i) = \tilde{y}_i(t_k) - G_i(p, t, \theta_i) u(t_k).$$

By taking Remark 4.7 into consideration, this residual is well approximated as

$$\varepsilon_i(t_k, \theta_i) = \frac{1}{A_i(p, \theta_i)} (A_i(p, \theta_i) \tilde{y}_i(t_k) - B_i(p, \theta_i) u(t_k)),$$

and subsequently  $A_i(p, \theta_i)$  and  $B_i(p, \theta_i)$  gives

$$\begin{aligned} \varepsilon_i(t_k, \theta_i) = \frac{1}{A_i(p, \theta_i)} & \left( a_{i,n_i}(t_k) p^{n_i} \tilde{y}_i(t_k) + \dots + a_{i,1}(t_k) p \tilde{y}_i(t_k) + \tilde{y}_i(t_k) \right. \\ & \left. - b_{i,m_i}(t_k) p^{m_i} u(t_k) - \dots - b_{i,1}(t_k) p u(t_k) - b_{i,0}(t_k) u(t_k) \right). \end{aligned}$$

The filtered residual output, filtered regressor vector, and parameter vector are defined as

$$\begin{aligned} \tilde{y}_{i,f}(t_k, \theta_i) &= \frac{1}{A_i(p, \theta_i)} \tilde{y}_i(t_k), \\ \varphi_{i,f}(t_k) &= \frac{1}{A_i(p, \theta_i)} \begin{bmatrix} -p \tilde{y}_i(t_k) & \dots & -p^{n_i} \tilde{y}_i(t_k) & u(t_k) & p u(t_k) & \dots & p^{m_i} u(t_k) \end{bmatrix}^\top, \\ \theta_i(t_k) &= \begin{bmatrix} a_{i,1}(t_k) & \dots & a_{i,n_i}(t_k) & b_{i,0}(t_k) & b_{i,1}(t_k) & \dots & b_{i,m_i}(t_k) \end{bmatrix}^\top, \end{aligned}$$

which allows to write  $\varepsilon_i$  as

$$\varepsilon_i(t_k, \theta_i) = \tilde{y}_{i,f}(t_k, \theta_i) - \varphi_{i,f}^\top(t_k, \theta_i) \theta_i(t_k), \quad (\text{A.2})$$

which is a pseudolinear regression problem in  $\theta_i$  that admits the refined instrumental variable (Garnier & Wang, 2008; Young, 2011) iterations. To this end, the parameter estimate of the previous SRIVC iterate  $s$  is taken in the denominator polynomial  $A_i$  which is used as prefilter for  $\tilde{y}_{i,f}$  and  $\varphi_{i,f}$ . Then, substitution in the first order optimality condition of (A.1), i.e.,

$$\sum_{k=1}^N \alpha_{N-k} \frac{\partial \varepsilon_i(t_k, \theta_i)}{\partial \theta_i} \varepsilon_i(t_k, \theta_i) = 0,$$

with the instrument vector  $\hat{\varphi}_{i,f}(t_k, \theta_i)$  gives

$$\sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_i) \left( \tilde{y}_{i,f}(t_k, \theta_i) - \varphi_{i,f}^\top(t_k, \theta_i) \theta_i(t_k) \right) = 0.$$

Hence, the iterations for a fixed block descent iteration  $l+1$  are of the form

$$\theta_{i,s+1}^{l+1}(t_N) = \left[ \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}) \right]^{-1} \times \left[ \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}) \right], \quad (\text{A.3})$$

with  $s \geq 1$ .

