Title

Multivariable control based on incomplete models via feedback linearization and continuous-time derivative estimation

Journal
International Journal of Robust and Nonlinear Control, 31(18)

ISSN
1049-8923

Authors
Rodrigues, Diogo
Mesbah, Ali

Publication Date
2021-12-01

DOI
10.1002/rnc.5762

Data Availability
The data associated with this publication are available at:
https://github.com/dfmrodrigues/SNSFproject- P2ELP2_184521

Peer reviewed

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Multivariable Control Based on Incomplete Models via Feedback Linearization and Continuous-Time Derivative Estimation

Diogo Rodrigues* | Ali Mesbah

Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

Correspondence
*Diogo Rodrigues
Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA
Email: d.rodrigues@berkeley.edu

Present Address
*Diogo Rodrigues
Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
Email: dfmr@tecnico.ulisboa.pt

Summary

Models of nonlinear dynamical systems are typically composed of unknown and known parts due to the existence of mismatch between the true nonlinear system dynamics and their models. This paper presents a multivariable control strategy for nonlinear systems that deals explicitly with the existence of incomplete models in a data-driven framework. The proposed strategy uses feedback linearization and continuous-time estimation of the unknown derivatives to achieve satisfactory closed-loop performance for fast sampling applications with simple tuning of design parameters in the absence of a fully known system model. It is shown that this strategy possesses useful steady-state properties, such as elimination of steady-state error and rejection of constant input disturbances without any integral term, and ensures exponential convergence to the setpoints. Another contribution of this paper is the development of methods for quantitative analysis of the stability and performance of the control strategy in the continuous-time case. The approach is illustrated via a simulated example of reactor control. The proposed control strategy is broadly applicable and simplifies control design significantly for a large variety of systems.

KEYWORDS:
Multivariable control, Feedback linearization, Derivative estimation, Nonlinear systems, Data-driven control

1 | INTRODUCTION

System identification and model-based control design are often seen as closely related tasks, since control laws are typically designed based on the knowledge of a system model[1]. However, in the context of nonlinear dynamical systems, the dynamic models that are used to predict the evolution of the system state for control are typically composed of parts that are well known and other ones that are unknown[2]. This situation occurs mostly due to the existence of structural mismatch between the true nonlinear system and its model in the sense that the true system may not be in the model set[3,4,5]. This can be the case not only for first-principles models but also for data-driven models[6]. Thus, although it is possible to use system identification or first-principles modeling to obtain a relatively accurate description of the true dynamical system, this description remains subject to some uncertainty, and the experimental effort required to reduce the model uncertainty may be time-consuming and expensive.

In addition, typical model-based control strategies based on linearization of nonlinear models at an equilibrium point cannot achieve globally exact linearization and may lead to unsatisfactory performance and robustness if the system is operated far from the point at which the model was linearized. For this reason, control methods for nonlinear systems based on exact linearization have been developed, namely, input-output feedback linearization and input-state feedback linearization[7,8,9,10].

These approaches via feedback linearization typically require an adequate model of the nonlinear system, otherwise the property of exact linearization is lost. However, in practice, significant modeling errors and disturbances may be present[11]. Hence, several approaches have been presented to increase the robustness of feedback-linearizing control with respect to errors and disturbances via adaptive control[12,13] and robust control[14,15,16,17], or to analyze the impact of the modeling errors on the stability and performance of the control strategy via feedback linearization[18]. However, the guarantees provided by these methods typically require some assumptions that the modeling errors and disturbances must satisfy, related to parametric uncertainty, conditions of similarity with respect to the nominal model (also known as matching conditions), relative degree, or bounded magnitude, which complicate their practical application. Since these earlier works on adaptive and robust feedback linearization, the work on control of nonlinear systems has mainly focused on model predictive control (MPC), which has found wide application in industrial contexts[19,20]. However, while nonlinear MPC is necessary to control highly nonlinear systems, it may be too computationally expensive for fast sampling applications, despite recent developments in this field[21], and it generally requires good models that may be difficult to obtain. On the other hand, although the ubiquitous proportional-integral-derivative (PID) control strategy is conceptually simple and computationally efficient for fast sampling applications, it remains challenging to design and tune adequate multi-loop PID control systems for multivariable and nonlinear systems. Hence, control approaches based on feedback linearization remain useful as computationally efficient advanced control strategies for uncertain, multivariable, and nonlinear systems. The applications of feedback linearization range from magnetic levitation[22] and power systems such as micro grids[23,24] to chemical and biological reactors[11] and cancer immunotherapy[25], only to name a few.

Nonetheless, it still remains an open problem how to identify models of nonlinear systems that yield satisfactory performance and robustness for feedback linearization control. However, since input-output feedback linearization relies mostly on the knowledge of values that can be related to the output derivatives, one could infer these derivatives directly from measurements via signal differentiation, without using a dynamic model. Some approaches for numerical differentiation for nonlinear systems, such as sliding mode[26] and Tikhonov regularization[27,28], can be highlighted in this context. More recently, efficient algebraic approaches for numerical differentiation via finite impulse response (FIR) filters have also been proposed, which result in an interesting deadbeat property of the derivative estimates[29,30]. This would be an alternative to the use of observers[31,32] or unknown input estimators[33] for measurement-based derivative estimation, which may be difficult to design so as to minimize the effect of measurement noise. However, possibly due to the scarcity of work on adaptive and robust control via feedback linearization in recent years, the use of derivative estimation via FIR filters for adaptive and robust control via feedback linearization has not been explored, to the best of our knowledge, despite the improved handling of measurement noise by these FIR filters.

The primary objective of this paper is to propose a multivariable control strategy for nonlinear systems that deals explicitly and systematically with the fact that models may be incomplete in the sense that there may exist parametric model uncertainty or even structural mismatch between the true nonlinear system and its model. The proposed control strategy is intuitive to understand and simple to tune. In addition, it provides elimination of steady-state error and rejection of constant input disturbances without any integral term and ensures exponential convergence to the setpoints. We use the concept of continuous-time derivative estimation to estimate the derivatives of the unknown system dynamics without model identification. This concept has already been developed and applied to chemical reactor control[34,35,36]. The control design uses an approach via input-output feedback linearization that was proposed originally for the case of fully known models[37], as it is typically the case whenever feedback linearization is used. The requirement of full knowledge of the dynamic model is avoided by the proposed control strategy since it is based on derivative estimation. As such, the proposed control strategy is similar to observer-based control, in which the unknown derivatives are estimated via high-gain observers. It has been shown that observer-based control recovers the performance of control via feedback linearization in the absence of model uncertainty and disturbances[32]. In contrast, in this work, the unknown derivatives are estimated via FIR filters, which provides advantages such as deadbeat estimation, instead of asymptotic convergence, and minimization of the effect of measurement noise. The quantitative analysis of the local stability and performance of this approach via feedback linearization and derivative estimation represents another contribution of this paper since this analysis has only been performed for the particular case of relative degree one of the controlled outputs with respect to the inputs and derivatives up to first order[38]. To the best of our knowledge, although continuous-time estimation of derivatives up to arbitrary order via FIR filters has been proposed before, this paper proposes and analyzes for the first time the use of continuous-time estimation of derivatives up to arbitrary order via FIR filters for control of nonlinear systems with incompletely known dynamics via input-output feedback linearization. In summary, the main contributions of the paper include:

1. Description of the multivariable control strategy that uses input-output feedback linearization and continuous-time estimation of derivatives up to arbitrary order via FIR filters for control of nonlinear systems with incompletely known dynamics and arbitrary relative degree of the controlled outputs with respect to the inputs.

2. Proof of the property of minimization of measurement noise provided by continuous-time derivative estimation via FIR filters, as well as several auxiliary results and generalizations related to these FIR filters that are necessary before these filters can be applied to multivariable control.