## Appendix B. Recursive computation of the SRIVC iterations

In order to reformulate (25) into its recursive form, define

$$R_{i,s}^{l+1}(t_N) = \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}), \quad (\text{B.1})$$

and

$$f_{i,s}^{l+1}(t_N) = \sum_{k=1}^N \alpha_{N-k} \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}). \quad (\text{B.2})$$

Using that  $\alpha_k = \lambda^k$  and that  $\alpha$  satisfies the multiplicative property  $\alpha_{k+1} = \lambda \alpha_k$ , with  $\alpha_0 = 1$ ,

$$R_{i,s}^{l+1}(t_k) = \lambda \bar{R}_i(t_{k-1}) + \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_k, \theta_{i,s}^{l+1}), \quad (\text{B.3})$$

and

$$f_{i,s}^{l+1}(t_k) = \lambda \bar{f}_i(t_{k-1}) + \hat{\varphi}_{i,f}(t_k, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_k, \theta_{i,s}^{l+1}), \quad (\text{B.4})$$

where  $\bar{R}_i(t_{k-1})$  and  $\bar{f}_i(t_{k-1})$  denote the final value of the iterations at the previous time step  $t_{k-1}$  (i.e., after iterating over the modes through  $i$  and  $l$ , and after the iterations in  $s$ ). Using that  $\theta_{i,s+1}^{l+1}(t_N) = (R_{i,s}^{l+1}(t_N))^{-1} f_{i,s}^{l+1}(t_N)$ , and  $\bar{f}_i(t_{k-1}) = \bar{R}_i(t_{k-1}) \bar{\theta}_i(t_{k-1})$ ,

$$\theta_{i,s+1}^{l+1}(t_N) = (R_{i,s}^{l+1}(t_N))^{-1} \left( \lambda \bar{R}_i(t_{N-1}) \bar{\theta}_i(t_{N-1}) + \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_N, \theta_{i,s}^{l+1}) \right),$$

which after substitution of  $\bar{R}_i(t_{k-1})$  from (B.3) is written as

$$\theta_{i,s+1}^{l+1}(t_N) = (R_{i,s}^{l+1}(t_N))^{-1} \left( \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \tilde{y}_{i,f}(t_N, \theta_{i,s}^{l+1}) + (R_{i,s}^{l+1}(t_N) - \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1})) \bar{\theta}_i(t_{N-1}) \right),$$

and thus

$$\theta_{i,s+1}^{l+1}(t_N) = \bar{\theta}_i(t_{N-1}) + (R_{i,s}^{l+1}(t_N))^{-1} \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \times \left( \tilde{y}_{i,f}(t_N, \theta_{i,s}^{l+1}) - \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{\theta}_i(t_{N-1}) \right). \quad (\text{B.5})$$

In order to avoid inverting  $R_{i,s}^{l+1}(t_N)$  at each time step, introduce  $P_{i,s}^{l+1}(t_N) = (R_{i,s}^{l+1}(t_N))^{-1}$ . The Sherman–Morrison formula (Horn & Johnson, 2012, Chap. 0.7.4), a generalization of the matrix inversion lemma, states that  $(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u}$ , provided the inverses exist, with  $A$  invertible and square, and column vectors  $u$  and  $v$ . Take  $A = \lambda \bar{R}_i(t_{N-1})$ ,  $u = \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})$ , and  $v = \varphi_{i,f}(t_N, \theta_{i,s}^{l+1})$ . Then,

$$\begin{aligned} (R_{i,s}^{l+1}(t_N))^{-1} &= \frac{1}{\lambda} (\bar{R}_i(t_{N-1}))^{-1} \\ &- \frac{\frac{1}{\lambda} (\bar{R}_i(t_{N-1}))^{-1} \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \frac{1}{\lambda} (\bar{R}_i(t_{N-1}))^{-1}}{1 + \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \frac{1}{\lambda} (\bar{R}_i(t_{N-1}))^{-1} \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})}, \end{aligned}$$

and thus with  $\bar{P}_i(t_{N-1}) = (\bar{R}_i(t_{N-1}))^{-1}$ ,

$$\begin{aligned} P_{i,s}^{l+1}(t_N) &= \frac{1}{\lambda} \bar{P}_i(t_{N-1}) \\ &- \frac{\frac{1}{\lambda} \bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1})}{\lambda + \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})}. \end{aligned} \quad (\text{B.6})$$

Define  $K_{i,s+1}^{l+1}(t_N) = P_{i,s}^{l+1}(t_N) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1})$ , or equivalently,

$$\begin{aligned} K_{i,s+1}^{l+1}(t_N) &= \bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \\ &\times \left( \lambda + \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1}) \hat{\varphi}_{i,f}(t_N, \theta_{i,s}^{l+1}) \right)^{-1}. \end{aligned} \quad (\text{B.7})$$