3. Proof of the properties of elimination of steady-state error and rejection of constant input disturbances without any integral term, and exponential convergence of the controlled outputs to the setpoints provided by the proposed control strategy.

4. Quantitative analysis of the local stability and performance of the proposed control strategy for a given system.

The paper is organized as follows. Section 2 describes the system under study and the control problem. Section 3 presents the control strategy via feedback linearization and derivative estimation. Section 4 shows the steady-state and convergence properties of the proposed control strategy, presents methods to assess its stability and performance, and addresses practical aspects related to the control strategy. Section 5 provides an illustrative example of reactor control without the use of kinetic models, followed by Section 6 that concludes the paper.

2 | PROBLEM DESCRIPTION

Consider a nonlinear dynamical system with n_x states $\mathbf{x}(t)$, n_u inputs $\mathbf{u}(t)$, and n_y outputs $\mathbf{y}(t)$. The state dynamics are composed of an unknown part and a known or available part, denoted by the subscripts u and a, while the outputs are known functions of the states. The true dynamical system is

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{1a}$$

$$\mathbf{y}(t) = \mathbf{g}\big(\mathbf{x}(t)\big), \tag{1b}$$

where $\mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ and $\mathbf{g}\big(\mathbf{x}(t)\big)$ are nonlinear functions of appropriate dimensions.

Suppose that the objective is to control $\mathbf{y}_c(t) := \mathbf{S}\mathbf{y}(t)$, the n_c linear combinations of the outputs $\mathbf{y}(t)$ given by the $n_c \times n_y$ matrix $\mathbf{S}$ with rank $(\mathbf{S}) = n_c$, which implies that $n_c \leq n_y$. Also, suppose that $\mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ and $\mathbf{g}\big(\mathbf{x}(t)\big)$ are $n+1$ times differentiable with respect to $\mathbf{x}(t)$ and the relative degree of the outputs $\mathbf{y}_c(t)$ with respect to $\mathbf{u}(t)$ is n. Consequently, the pth time derivatives of $\mathbf{y}_c(t)$ denoted as $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, are absolutely continuous functions of t that depend only on the absolutely continuous states $\mathbf{x}(t)$, which implies that the time derivatives $\dot{\mathbf{y}}_c^{(p)}(t) = \mathbf{y}_c^{(p+1)}(t)$, for $p = 0, \ldots, n-1$, exist almost everywhere and are Lebesgue integrable. Furthermore, suppose that

$$\dot{\mathbf{y}}_c^{(n-1)}(t) = \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t), \tag{2}$$

with

$$\mathbf{y}_u^{(n)}(t) = \mathbf{s}_u\big(\mathbf{x}(t)\big), \tag{3}$$

where $\mathbf{s}_u\big(\mathbf{x}(t)\big)$ is a fully unknown function of the absolutely continuous states $\mathbf{x}(t)$ only, and

$$\mathbf{y}_a^{(n)}(t) = \mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big), \tag{4}$$

where $\mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big)$ is a function of the current outputs $\mathbf{y}(t)$ and inputs $\mathbf{u}(t)$ that is partially known with some modeling error. In addition, suppose that $\mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big)$ is an affine function of the inputs, that is,

$$\mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big) = \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{u}(t), \tag{5}$$

for some vector $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ and some invertible matrix $\mathbf{B}_a\big(\mathbf{y}(t)\big)$, which implies that $n_c = n_u$. As mentioned, $\mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big)$ is partially known with some modeling error as

$$\tilde{\mathbf{s}}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big) = \tilde{\boldsymbol{\beta}}_a\big(\mathbf{y}(t)\big) + \tilde{\mathbf{B}}_a\big(\mathbf{y}(t)\big)\mathbf{u}(t), \tag{6}$$

where $\tilde{\boldsymbol{\beta}}_a\big(\mathbf{y}(t)\big)$ and $\tilde{\mathbf{B}}_a\big(\mathbf{y}(t)\big)$ are the modeled counterparts of $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ and $\mathbf{B}_a\big(\mathbf{y}(t)\big)$.

The goal is to control the outputs $\mathbf{y}_c(t)$ to the setpoints $\mathbf{r}(t)$ despite the fact that the function $\mathbf{s}_u\big(\mathbf{x}(t)\big)$ that corresponds to the derivative $\mathbf{y}_u^{(n)}(t)$ is fully unknown, the functions $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ and $\mathbf{B}_a\big(\mathbf{y}(t)\big)$ are partially known with some modeling error, the derivatives $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, are unknown, and the inputs and outputs $\mathbf{u}(t)$ and $\mathbf{y}(t)$ are subject to disturbances. The proposed control strategy must also ensure elimination of steady-state error and rejection of constant input disturbances without any integral term, as well as fast convergence of the controlled outputs to the setpoints. It must also be possible to assess the stability and performance of this strategy for a given system.

3 | CONTROL STRATEGY

This section describes the control strategy via feedback linearization and continuous-time estimation of unknown derivatives.

3.1 | Control via feedback linearization

If the true system (1) were perfectly known, one could use the input-output feedback linearization law[10]

$$\mathbf{u}(t) = \mathbf{B}_a\big(\mathbf{y}(t)\big)^{-1}\Big(\mathbf{v}(t) - \mathbf{y}_u^{(n)}(t) - \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)\Big) \tag{7}$$

to set the rates of variation $\mathbf{v}(t)$ for $\mathbf{y}_c^{(n-1)}(t)$ as follows:

$$\dot{\mathbf{y}}_c^{(n-1)}(t) = \mathbf{v}(t). \tag{8}$$

In the remainder, we assume that $\mathbf{r}^{(n-1)}(t)$ is piecewise-constant but may be discontinuous, which implies that $\dot{\mathbf{r}}^{(n-1)}(t) = \mathbf{0}_{n_c}$ between time instants with discontinuity of $\mathbf{r}^{(n-1)}(t)$, while $\mathbf{r}^{(p)}(t)$, for $p = 0, \ldots, n-2$, are absolutely continuous functions of t, which implies that the time derivatives $\dot{\mathbf{r}}^{(p)}(t) = \mathbf{r}^{(p+1)}(t)$, for $p = 0, \ldots, n-2$, exist almost everywhere and are Lebesgue integrable. Then, for the rates of variation $\mathbf{v}(t)$, one could choose the control law

$$\begin{aligned}\mathbf{v}(t) &= \tau_c^{-n}\big(\mathbf{r}(t) - \mathbf{y}_c(t)\big) + \textstyle\sum_{p=1}^{n-1}\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \mathbf{y}_c^{(p)}(t)\big)\\ &= \alpha\tau_c^{-n}\big(\mathbf{r}(t) - \mathbf{y}_c(t)\big) + \textstyle\sum_{p=0}^{n-1}\left(1 - \alpha 0^p\right)\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \mathbf{y}_c^{(p)}(t)\big),\end{aligned} \tag{9}$$

with $0 \leq \alpha \leq 1$, which would imply that

$$\Big[\dot{\mathbf{r}}^{(n-1)}(t)^{\mathrm{T}} - \dot{\mathbf{y}}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \dot{\mathbf{r}}(t)^{\mathrm{T}} - \dot{\mathbf{y}}_c(t)^{\mathrm{T}}\Big]^{\mathrm{T}} = \mathbf{A}_c\Big[\mathbf{r}^{(n-1)}(t)^{\mathrm{T}} - \mathbf{y}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \mathbf{r}(t)^{\mathrm{T}} - \mathbf{y}_c(t)^{\mathrm{T}}\Big]^{\mathrm{T}}, \tag{10}$$