This expression allows us to write (B.6) as

$$P_{i,s}^{l+1}(t_N) = \frac{1}{\lambda} \left( \bar{P}_i(t_{N-1}) - K_{i,s+1}^{l+1}(t_N) \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{P}_i(t_{N-1}) \right),$$

and (B.5) as

$$\theta_{i,s+1}^{l+1}(t_N) = \bar{\theta}_i(t_{N-1}) + K_{i,s+1}^{l+1}(t_N) \left( \tilde{y}_{i,f}(t_N, \theta_{i,s}^{l+1}) - \varphi_{i,f}^\top(t_N, \theta_{i,s}^{l+1}) \bar{\theta}_i(t_{N-1}) \right).$$

Hence, (B.7), (B.6), and (B.5), form a recursive equivalent of (A.3).

## Appendix C. Generic consistency proof with $\alpha_k = 1$ for all $k$

In this appendix, the proof of Theorem 2 is provided.

**Proof.** As the number of SRIVC iterations  $s$  tends to infinity,

$$\begin{aligned} \theta_{i,\infty}^{l+1} &= \frac{1}{N} \left[ \sum_{k=1}^N \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) \varphi_{i,f}^\top(t_k, \theta_{i,\infty}^{l+1}) \right]^{-1} \\ &\times \frac{1}{N} \left[ \sum_{k=1}^N \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) \tilde{y}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) \right]. \end{aligned} \quad (\text{C.1})$$

That is,

$$\frac{1}{N} \sum_{k=1}^N \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) \left[ \tilde{y}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) - \varphi_{i,f}^\top(t_k, \theta_{i,\infty}^{l+1}) \theta_{i,\infty}^{l+1} \right] = 0, \quad (\text{C.2})$$

where

$$\begin{aligned} \tilde{y}_{i,f}(t_k, \theta_{i,\infty}^{l+1}) - \varphi_{i,f}^\top(t_k, \theta_{i,\infty}^{l+1}) \theta_{i,\infty}^{l+1} &= \frac{1}{A_i(t_k, \theta_{i,\infty}^{l+1})} \tilde{y}_i(t_k) + \frac{p a_{i,\infty,1}^{l+1} + \dots + p^n a_{i,\infty,m_i}^{l+1}}{A_i(t_k, \theta_{i,\infty}^{l+1})} \tilde{y}_i(t_k) \\ &- \frac{b_{i,\infty,0}^{l+1} + p b_{i,\infty,1}^{l+1} + \dots + p^{m_i} b_{i,\infty,m_i}^{l+1}}{A_i(t_k, \theta_{i,\infty}^{l+1})} u(t_k) \end{aligned} \quad (\text{C.3})$$

$$= \tilde{y}_i(t_k) - \frac{B_i(t_k, \theta_{i,\infty}^{l+1})}{A_i(t_k, \theta_{i,\infty}^{l+1})} u(t_k) \quad (\text{C.4})$$

$$\begin{aligned} &= y(t_k) - \sum_{j=1}^{i-1} G_j(p, t_k, \theta_j^{l+1}) u(t_k) - \sum_{j=i+1}^K G_j(p, t_k, \theta_j^l) u(t_k) \\ &- \frac{B_i(t_k, \theta_{i,\infty}^{l+1})}{A_i(t_k, \theta_{i,\infty}^{l+1})} u(t_k). \end{aligned} \quad (\text{C.5})$$

To simplify the notation, in the subsequent analysis, the  $t_k$  dependency on the transfer functions is omitted. Denoting the parameter vector  $\theta_{i,\infty}^l$  as the limit of  $\theta_{i,\infty}^{l+1}$  when the number of descent iterations tends to infinity, then, as  $l \rightarrow \infty$ ,

$$\tilde{y}_{i,f}(t_k, \theta_{i,\infty}^\infty) - \varphi_{i,f}^\top(t_k, \theta_{i,\infty}^\infty) \theta_{i,\infty}^\infty = y(t_k) - \sum_{j=1}^K G_j(p, \theta_{j,\infty}^\infty) u(t_k). \quad (\text{C.6})$$

On the other hand, note that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^\infty) v(t_k) = \mathbb{E} \left\{ \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^\infty) v(t_k) \right\} = 0, \quad (\text{C.7})$$

where the first equality is due to the ergodicity property of the input and noise signals (see Theorem 2B.1 of Butcher et al. (2008) for more details), and the second equality holds since  $v$  and  $\hat{\varphi}_{i,f}$  are uncorrelated due to Assumption 4. So, for  $i = 1, \dots, K$ , combining (C.2), (C.6), (13) and (C.7) leads to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^K \sum_{k=1}^N \hat{\varphi}_{i,f}(t_k, \theta_{i,\infty}^\infty) \left[ G_j^*(p) - G_j(p, \theta_{j,\infty}^\infty) \right] u(t_k) = 0. \quad (\text{C.8})$$

Now,

$$G_j^*(p) - G_j(p, \theta_{j,\infty}^\infty) = \frac{B_j^*(p)}{A_j^*(p)} - \frac{B_j(p, \theta_{j,\infty}^\infty)}{A_j(p, \theta_{j,\infty}^\infty)}$$

$$\begin{aligned}
&= \frac{B_j^*(p)A_j(p, \theta_{j,\infty}^\infty) - A_j^*(p)B_j(p, \theta_{j,\infty}^\infty)}{A_j^*(p)A_j(p, \theta_{j,\infty}^\infty)} \\
&= \frac{[1, \quad p, \quad \dots, \quad p^{n_i+m_j}]}{A_j^*(p)A_j(p, \theta_{j,\infty}^\infty)} h_j,
\end{aligned} \quad (\text{C.9})$$

where  $h_j \in \mathbb{R}^{n_i+m_j+1}$  is the vector that contains the coefficients of  $B_j^*(p)A_j(p, \theta_{j,\infty}^\infty) - A_j^*(p)B_j(p, \theta_{j,\infty}^\infty)$  in increasing order of degree. With regards to the instrument vector, the following decomposition holds (see, e.g., Eq. (16) of Pan, González et al., 2020)

$$\hat{\varphi}_{i,t}(t_k, \theta_{i,\infty}^\infty) = S(-B_i, A_i) \frac{1}{A_i^2(p, \theta_{i,\infty}^\infty)} U_i(t_k), \quad (\text{C.10})$$

where  $U_i(t_k) = [1 \quad p \quad \dots \quad p^{n_i+m_i}]^\top u(t_k)$ , and where  $S(-B_i, A_i)$  is a Sylvester matrix formed by the coefficients of the polynomials  $-B_i(p, \theta_{i,\infty}^\infty)$  and  $A_i(p, \theta_{i,\infty}^\infty)$ , similar to Eq. (13) of Pan, González et al. (2020). Thus, (C.8) is equivalent to

$$\begin{aligned}
&\sum_{j=1}^K \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N S(-B_i, A_i) \frac{1}{A_i^2(t_k, \theta_{i,\infty}^\infty)} U_i(t_k) \\
&\quad \times \frac{1}{A_j^*(p)A_j(p, \theta_{j,\infty}^\infty)} U_j^\top(t_k) h_j = 0,
\end{aligned} \quad (\text{C.11})$$

for  $i = 1, \dots, K$ . After applying Lemma 1 in Appendix C.1 and the ergodicity result in Theorem 2B.1 of Butcher et al. (2008), which hold since the system and model denominators describe uniformly stable transfer functions and the input is assumed quasi-stationary, the following is obtained.

$$\begin{aligned}
S(-\bar{B}_i, \bar{A}_i) \sum_{j=1}^K \mathbb{E} \left\{ \frac{1}{A_i^2(t_k, \theta_{i,\infty}^\infty)} U_i(t_k) \right. \\
\left. \times \frac{1}{A_j^*(p)A_j(p, \theta_{j,\infty}^\infty)} U_j^\top(t_k) \right\} \bar{h}_j = 0,
\end{aligned} \quad (\text{C.12})$$