between time instants with discontinuity of $\mathbf{r}^{(n-1)}(t)$, with

$$\mathbf{A}_c = \begin{bmatrix} \mathbf{A}_{c,n-1} & \cdots & \mathbf{A}_{c,0} \\ \mathbf{I}_{n_c} & \mathbf{0}_{n_c\times(n-2)n_c} & \mathbf{0}_{n_c\times n_c} \\ \mathbf{0}_{(n-2)n_c\times n_c} & \mathbf{I}_{(n-2)n_c} & \mathbf{0}_{(n-2)n_c\times n_c}\end{bmatrix}, \quad \mathbf{A}_{c,p} = -\binom{n}{p}\tau_c^{p-n}\mathbf{I}_{n_c}, \quad p = 0, \ldots, n-1, \tag{11}$$

and all the eigenvalues of $\mathbf{A}_c$ equal to $-\tau_c^{-1}$. This would result in the solution

$$\Big[\mathbf{r}^{(n-1)}(t)^{\mathrm{T}} - \mathbf{y}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \mathbf{r}(t)^{\mathrm{T}} - \mathbf{y}_c(t)^{\mathrm{T}}\Big]^{\mathrm{T}} = \exp\big(\mathbf{A}_c(t - t_0)\big)\Big[\mathbf{r}^{(n-1)}(t_0)^{\mathrm{T}} - \mathbf{y}_c^{(n-1)}(t_0)^{\mathrm{T}} \ \cdots \ \mathbf{r}(t_0)^{\mathrm{T}} - \mathbf{y}_c(t_0)^{\mathrm{T}}\Big]^{\mathrm{T}}, \tag{12}$$

where t_0 is the previous time instant with discontinuity of $\mathbf{r}^{(n-1)}(t)$, which would ensure exponential convergence of $\mathbf{y}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0, \ldots, n-1$, with time constants τ_c in the ideal case of perfectly known functions $\mathbf{s}_u\big(\mathbf{x}(t)\big)$, $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$, and $\mathbf{B}_a\big(\mathbf{y}(t)\big)$, known derivatives $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and no input or output disturbances.

However, as mentioned, the function $\mathbf{s}_u\big(\mathbf{x}(t)\big)$ that corresponds to the derivative $\mathbf{y}_u^{(n)}(t)$ is fully unknown, while the functions $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ and $\mathbf{B}_a\big(\mathbf{y}(t)\big)$ are partially known with some modeling error. Likewise, the derivatives $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, are unknown. Furthermore, there is a difference between the system inputs $\mathbf{u}(t)$ and the actuator inputs $\tilde{\mathbf{u}}(t)$ due to the input disturbances $\mathbf{d}(t) := \mathbf{u}(t) - \tilde{\mathbf{u}}(t)$ and a difference between the sensor outputs $\tilde{\mathbf{y}}(t)$ and the system outputs $\mathbf{y}(t)$ due to the output disturbances $\mathbf{w}(t) := \tilde{\mathbf{y}}(t) - \mathbf{y}(t)$. We also define the sensor outputs $\tilde{\mathbf{y}}_c(t) := \mathbf{S}\tilde{\mathbf{y}}(t)$ and the disturbances $\mathbf{w}_c(t) := \tilde{\mathbf{y}}_c(t) - \mathbf{y}_c(t)$.

In this case, it is not possible to use directly the control laws (7) and (9). However, one can approximate them by replacing $\mathbf{u}(t)$ and $\mathbf{y}(t)$ by $\tilde{\mathbf{u}}(t)$ and $\tilde{\mathbf{y}}(t)$, $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ and $\mathbf{B}_a\big(\mathbf{y}(t)\big)$ by $\tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big)$ and $\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)$, as well as the unknown derivatives $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\mathbf{y}_u^{(n)}(t)$ by their estimates $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$, which results in:

$$\tilde{\mathbf{u}}(t) = \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)^{-1}\Big(\tilde{\mathbf{v}}(t) - \hat{\mathbf{y}}_u^{(n)}(t) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big), \tag{13}$$

$$\tilde{\mathbf{v}}(t) = \alpha\tau_c^{-n}\big(\mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)\big) + \textstyle\sum_{p=0}^{n-1}\left(1 - \alpha 0^p\right)\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \hat{\mathbf{y}}_c^{(p)}(t)\big). \tag{14}$$

The estimation of these derivatives is investigated in the next subsection.

3.2 | Estimation of unknown derivatives

One can use (2)–(4) to estimate the values $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$ of the unknown derivatives by applying an FIR filter to the sensor outputs $\tilde{\mathbf{y}}_c(t)$ and the available rates

$$\tilde{\mathbf{y}}_a^{(n)}(t) = \tilde{\mathbf{s}}_a\big(\tilde{\mathbf{y}}(t), \tilde{\mathbf{u}}(t)\big), \tag{15}$$

in the interval $[t-\Delta t, t]$, where Δt is the size of the filter window.

Remark 1. One can also apply an infinite impulse response (IIR) filter to $\tilde{\mathbf{y}}_c(t)$ and $\tilde{\mathbf{y}}_a^{(n)}(t)$, which results in a linear observer[31,32]. However, the following results about derivative estimation refer to the FIR filter. These results were discussed previously for derivatives up to first order in the discrete-time case[36] and in the continuous-time case[38] and are extended here to derivatives up to arbitrary order in the continuous-time case such that closed-form expressions for the FIR filter and the corresponding differences between derivative estimates and true derivatives are still obtained. As discussed in the remainder of this section, the advantages of the FIR filter with respect to an IIR filter are the property of deadbeat estimation of the unknown derivatives after the window size Δt, instead of asymptotic convergence, and the minimization of the effect of measurement noise $\mathbf{w}_c(t)$ for a given Δt without the need to tune the observer parameters.

Remark 2. Since $\mathbf{r}^{(n-1)}(t)$ may be discontinuous, the same is valid for $\tilde{\mathbf{u}}(t)$ and $\tilde{\mathbf{y}}_a^{(n)}(t)$ and consequently for $\mathbf{u}(t)$, $\mathbf{y}_a^{(n)}(t)$, and $\mathbf{y}_c^{(n)}(t)$. On the other hand, since $\mathbf{x}(t)$ is absolutely continuous, the same is valid for $\mathbf{y}(t)$, $\mathbf{y}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\mathbf{y}_u^{(n)}(t)$. We also assume that $\mathbf{w}^{(p)}(t)$, for $p = 0, \ldots, n-1$ are absolutely continuous, thus the same is valid for $\tilde{\mathbf{y}}(t)$ and $\tilde{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$. The fact that these functions are absolutely continuous is used in several technical results in this section.

For any distance z from the beginning of the filter window, we propose to compute the estimates $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ as the quantities that minimize the mean squared error between the measured sensor outputs $\tilde{\mathbf{y}}_c(t-\Delta t+\tau)$ and the predicted outputs $\hat{\mathbf{y}}_c(t-\Delta t+\tau)$ for $\tau \in [0, \Delta t]$, with the goal of minimizing the effect of $\mathbf{w}_c(t)$ for a given Δt. This mean squared error is defined as

$$J_{n,z}(t) = \frac{1}{\Delta t}\int_0^{\Delta t} \hat{\boldsymbol{\epsilon}}_{n,z}(t-\Delta t+\tau)^{\mathrm{T}} \hat{\boldsymbol{\epsilon}}_{n,z}(t-\Delta t+\tau)\mathrm{d}\tau, \tag{16}$$

where the prediction errors are

$$\hat{\boldsymbol{\epsilon}}_{n,z}(t-\Delta t+\tau) = \tilde{\mathbf{y}}_c(t-\Delta t+\tau) - \sum_{l=0}^{n-1} \frac{(\tau-z)^l}{l!}\hat{\mathbf{y}}_c^{(l)}(t-\Delta t+z) - \int_z^{\tau} \frac{(\tau-\zeta)^{n-1}}{(n-1)!}\hat{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)\mathrm{d}\zeta, \tag{17}$$

with

$$\hat{\mathbf{y}}_c^{(n)}(t-\Delta t+\tau) = \hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z) + \tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau). \tag{18}$$

These derivative estimates are given by the following lemma.

Lemma 1. The predicted outputs $\hat{\mathbf{y}}_c(t)$ that minimize the mean squared error $J_{n,z}(t)$ are given by

$$\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z) = \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t} \left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} \boldsymbol{v}_{n,z}(t-\Delta t+\tau)\mathrm{d}\tau, \qquad p = 0, \ldots, n-1, \tag{19}$$

$$\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z) = \frac{n!}{\Delta t^{n+1}}\int_0^{\Delta t} \left(\mathbf{c}_{n,z}(\tau)\right)_{n+1} \boldsymbol{v}_{n,z}(t-\Delta t+\tau)\mathrm{d}\tau, \tag{20}$$

where

$$\boldsymbol{v}_{n,z}(t-\Delta t+\tau) = \tilde{\mathbf{y}}_c(t-\Delta t+\tau) - \int_z^{\tau} \frac{(\tau-\zeta)^{n-1}}{(n-1)!}\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)\mathrm{d}\zeta, \tag{21}$$

the $(n+1)$-dimensional vector function $\mathbf{c}_{n,z}(\tau)$ is defined as

$$\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} = \sum_{i=0}^{n} \left(\frac{\tau-z}{\Delta t}\right)^i \left(\mathbf{A}_n^r\right)_{i+1,p+1}, \qquad p = 0, \ldots, n, \tag{22}$$

and the $(n+1)\times(n+1)$ matrix $\mathbf{A}_n^r$ is the inverse of $\mathbf{G}_n^r$ defined as

$$\left(\mathbf{G}_n^r\right)_{l+1,i+1} = \int_{r-1}^{r} x^{i+l}\mathrm{d}x = \frac{r^{i+l+1}-(r-1)^{i+l+1}}{i+l+1}, \qquad l = 0, \ldots, n, \qquad i = 0, \ldots, n, \tag{23}$$

with $r = 1 - \frac{z}{\Delta t}$. Furthermore, $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ are the linear unbiased estimators of $\mathbf{y}_c^{(p)}(t-\Delta t+z)$, for $p = 0, \ldots, n-1$, and $\mathbf{y}_u^{(n)}(t-\Delta t+z)$ that provide minimal variance if the following assumptions hold:

1. The unknown derivatives $\mathbf{y}_u^{(n)}(t-\Delta t+\tau)$ are constant for $\tau \in [0, \Delta t]$.
2. The mismatch between $\tilde{\mathbf{y}}_a^{(n)}(t)$ and $\mathbf{y}_a^{(n)}(t)$ is negligible in comparison with the measurement noise $\mathbf{w}_c(t)$.
3. The measured outputs $\tilde{\mathbf{y}}_c(t)$ are corrupted by zero-mean noise $\mathbf{w}_c(t)$.
4. The noise $\mathbf{w}_c(t)$ is uncorrelated and homoscedastic with covariance $\boldsymbol{\Sigma}_{\mathbf{w}_c}$, which implies that, for any L^2 function $g(t)$,

$$\mathrm{Var}\left[\frac{1}{\Delta t}\int_0^{\Delta t} \mathbf{w}_c(t-\Delta t+\tau)g(\tau)\mathrm{d}\tau\right] = \frac{1}{\Delta t}\int_0^{\Delta t} g(\tau)^2\mathrm{d}\tau\boldsymbol{\Sigma}_{\mathbf{w}_c}. \tag{24}$$

Proof. See Appendix A. □

Remark 3. Lemma 1 can be regarded as an adaptation of Gauss-Markov Theorem to the context of continuous-time observations since Gauss-Markov Theorem also states that the estimator that minimizes the mean squared error is the linear unbiased estimator that provides minimal variance if the observation errors are uncorrelated and homoscedastic and have mean equal to zero.

Hence, to determine $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p=0,\dots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$, it is necessary to know $\mathbf{A}_n^r$. For this, we use the concept of shifted Legendre polynomials, which allows us to obtain expressions for the elements of $\mathbf{A}_n^r$ according to the following lemma.

Lemma 2. The elements of $\mathbf{A}_n^r$ are given by

$$\left(\mathbf{A}_n^1\right)_{i+1,m+1} = (-1)^{i+m}\frac{\prod_{k=0}^n (n+i+1-k)(n+m+1-k)}{(i+m+1)i!(n-i)!m!(n-m)!}, \tag{25}$$

$$\left(\mathbf{A}_n^{1/2}\right)_{i+1,m+1} = 2^{i+m}\frac{\prod_{k=0}^n (n+i+1-2k)(n+m+1-2k)+\prod_{k=0}^n (n+i-2k)(n+m-2k)}{(i+m+1)i!(n-i)!m!(n-m)!}, \tag{26}$$

$$\left(\mathbf{A}_n^0\right)_{i+1,m+1} = \frac{\prod_{k=0}^n (n+i+1-k)(n+m+1-k)}{(i+m+1)i!(n-i)!m!(n-m)!}, \tag{27}$$

for $r\in\{0,1/2,1\}$. In addition, for all $0\le s\le 1$,

$$\left(\mathbf{A}_n^s\right)_{i+1,p+1} = \sum_{j=i}^n\sum_{m=p}^n \binom{j}{i}\binom{m}{p}(r-s)^{j-i+m-p}\left(\mathbf{A}_n^r\right)_{j+1,m+1}, \qquad i=0,\dots,n, \qquad p=0,\dots,n. \tag{28}$$

Proof. See Appendix B. □

The following lemma builds on the previous results and shows the outcome of applying a more general version of the FIR filter in Lemma 1 to $\tilde{\mathbf{y}}_c$, instead of $\boldsymbol{v}_{n,z}$, using the definition

$$\tilde{\mathbf{y}}_u^{(n)}(t) = \tilde{\mathbf{y}}_c^{(n)}(t) - \tilde{\mathbf{y}}_a^{(n)}(t) = \mathbf{w}_c^{(n)}(t) + \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t) - \tilde{\mathbf{y}}_a^{(n)}(t). \tag{29}$$

Lemma 3. Let $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c,t)$ be the pth differentiation continuous-time FIR filter of order n and window size Δt given for the point at a distance z from the beginning of the filter window and applied to the function $\tilde{\mathbf{y}}_c$ on the interval $[t-\Delta t,t]$. If $\tilde{\mathbf{y}}_c^{(n-1)}$ is absolutely continuous and $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c,t)$ is defined as

$$\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c,t) := \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\tilde{\mathbf{y}}_c(t-\Delta t+\tau)\mathrm{d}\tau, \tag{30}$$

where, for $r\in\{0,1/2,1\}$ and all $0\le s\le 1$,

$$\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} = \sum_{m=p}^n\binom{m}{p}\left(\frac{z}{\Delta t}-1+r\right)^{m-p}\left(\mathbf{c}_n(\tau)\right)_{m+1} = \sum_{i=0}^n\left(\frac{\tau}{\Delta t}-1+s\right)^i\left(\mathbf{a}_{p,n,z}^s\right)_{i+1}, \tag{31}$$

with

$$\left(\mathbf{c}_n(\tau)\right)_{m+1} = \sum_{i=0}^n\left(\frac{\tau}{\Delta t}-1+s\right)^i\sum_{j=i}^n\binom{j}{i}(r-s)^{j-i}\left(\mathbf{A}_n^r\right)_{j+1,m+1}, \qquad m=0,\dots,n, \tag{32}$$

$$\left(\mathbf{a}_{p,n,z}^s\right)_{i+1} = \sum_{j=i}^n\binom{j}{i}(r-s)^{j-i}\sum_{m=p}^n\binom{m}{p}\left(\frac{z}{\Delta t}-1+r\right)^{m-p}\left(\mathbf{A}_n^r\right)_{j+1,m+1}, \qquad i=0,\dots,n, \tag{33}$$

where $\mathbf{A}_n^r$ is the inverse of $\mathbf{G}_n^r$, then

$$\begin{aligned}\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c,t) &= \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z) + \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\ &= \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z) + \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau, \quad p=0,\dots,n-1,\end{aligned} \tag{34}$$

$$\begin{aligned}\mathcal{H}_{n,n,z}(\tilde{\mathbf{y}}_c,t) &= \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z) + \frac{n!}{\Delta t^{n+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{n+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\ &= \frac{n!}{\Delta t^{n+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{n+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau.\end{aligned} \tag{35}$$