for  $i = 1, \dots, K$ , where  $(\bar{A}_i, \bar{B}_i) = \lim_{N \rightarrow \infty} (A_i(p, \theta_{i,\infty}^\infty), B_i(p, \theta_{i,\infty}^\infty))$ , and  $\bar{h}_j = \lim_{N \rightarrow \infty} h_j$ . The  $K$  conditions in (C.12) are jointly expressed as  $S\Phi_u \eta = 0$ , where  $S$  is a block-diagonal matrix whose block-diagonal is comprised by  $\{S(-\bar{B}_i, \bar{A}_i)\}_{i=1}^K$ , while  $\eta = [\bar{h}_1^\top, \dots, \bar{h}_K^\top]^\top$ , and  $\Phi_u$  has block entries given by

$$\Phi_u^{i,j} = \mathbb{E} \left\{ \frac{1}{A_i^2(t_k, \theta_{i,\infty}^\infty)} U_i(t_k) \frac{1}{A_j^*(p)A_j(p, \theta_{j,\infty}^\infty)} U_j(t_k) \right\}.$$

**Assumption 3** regarding the coprimeness of  $\bar{A}_i$  and  $\bar{B}_i$  implies that  $S$  is nonsingular, as shown in Lemma A3.1 of Söderström and Stoica (1983). Moreover, Lemma 2 in Appendix C.1 applied over each entry of  $\Phi_u$  indicates that the non-singularity of  $\Phi_u$  can be studied directly from the cross covariance of the inputs filtered by the LTI systems obtained at the limit as  $N \rightarrow \infty$ . Thus, following the same reasoning as in Lemma 3 of González, Classens et al. (2025), Assumptions 1 to 4 imply the generic nonsingularity of  $\Phi_u$  with respect to the converging system and model denominator parameters. Consequently,  $\eta = 0$ , i.e.,  $\lim_{N \rightarrow \infty} h_j = 0$  for all  $i = 1, \dots, K$ , as desired.  $\square$

### C.1. Technical lemmas

**Lemma 1.** *If the matrix  $\mathcal{O}_k$  has a bounded norm for all  $k$  with probability 1, then, with probability 1,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N S_k \mathcal{O}_k h_k = \bar{S} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathcal{O}_k \bar{h}, \quad (\text{C.13})$$

where  $S_k$  and  $h_k$  are matrices and vectors of appropriate dimensions, such that  $\lim_{k \rightarrow \infty} S_k = \bar{S}$  and  $\lim_{k \rightarrow \infty} h_k = \bar{h}$ .

**Proof.** Note that the following inequalities hold:

$$\frac{1}{N} \left\| \sum_{k=1}^N (S_k - \bar{S}) \mathcal{O}_k h_k \right\| \leq \frac{1}{N} \sum_{k=1}^N \|S_k - \bar{S}\| \|\mathcal{O}_k h_k\|$$

$$\leq \frac{M_1}{N} \sum_{k=1}^N \|S_k - \bar{S}\|, \quad (\text{C.14})$$

where  $M_1 = \sup_k \|\mathcal{O}_k h_k\| < \infty$  with probability 1, since  $\mathcal{O}_k$  and  $h_k$  are bounded. Since  $\|S_k - \bar{S}\| \rightarrow 0$  as  $k \rightarrow \infty$ , then the Cesàro mean in (C.14) also goes to zero due to Theorem 4.2.3 of Cover and Thomas (2006). By the same reasoning, it follows that

$$\begin{aligned}
\frac{1}{N} \left\| \sum_{k=1}^N \bar{S} \mathcal{O}_k (h_k - \bar{h}) \right\| &\leq \frac{1}{N} \sum_{k=1}^N \|\bar{S} \mathcal{O}_k\| \|h_k - \bar{h}\| \\
&\leq \frac{M_2}{N} \sum_{k=1}^N \|h_k - \bar{h}\| \xrightarrow{N \rightarrow \infty} 0.
\end{aligned}$$

Thus, the result in (C.13) follows from decomposing  $S_k \mathcal{O}_k h_k$  as

$$S_k \mathcal{O}_k h_k = (S_k - \bar{S}) \mathcal{O}_k h_k + \bar{S} \mathcal{O}_k (h_k - \bar{h}) + \bar{S} \mathcal{O}_k \bar{h},$$

summing over  $k$ , dividing over  $N$ , and computing the limit as  $N \rightarrow \infty$ .  $\square$