Proof. See Appendix C. □

Remark 4. The function $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c,t)$ corresponds to the continuous-time version of the pth-differentiation Savitzky-Golay filter of order n, which was developed originally for the discrete-time case[39] and has been called algebraic time-derivative estimation in the continuous-time case[30]. Even though the FIR filters in this section coincide with the FIR filters in previous work[29,30], this section provides several new auxiliary results and generalizations that are necessary before these FIR filters can be applied to multivariable control and their implications on the performance of the control strategy can be properly studied. Also, the use of a smaller window size Δt of the differentiation filter amplifies the effect of measurement noise in $\mathbf{w}_c(t)$.

Lemma 3 uses the fact that $\tilde{\mathbf{y}}_c^{(n-1)}$ is absolutely continuous to derive expressions for all the necessary derivative estimates and relates the remainder $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) - \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$ (or $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ if $p=n$) to a double integral of $\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z) = \tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta) + \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ for $\zeta \in [0, \Delta t]$. Now we aim to obtain a more explicit expression for this remainder in terms of a single convolution integral of the known function $\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)$ for $\tau \in [0, \Delta t]$ and of the unknown function $\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ for $\tau \in [0, \Delta t]$, which is done in the following lemma.

Lemma 4. If $\tilde{\mathbf{y}}_c^{(n-1)}$ is absolutely continuous and $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t)$ is defined as in Lemma 3, then

$$\begin{aligned}\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) &= \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z) + \mathcal{V}_{p,n,z}(\tilde{\mathbf{y}}_u^{(n)}, t) + \mathcal{Z}_{p,n,z}(\tilde{\mathbf{y}}_a^{(n)}, t) \\ &= \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z) + \mathcal{Z}_{p,n,z}(\tilde{\mathbf{y}}_u^{(n)}, t) + \mathcal{Z}_{p,n,z}(\tilde{\mathbf{y}}_a^{(n)}, t), \quad p = 0,\ldots,n-1, \end{aligned} \tag{36}$$

$$\begin{aligned}\mathcal{H}_{n,n,z}(\tilde{\mathbf{y}}_c, t) &= \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z) + \mathcal{V}_{n,n,z}(\tilde{\mathbf{y}}_u^{(n)}, t) + \mathcal{Z}_{n,n,z}(\tilde{\mathbf{y}}_a^{(n)}, t) \\ &= \mathcal{Z}_{n,n,z}(\tilde{\mathbf{y}}_u^{(n)}, t) + \mathcal{Z}_{n,n,z}(\tilde{\mathbf{y}}_a^{(n)}, t),\end{aligned} \tag{37}$$

with

$$\begin{aligned}\mathcal{V}_{p,n,z}(\tilde{\mathbf{y}}_u^{(n)}, t) &= \tfrac{p!}{\Delta t^{p+1}}\left(\int_0^z \left(\mathbf{c}_{n,z,1}^{(-n)}(\tau)\right)_{p+1} \tfrac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^n}\mathrm{d}\tau + \int_z^{\Delta t} \left(\mathbf{c}_{n,z,0}^{(-n)}(\tau)\right)_{p+1} \tfrac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^n}\mathrm{d}\tau\right) \\ &= \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t} \left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} \int_z^{\tau} \tfrac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\,\mathrm{d}\zeta\mathrm{d}\tau,\end{aligned} \tag{38}$$

$$\begin{aligned}\mathcal{Z}_{p,n,z}(\tilde{\mathbf{y}}_a^{(n)}, t) &= \tfrac{p!}{\Delta t^{p+1}}\left(\int_0^z \left(\mathbf{c}_{n,z,1}^{(-n)}(\tau)\right)_{p+1} \tfrac{\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)}{(-1)^n}\mathrm{d}\tau + \int_z^{\Delta t} \left(\mathbf{c}_{n,z,0}^{(-n)}(\tau)\right)_{p+1} \tfrac{\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)}{(-1)^n}\mathrm{d}\tau\right) \\ &= \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t} \left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} \int_z^{\tau} \tfrac{\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)}{(n-1)!}(\tau-\zeta)^{n-1}\,\mathrm{d}\zeta\mathrm{d}\tau,\end{aligned} \tag{39}$$

where

$$\left(\mathbf{c}_{n,z,s}^{(-n)}(\tau)\right)_{p+1} = \sum_{i=0}^{n}\left(\tfrac{\tau}{\Delta t} - 1 + s\right)^{i+n} \tfrac{\Delta t^n i!}{(i+n)!}\left(\mathbf{a}_{p,n,z}^s\right)_{i+1} \tag{40}$$

is the nth antiderivative of $\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}$ such that $\left(\mathbf{c}_{n,z,s}^{(-n)}((1-s)\Delta t)\right)_{p+1} = 0$.

Proof. See Appendix D. □

Lemma 4 shows that $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) - \tilde{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$ (or $\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ if $p=n$) discounted by a single convolution integral of the known function $\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)$ for $\tau \in [0,\Delta t]$ equals the same convolution integral of the unknown function $\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ for $\tau \in [0,\Delta t]$. The particular case $p=1$, $n=1$, and $z=\Delta t$ is shown below.

Lemma 5. If $\tilde{\mathbf{y}}_c$ is absolutely continuous and $\mathcal{D}(\tilde{\mathbf{y}}_c, t)$ is defined as

$$\mathcal{D}(\tilde{\mathbf{y}}_c, t) := \mathcal{H}_{1,1,\Delta t}(\tilde{\mathbf{y}}_c, t) = \tfrac{1}{\Delta t^2}\int_0^{\Delta t}\left(\mathbf{c}_{1,\Delta t}(\tau)\right)_2 \tilde{\mathbf{y}}_c(t-\Delta t+\tau)\mathrm{d}\tau, \tag{41}$$

where

$$\left(\mathbf{c}_{1,\Delta t}(\tau)\right)_2 = \left(\mathbf{a}_{1,1,\Delta t}^1\right)_1 + \left(\tfrac{\tau}{\Delta t}\right)\left(\mathbf{a}_{1,1,\Delta t}^1\right)_2, \tag{42}$$

with

$$\left(\mathbf{a}_{1,1,\Delta t}^1\right)_1 = -6, \qquad \left(\mathbf{a}_{1,1,\Delta t}^1\right)_2 = 12, \tag{43}$$

then

$$\mathcal{D}(\tilde{\mathbf{y}}_c, t) = \tilde{\mathbf{y}}_u(t) + \mathcal{R}(\tilde{\mathbf{y}}_u, t) + \mathcal{W}(\tilde{\mathbf{y}}_a, t) = \mathcal{W}(\tilde{\mathbf{y}}_u, t) + \mathcal{W}(\tilde{\mathbf{y}}_a, t), \tag{44}$$

with

$$\mathcal{R}(\tilde{\mathbf{y}}_u, t) := \mathcal{V}_{1,1,\Delta t}(\tilde{\mathbf{y}}_u, t) = \int_0^{\Delta t} 6\left(\tfrac{\tau}{\Delta t} - \left(\tfrac{\tau}{\Delta t}\right)^2\right)\tfrac{\tilde{\mathbf{y}}_u(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u(t)}{\Delta t}\mathrm{d}\tau, \tag{45}$$

$$\mathcal{W}(\tilde{\mathbf{y}}_a, t) := \mathcal{Z}_{1,1,\Delta t}(\tilde{\mathbf{y}}_a, t) = \int_0^{\Delta t} 6\left(\tfrac{\tau}{\Delta t} - \left(\tfrac{\tau}{\Delta t}\right)^2\right)\tfrac{\tilde{\mathbf{y}}_a(t-\Delta t+\tau)}{\Delta t}\mathrm{d}\tau. \tag{46}$$

Proof. See Appendix E. □

Then, we can note that, according to the definitions in Lemmas 3 and 4,

$$\hat{\mathbf{y}}_c^{(p)}(t) = \mathcal{H}_{p,n,\Delta t}(\tilde{\mathbf{y}}_c, t) - \mathcal{Z}_{p,n,\Delta t}(\tilde{\mathbf{y}}_a^{(n)}, t), \quad p = 0,\ldots,n-1, \tag{47}$$

$$\hat{\mathbf{y}}_u^{(n)}(t) = \mathcal{H}_{n,n,\Delta t}(\tilde{\mathbf{y}}_c, t) - \mathcal{Z}_{n,n,\Delta t}(\tilde{\mathbf{y}}_a^{(n)}, t), \tag{48}$$

which implies that, from (36) and (37),

$$\hat{\mathbf{y}}_c^{(p)}(t) = \tilde{\mathbf{y}}_c^{(p)}(t) + \mathcal{V}_{p,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t) = \tilde{\mathbf{y}}_c^{(p)}(t) + \mathcal{Z}_{p,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t), \quad p = 0, \ldots, n-1, \tag{49}$$

$$\hat{\mathbf{y}}_u^{(n)}(t) = \tilde{\mathbf{y}}_u^{(n)}(t) + \mathcal{V}_{n,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t) = \mathcal{Z}_{n,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t). \tag{50}$$

In particular, if $n = 1$, we can note that, according to the definitions in Lemmas 3 and 4,

$$\hat{\mathbf{y}}_u(t) = \mathcal{H}_{1,1,\Delta t}(\tilde{\mathbf{y}}_c, t) - \mathcal{Z}_{1,1,\Delta t}(\tilde{\mathbf{y}}_a, t), \tag{51}$$

which implies that, from (36) and (37),

$$\hat{\mathbf{y}}_u(t) = \tilde{\mathbf{y}}_u(t) + \mathcal{V}_{1,1,\Delta t}(\tilde{\mathbf{y}}_u, t) = \mathcal{Z}_{1,1,\Delta t}(\tilde{\mathbf{y}}_u, t). \tag{52}$$

Remark 5. From (49), (50), and the definition of $\mathcal{V}_{p,n,z}$ in Lemma 4, one can observe that, if $\tilde{\mathbf{y}}_u^{(n)}(t - \Delta t + \tau)$ is constant for $\tau \in [0, \Delta t]$, the estimates $\hat{\mathbf{y}}_c^{(p)}(t)$ and $\hat{\mathbf{y}}_u^{(n)}(t)$ yield the desired values $\tilde{\mathbf{y}}_c^{(p)}(t)$ and $\tilde{\mathbf{y}}_u^{(n)}(t)$. This results in the deadbeat property claimed in Remark 1. In addition, as shown in Lemma 1, these estimates minimize both the mean squared error and the effect of measurement noise $\mathbf{w}_c(t)$ for a given Δt as claimed in Remark 1.

In summary, the proposed controller via feedback linearization and continuous-time derivative estimation is given by (13), (14) with (6), (15), (47), (48) for the system (2)–(5). The next section shows some properties provided by this control strategy.

3.3 | Laplace transform of the derivative estimates

Before we describe the properties of control via feedback linearization and continuous-time derivative estimation in the next section, it is necessary to convert the estimation of unknown derivatives developed for the time domain in this section to the frequency domain by applying the Laplace transform. This result will be used to establish the convergence properties of the control strategy in Theorem 8. It is assumed that, for each time-varying signal $\mathbf{s}(t)$, the initial conditions $\mathbf{s}(0)$ correspond to the equilibrium point $\bar{\mathbf{s}}$, deviation variables are defined as $\delta\mathbf{s}(t) := \mathbf{s}(t) - \bar{\mathbf{s}}$, which implies that the initial conditions $\delta\mathbf{s}(0)$ are zero and can be omitted, and the corresponding Laplace transforms are denoted as $\mathbf{S}(s)$. The Laplace transform of the derivative estimates $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$ is obtained in the following lemma.

Lemma 6. The estimation of $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$ in (47) and (48), which implies (49) and (50), is expressed in the frequency domain by applying the Laplace transform as

$$\begin{aligned}
\hat{\mathbf{Y}}_c^{(p)}(s) = {} & \tfrac{p!}{\Delta t^{p+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^i \sum_{j=i}^{n} \binom{j}{i} (-1)^{j-i} \left(\mathbf{A}_n^0\right)_{j+1,p+1} \exp(-st)\, \mathrm{d}t \mathbf{S}\tilde{\mathbf{Y}}(s) \\
& - \tfrac{p!}{\Delta t^{p+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^{i+n} \tfrac{\Delta t^n i!}{(i+n)!} \sum_{j=i}^{n} \binom{j}{i} (-1)^{n+j-i} \left(\mathbf{A}_n^0\right)_{j+1,p+1} \exp(-st)\, \mathrm{d}t \tilde{\mathbf{Y}}_a^{(n)}(s) \\
= {} & \tilde{\mathbf{Y}}_c^{(p)}(s) \\
& + \tfrac{p!}{\Delta t^{p+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^{i+n} \tfrac{\Delta t^n i!}{(i+n)!} \sum_{j=i}^{n} \binom{j}{i} (-1)^{n+j-i} \left(\mathbf{A}_n^0\right)_{j+1,p+1} \exp(-st)\, \mathrm{d}t \tilde{\mathbf{Y}}_u^{(n)}(s), \quad p = 0, \ldots, n-1,
\end{aligned} \tag{53}$$

$$\begin{aligned}
\hat{\mathbf{Y}}_u^{(n)}(s) = {} & \tfrac{n!}{\Delta t^{n+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^i \sum_{j=i}^{n} \binom{j}{i} (-1)^{j-i} \left(\mathbf{A}_n^0\right)_{j+1,n+1} \exp(-st)\, \mathrm{d}t \mathbf{S}\tilde{\mathbf{Y}}(s) \\
& - \tfrac{n!}{\Delta t^{n+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^{i+n} \tfrac{\Delta t^n i!}{(i+n)!} \sum_{j=i}^{n} \binom{j}{i} (-1)^{n+j-i} \left(\mathbf{A}_n^0\right)_{j+1,n+1} \exp(-st)\, \mathrm{d}t \tilde{\mathbf{Y}}_a^{(n)}(s) \\
= {} & \tfrac{n!}{\Delta t^{n+1}} \int_0^{\Delta t} \textstyle\sum_{i=0}^{n} \left(1 - \tfrac{t}{\Delta t}\right)^{i+n} \tfrac{\Delta t^n i!}{(i+n)!} \sum_{j=i}^{n} \binom{j}{i} (-1)^{n+j-i} \left(\mathbf{A}_n^0\right)_{j+1,n+1} \exp(-st)\, \mathrm{d}t \tilde{\mathbf{Y}}_u^{(n)}(s).
\end{aligned} \tag{54}$$

Proof. See Appendix F. □

4 | PROPERTIES OF THE CONTROL STRATEGY

This section shows the steady-state and convergence properties provided by the control strategy developed in the previous section, presents methods to assess its stability and performance, and addresses practical aspects related to the control strategy.