**Lemma 2.** *Let*

$$w(t) = \sum_{k=0}^{\infty} \alpha_t(k) e(t-k), \quad v(t) = \sum_{k=0}^{\infty} \beta_t(k) e(t-k), \quad (\text{C.15})$$

where the families of filters  $\alpha_t(k), \beta_t(k)$ ,  $t = 1, 2, \dots$  are uniformly stable, and  $\{e(t)\}$  is a quasi-stationary signal. In addition, let  $\bar{w}(t)$  and  $\bar{v}(t)$  be defined as in (C.15) but for the LTI filters  $\bar{\alpha}(k)$  and  $\bar{\beta}(k)$  respectively, where  $\alpha_t(k) \rightarrow \bar{\alpha}(k)$  and  $\beta_t(k) \rightarrow \bar{\beta}(k)$  for every  $k$  as  $t \rightarrow \infty$ . Then,

$$|\mathbb{E}\{w(t)v(t)\} - \mathbb{E}\{\bar{w}(t)\bar{v}(t)\}| \rightarrow 0 \quad \text{as } N \rightarrow \infty. \quad (\text{C.16})$$

**Proof.** The following inequalities hold:

$$|\mathbb{E}\{w(t)v(t)\} - \mathbb{E}\{\bar{w}(t)\bar{v}(t)\}| = \left| \frac{1}{N} \sum_{t=1}^N \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \mathbb{E}\{e(t-k)e(t-\ell)\} (\alpha_t(k)\beta_t(\ell) - \bar{\alpha}(k)\bar{\beta}(\ell)) \right| \quad (\text{C.17})$$

$$\leq \frac{C}{N} \sum_{t=1}^N \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \left| (\alpha_t(k)(\beta_t(\ell) - \bar{\beta}(\ell)) + \bar{\beta}(\ell)(\alpha_t(k) - \bar{\alpha}(k))) \right| \quad (\text{C.18})$$

$$\begin{aligned}
&\leq \frac{C}{N} \sup_{t \in \mathbb{N}} \sum_{k=0}^{\infty} |\alpha_t(k)| \sum_{\ell=0}^{\infty} |\beta_t(\ell) - \bar{\beta}(\ell)| \\
&\quad + \frac{C}{N} \sum_{\ell=0}^{\infty} |\bar{\beta}(\ell)| \sum_{t=1}^N \sum_{k=0}^{\infty} |\alpha_t(k) - \bar{\alpha}(k)|,
\end{aligned} \quad (\text{C.19})$$

where (C.17) holds by definition, (C.18) uses the quasi-stationary property  $|\mathbb{E}\{e(t-k)e(t-\ell)\}| \leq C$  for any indices  $t, k$  and  $\ell$ , and (C.19) is due to the triangle inequality. Since  $\alpha_t(k)$  and  $\beta_t(\ell)$  describe uniformly stable families of filters, the two series  $\sup_{t \in \mathbb{N}} \sum_{k=0}^{\infty} |\alpha_t(k)|$  and  $\sum_{k=0}^{\infty} |\bar{\beta}(k)|$  converge. Thus, to prove (C.16) we require to show that

$$\frac{1}{N} \sum_{t=1}^N \sum_{k=0}^{\infty} |\alpha_t(k) - \bar{\alpha}(k)| \rightarrow 0 \quad \text{as } N \rightarrow \infty, \quad (\text{C.20})$$

and similarly with  $\beta_t(k)$ . Due to the pointwise convergence,  $\lim_{t \rightarrow \infty} |\alpha_t(k) - \bar{\alpha}(k)| = 0$ , and due to the uniform stability of  $\alpha_t(k)$ ,

$$\sum_{k=0}^{\infty} |\alpha_t(k) - \bar{\alpha}(k)| \leq 2 \sum_{k=0}^{\infty} \sup_{t \in \mathbb{N}} |\alpha_t(k)| < \infty.$$