4.1 | Steady-state and convergence properties

We start by proving that, although this controller does not include any integral term, it eliminates steady-state error and rejects constant input disturbances, as shown by the following theorem.

Theorem 7. The controller given by (13), (14) with (6), (15), (47), (48) for the system (2)–(5) eliminates steady-state error and rejects constant input disturbances if steady state is reached.

Proof. Firstly, note that the controller with (47), (48) enforces (49), (50) with (29). The evaluation is performed at steady state, thus $\mathbf{r}(t) = \bar{\mathbf{r}}$, $\mathbf{d}(t) = \bar{\mathbf{d}}$, $\mathbf{w}(t) = \bar{\mathbf{w}}$, $\mathbf{u}(t) = \bar{\mathbf{u}}$, $\mathbf{y}(t) = \bar{\mathbf{y}}$, $\mathbf{x}(t) = \bar{\mathbf{x}}$, and $\mathbf{y}_c^{(p)}(t) = \mathbf{r}^{(p)}(t) = \mathbf{w}_c^{(p)}(t) = \mathbf{0}_{n_c}$ for $p = 1, \ldots, n$, that is, $\mathbf{y}_c^{(p)}(t) = 0^p \bar{\mathbf{y}}_c$, $\mathbf{r}^{(p)}(t) = 0^p \bar{\mathbf{r}}$, $\mathbf{w}_c^{(p)}(t) = 0^p \bar{\mathbf{w}}_c$ for $p = 0, \ldots, n$. Then, (13), (14), (15), (29), (49), (50) are written as

$$\bar{\mathbf{u}} - \bar{\mathbf{d}} = \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})^{-1}\Big(\tilde{\mathbf{v}}(t) - \hat{\mathbf{y}}_u^{(n)}(t) - \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})\Big), \tag{55}$$

$$\tilde{\mathbf{v}}(t) = \alpha\tau_c^{-n}(\bar{\mathbf{r}} - \bar{\mathbf{y}}_c - \bar{\mathbf{w}}_c) + \textstyle\sum_{p=0}^{n-1}(1 - \alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(0^p\bar{\mathbf{r}} - \hat{\mathbf{y}}_c^{(p)}(t)\big), \tag{56}$$

$$\tilde{\mathbf{y}}_a^{(n)}(t) = \tilde{\mathbf{s}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}, \bar{\mathbf{u}} - \bar{\mathbf{d}}), \tag{57}$$

$$\tilde{\mathbf{y}}_u^{(n)}(t) = \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t) - \tilde{\mathbf{y}}_a^{(n)}(t), \tag{58}$$

$$\hat{\mathbf{y}}_c^{(p)}(t) = 0^p\bar{\mathbf{y}}_c + 0^p\bar{\mathbf{w}}_c + \mathcal{V}_{p,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t), \quad p = 0, \ldots, n-1, \tag{59}$$

$$\hat{\mathbf{y}}_u^{(n)}(t) = \tilde{\mathbf{y}}_u^{(n)}(t) + \mathcal{V}_{n,n,\Delta t}(\tilde{\mathbf{y}}_u^{(n)}, t), \tag{60}$$

while (2)–(5) and (6) are written as

$$\mathbf{0}_{n_c} = \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t), \tag{61}$$

$$\mathbf{y}_u^{(n)}(t) = \mathbf{s}_u(\bar{\mathbf{x}}), \tag{62}$$

$$\mathbf{y}_a^{(n)}(t) = \mathbf{s}_a(\bar{\mathbf{y}}, \bar{\mathbf{u}}), \tag{63}$$

$$\mathbf{s}_a(\bar{\mathbf{y}}, \bar{\mathbf{u}}) = \boldsymbol{\beta}_a(\bar{\mathbf{y}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{u}}, \tag{64}$$

$$\tilde{\mathbf{s}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}, \bar{\mathbf{u}} - \bar{\mathbf{d}}) = \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}) + \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})(\bar{\mathbf{u}} - \bar{\mathbf{d}}). \tag{65}$$

This implies that

$$\tilde{\mathbf{y}}_u^{(n)}(t) = \mathbf{s}_u(\bar{\mathbf{x}}) + \boldsymbol{\beta}_a(\bar{\mathbf{y}}) - \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{d}} + \Big(\mathbf{B}_a(\bar{\mathbf{y}}) - \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})\Big)(\bar{\mathbf{u}} - \bar{\mathbf{d}}) \tag{66}$$

is constant, which means that, from the definition of $\mathcal{V}_{p,n,z}$ in Lemma 4,

$$\hat{\mathbf{y}}_c^{(p)}(t) = 0^p\bar{\mathbf{y}}_c + 0^p\bar{\mathbf{w}}_c, \quad p = 0, \ldots, n-1, \tag{67}$$

$$\hat{\mathbf{y}}_u^{(n)}(t) = \tilde{\mathbf{y}}_u^{(n)}(t). \tag{68}$$

Consequently, if steady state is reached,

$$\begin{aligned}
\mathbf{0}_{n_c} &= \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t)\\
&= \mathbf{s}_u(\bar{\mathbf{x}}) + \mathbf{s}_a(\bar{\mathbf{y}}, \bar{\mathbf{u}})\\
&= \mathbf{s}_u(\bar{\mathbf{x}}) + \boldsymbol{\beta}_a(\bar{\mathbf{y}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{u}}\\
&= \mathbf{s}_u(\bar{\mathbf{x}}) + \boldsymbol{\beta}_a(\bar{\mathbf{y}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{u}} + \tilde{\mathbf{v}}(t) - \hat{\mathbf{y}}_u^{(n)}(t) - \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}) - \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})(\bar{\mathbf{u}} - \bar{\mathbf{d}})\\
&= \mathbf{s}_u(\bar{\mathbf{x}}) + \alpha\tau_c^{-n}(\bar{\mathbf{r}} - \bar{\mathbf{y}}_c - \bar{\mathbf{w}}_c) + \textstyle\sum_{p=0}^{n-1}(1 - \alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(0^p\bar{\mathbf{r}} - \hat{\mathbf{y}}_c^{(p)}(t)\big) - \hat{\mathbf{y}}_u^{(n)}(t)\\
&\quad + \boldsymbol{\beta}_a(\bar{\mathbf{y}}) - \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{d}} + \Big(\mathbf{B}_a(\bar{\mathbf{y}}) - \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})\Big)(\bar{\mathbf{u}} - \bar{\mathbf{d}})\\
&= \mathbf{s}_u(\bar{\mathbf{x}}) + \tau_c^{-n}(\bar{\mathbf{r}} - \bar{\mathbf{y}}_c - \bar{\mathbf{w}}_c) - \tilde{\mathbf{y}}_u^{(n)}(t)\\
&\quad + \boldsymbol{\beta}_a(\bar{\mathbf{y}}) - \tilde{\boldsymbol{\beta}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}}) + \mathbf{B}_a(\bar{\mathbf{y}})\bar{\mathbf{d}} + \Big(\mathbf{B}_a(\bar{\mathbf{y}}) - \tilde{\mathbf{B}}_a(\bar{\mathbf{y}} + \bar{\mathbf{w}})\Big)(\bar{\mathbf{u}} - \bar{\mathbf{d}})\\
&= \tau_c^{-n}(\bar{\mathbf{r}} - \bar{\mathbf{y}}_c - \bar{\mathbf{w}}_c),
\end{aligned} \tag{69}$$

which shows that $\bar{\mathbf{y}}_c + \bar{\mathbf{w}}_c = \bar{\mathbf{r}}$ at steady state regardless of the values of $\bar{\mathbf{r}}$, $\bar{\mathbf{d}}$, and $\bar{\mathbf{w}}$ and implies that the controller eliminates steady-state error and rejects constant input disturbances. □

Theorem 7 considers the steady-state case, which implies that the setpoints and disturbances must be constant. For this case, Theorem 7 shows that the control strategy in this paper eliminates steady-state error and rejects constant input disturbances if steady state is reached, in the same way as a proportional-integral controller, although the control strategy in this paper does not include any integral term. The discussion about the effect of time-varying setpoints and disturbances on the closed-loop stability and performance will be presented in the next subsection.