Thus, by the Weierstrass M-test (Billingsley, 1995, Theorem A28),

$$\lim_{t \rightarrow \infty} \sum_{k=0}^{\infty} |\alpha_t(k) - \bar{\alpha}(k)| = \sum_{k=0}^{\infty} \lim_{t \rightarrow \infty} |\alpha_t(k) - \bar{\alpha}(k)| = 0.$$

Hence, (C.20) can be interpreted as a Cesàro mean with sequence in  $t$  given by  $\sum_{k=0}^{\infty} |\alpha_t(k) - \bar{\alpha}(k)|$ , which goes to zero as  $t$  tends to infinity. Thus, by Theorem 4.2.3 of Cover and Thomas (2006), (C.20) holds. The same derivation holds for  $\beta_t(\ell)$ , which implies that the right hand side of (C.19) tends to zero as  $N$  tends to infinity, thus proving (C.16).  $\square$



## References

- Bai, E.-W. (2005). Identification of nonlinear additive FIR systems. *Automatica*, 41(7), 1247–1253.
- Bai, E.-W., & Chan, K.-S. (2008). Identification of an additive nonlinear system and its applications in generalized Hammerstein models. *Automatica*, 44(2), 430–436.
- Billingsley, P. (1995). *Probability and measure* (3rd ed.). John Wiley & Sons.
- Boeren, F., Blanken, L., Bruijnen, D., & Oomen, T. (2018). Optimal estimation of rational feedforward control via instrumental variables: With application to a wafer stage. *Asian Journal of Control*, 20(3), 975–992.
- Bruce, A. L., Goel, A., & Bernstein, D. S. (2020). Convergence and consistency of recursive least squares with variable-rate forgetting. *Automatica*, 119, Article 109052.
- Butcher, M., Karimi, A., & Longchamp, R. (2008). On the consistency of certain identification methods for linear parameter varying systems. *IFAC Proceedings Volumes*, 41(2), 4018–4023.
- Classens, K., van de Wijdeven, J., Heemels, W. P. M. H., & Oomen, T. (2023). Opportunities of digital twins for high-tech systems: From fault diagnosis and predictive maintenance to control reconfiguration. *Mikroniek*, 63(5), 5–12.
- Classens, K., Heemels, W. P. M. H., & Oomen, T. (2021). A closed-loop perspective on fault detection for precision motion control: With application to an overactuated system. In *2021 IEEE international conference on mechatronics (ICM)* (pp. 1–6). IEEE.
- Classens, K., Mostard, M., van de Wijdeven, J., Heemels, W. P. M. H., & Oomen, T. (2022). Fault detection for precision mechatronics: Online estimation of mechanical resonances. *IFAC-PapersOnLine*, 55(37), 746–751.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). John Wiley & Sons.
- Garnier, H. (2015). Direct continuous-time approaches to system identification. Overview and benefits for practical applications. *European Journal of Control*, 24, 50–62.
- Garnier, H., & Wang, L. (2008). *Identification of continuous-time models from sampled data*. Springer.
- Garnier, H., & Young, P. C. (2014). The advantages of directly identifying continuous-time transfer function models in practical applications. *International Journal of Control*, 87(7), 1319–1338.
- Gawronski, W. K. (2004). *Advanced structural dynamics and active control of structures*. Springer.
- Gilson, M., Garnier, H., Young, P. C., & Van den Hof, P. M. J. (2008). Instrumental variable methods for closed-loop continuous-time model identification. In H. Garnier, & L. Wang (Eds.), *Identification of continuous-time models from sampled data* (pp. 133–160). Springer.
- Gilson, M., Garnier, H., Young, P. C., & Van den Hof, P. M. J. (2011). Optimal instrumental variable method for closed-loop identification. *IET Control Theory & Applications*, 5(10), 1147–1154.
- González, R. A., Classens, K., Rojas, C. R., Welsh, J. S., & Oomen, T. (2024). Statistical analysis of block coordinate descent algorithms for linear continuous-time system identification. *IEEE Control Systems Letters*.
- González, R. A., Classens, K., Rojas, C. R., Welsh, J. S., & Oomen, T. (2025). Identification of additive continuous-time systems in open and closed loop. *Automatica*, 173, Article 112013.
- González, R. A., Pan, S., Rojas, C. R., & Welsh, J. S. (2025). Consistency analysis of refined instrumental variable methods for continuous-time system identification in closed-loop. *Automatica*, 166, Article 111697.
- González, R. A., Rojas, C. R., Pan, S., & Welsh, J. S. (2023). Parsimonious identification of continuous-time systems: A block-coordinate descent approach. In *15th IFAC world congress, yokohama, Japan*.
- Ha, H., & Welsh, J. S. (2014). Ensuring stability in continuous time system identification instrumental variable method for over-parameterized models. In *53rd IEEE conference on decision and control (CDC)* (pp. 2597–2602).
- Härdle, W., Huet, S., Mammen, E., & Sperlich, S. (2004). Bootstrap inference in semiparametric generalized additive models. *Econometric Theory*, 20(2), 265–300.
- Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, 1(3), 297–310.
- Horn, R. A., & Johnson, C. R. (2012). *Matrix analysis* (2nd ed.). Cambridge University Press.
- Ljung, L. (1999). *System identification: Theory for the user* (2nd ed.). Prentice-Hall.
- Ljung, L., & Söderström, T. (1983). *Theory and practice of recursive identification*. MIT Press.
- Luenberger, D. G., & Ye, Y. (2008). *Linear and nonlinear programming* (3rd ed.). Springer.
- Mooren, N., Witvoet, G., & Oomen, T. (2023). On-line instrumental variable-based feedforward tuning for non-resetting motion tasks. *International Journal of Robust and Nonlinear Control*, 1–19.
- Oomen, T. (2018). Advanced motion control for precision mechatronics: Control, identification, and learning of complex systems. *IEEE Journal of Industry Applications*, 7(2), 127–140.
- Opsomer, J. D., & Ruppert, D. (1999). A root-n consistent backfitting estimator for semiparametric additive modeling. *Journal of Computational and Graphical Statistics*, 8(4), 715–732.
- Padilla, A., Garnier, H., Young, P. C., Chen, F., & Yuz, J. (2019). Identification of continuous-time models with slowly time-varying parameters. *Control Engineering Practice*, 93, Article 104165.
- Padilla, A., Garnier, H., Young, P. C., & Yuz, J. (2017). Recursive online IV method for identification of continuous-time slowly time-varying models in closed loop. *IFAC-PapersOnLine*, 50(1), 4008–4013.
- Pan, S., González, R. A., Welsh, J. S., & Rojas, C. R. (2020). Consistency analysis of the simplified refined instrumental variable method for continuous-time systems. *Automatica*, 113.
- Pan, S., Welsh, J. S., & Fu, M. (2021). Identification of continuous-time linear time-varying systems with abrupt changes in parameters. *IFAC-PapersOnLine*, 54(7), 339–344.
- Pan, S., Welsh, J. S., González, R. A., & Rojas, C. R. (2020). Efficiency analysis of the simplified refined instrumental variable method for continuous-time systems. *Automatica*, 121, Article 109196.
- Rugh, W. J. (1996). *Linear system theory*. Prentice-Hall.
- Söderström, T., & Stoica, P. (1983). *Instrumental variable methods for system identification*. Berlin: Springer-Verlag.
- Söderström, T., & Stoica, P. (1989). *System identification*. Prentice-Hall.
- Tóth, R. (2010). *Modeling and identification of linear parameter-varying systems*. Springer.
- Van den Hof, P. M. J. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, 22, 173–186.
- Voorhoeve, R., Aangenent, W., & Oomen, T. (2020). Identifying position-dependent mechanical systems: A modal approach applied to a flexible wafer stage. *IEEE Transactions on Control Systems Technology*, 22(1), 194–206.
- Young, P. C. (2011). *Recursive estimation and time-series analysis: An introduction for the student and practitioner*. Springer.
- Young, P. C. (2015). Refined instrumental variable estimation: Maximum Likelihood optimization of a unified Box–Jenkins model. *Automatica*, 52, 35–46.
- Young, P. C., & Jakeman, A. J. (1980). Refined instrumental variable methods of recursive time-series analysis part III. Extensions. *International Journal of Control*, 31(4), 741–764.