The presented control strategy requires only two design parameters with a well-defined meaning: (i) Δt is the window size of the differentiation filter, which should be sufficiently small so that the unknown function $\mathbf{s}_u\big(\mathbf{x}(t)\big)$ is approximately constant in this window and the resulting derivative estimates are sufficiently accurate, but not too small to prevent amplification of measurement noise in $\mathbf{w}_c(t)$; and (ii) τ_c is the inverse of the nth root of the controller gain in (14), which means that it should not be too small due to the noise, but sufficiently small since it is expected to be approximately equal to the dominant closed-loop time constant if the derivative estimation is sufficiently accurate and the variations of the disturbances $\mathbf{w}_c^{(n)}(t)$ and $\mathbf{d}(t)$ and of the modeling error in the known part of the model are bounded. The latter property is shown in the following theorem.

Theorem 8. Denote the supremum of the largest singular value of the matrix $\mathbf{K}_a(t)$ as $k_a := \sup_t \bar{\sigma}\big(\mathbf{K}_a(t)\big)$ and the $\mathcal{H}_\infty$-norm of the transfer function $G_e(s)\mathbf{I}_{n_c} = \hat{\mathbf{Y}}_u^{(*)}(s)\tilde{\mathbf{Y}}_u^{(n)}(s)^{-1}$ inferred from Lemma 6 as $\|G_e(s)\|_\infty := \sup_{\omega\in\mathbb{R}} |G_e(\omega j)|$, where

$$\mathbf{K}_a(t) := \Big(\mathbf{B}_a\big(\mathbf{y}(t)\big) - \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big)\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)^{-1}, \tag{70}$$

$$\hat{\mathbf{Y}}_u^{(*)}(s) := \textstyle\sum_{p=0}^{n-1} (1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(\hat{\mathbf{Y}}_c^{(p)}(s) - \tilde{\mathbf{Y}}_c^{(p)}(s)\big) + \hat{\mathbf{Y}}_u^{(n)}(s). \tag{71}$$

Suppose that (i) $k_a < \frac{1}{\|G_e(s)\|_\infty}$, (ii) the control of $\tilde{\mathbf{y}}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0, \dots, n-1$, keeps $\dot{\mathbf{x}}(t)$ bounded, and (iii) the variations of the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$ and $\mathbf{B}_a\big(\mathbf{y}(t-\Delta t+\tau)\big)\mathbf{d}(t-\Delta t+\tau)$ and of the modeling error in the known function $\boldsymbol{\beta}_a\big(\mathbf{y}(t-\Delta t+\tau)\big)$ are bounded. Then, there exists a bounded vector $\boldsymbol{\delta}_s$ and a scalar $\Delta\bar{t} > 0$ such that, for any sufficiently small window size $\Delta t \leq \Delta\bar{t}$, (i) the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$, the unknown function $\mathbf{s}_u\big(\mathbf{x}(t-\Delta t+\tau)\big)$, and the modeling error in the known function $\mathbf{s}_a\big(\mathbf{y}(t-\Delta t+\tau), \mathbf{u}(t-\Delta t+\tau)\big)$ do not vary more than $\boldsymbol{\delta}_s$ in the interval $\tau \in [0, \Delta t]$, and (ii) the controller given by (13), (14) with (6), (15), (47), (48) for the system (2)–(5) forces $\tilde{\mathbf{y}}_c^{(p)}(t)$ to converge exponentially to a distance linear in $\|\boldsymbol{\delta}_s\|$ from $\mathbf{r}^{(p)}(t)$, for $p = 0, \dots, n-1$, with time constants τ_c, that is, there exists a nondecreasing scalar function $c_s(\Delta t/\tau_c)$ and a matrix $\mathbf{K}_c\Big(\frac{t-\zeta}{\tau_c}, \tau_c\Big)$ with elements that depend polynomially on $\frac{t-\zeta}{\tau_c}$ such that

$$\begin{aligned}&\bigg\|\Big[\mathbf{e}_c^{(n-1)}(t)^{\mathrm{T}} \cdots \mathbf{e}_c(t)^{\mathrm{T}}\Big]^{\mathrm{T}} - \exp\Big(-\tfrac{t-t_0}{\tau_c}\Big)\mathbf{K}_c\Big(\tfrac{t-t_0}{\tau_c}, \tau_c\Big)\Big[\mathbf{e}_c^{(n-1)}(t_0)^{\mathrm{T}} \cdots \mathbf{e}_c(t_0)^{\mathrm{T}}\Big]^{\mathrm{T}}\bigg\| \\ &\leq \int_{t_0}^{t} \exp\Big(-\tfrac{t-\zeta}{\tau_c}\Big)\Big\|\mathbf{K}_c\Big(\tfrac{t-\zeta}{\tau_c}, \tau_c\Big)\Big\| c_s(\Delta t/\tau_c)\|\boldsymbol{\delta}_s\|\mathrm{d}\zeta,\end{aligned} \tag{72}$$

where $\mathbf{e}_c(t) := \mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)$ and t_0 is the previous time instant with discontinuity of $\mathbf{r}^{(n-1)}(t)$.

Proof. Again, (2)–(5), (6), (13), (14), (15), (29), (49), (50) are valid for the system and the controller, which implies that

$$\tilde{\mathbf{y}}_u^{(n)}(t) = \mathbf{w}_c^{(n)}(t) + \mathbf{s}_u\big(\mathbf{x}(t)\big) + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t) + \Big(\mathbf{B}_a\big(\mathbf{y}(t)\big) - \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big)\tilde{\mathbf{u}}(t), \tag{73}$$

where

$$\begin{aligned}\tilde{\mathbf{u}}(t) &= \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)^{-1}\Big(\alpha\tau_c^{-n}\big(\mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)\big) + \textstyle\sum_{p=0}^{n-1}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \hat{\mathbf{y}}_c^{(p)}(t)\big) - \hat{\mathbf{y}}_u^{(n)}(t) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big) \\ &= \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)^{-1}\Big(\textstyle\sum_{p=0}^{n-1}\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \tilde{\mathbf{y}}_c^{(p)}(t)\big) - \hat{\mathbf{y}}_u^{(*)}(t) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big),\end{aligned} \tag{74}$$

with

$$\hat{\mathbf{y}}_u^{(*)}(t) := \textstyle\sum_{p=0}^{n-1}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(\hat{\mathbf{y}}_c^{(p)}(t) - \tilde{\mathbf{y}}_c^{(p)}(t)\big) + \hat{\mathbf{y}}_u^{(n)}(t). \tag{75}$$

The previous equations indicate the existence of a linear time-variant relation between $\tilde{\mathbf{y}}_u^{(n)}(t)$ and $\hat{\mathbf{y}}_u^{(*)}(t)$ given by

$$\tilde{\mathbf{y}}_u^{(n)}(t) = -\mathbf{K}_a(t)\hat{\mathbf{y}}_u^{(*)}(t) + \mathbf{u}_u(t), \tag{76}$$

with

$$\begin{aligned}\mathbf{u}_u(t) :=\ &\mathbf{w}_c^{(n)}(t) + \mathbf{s}_u\big(\mathbf{x}(t)\big) + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t) \\ &+ \mathbf{K}_a(t)\Big(\textstyle\sum_{p=0}^{n-1}\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \tilde{\mathbf{y}}_c^{(p)}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big).\end{aligned} \tag{77}$$

To show the boundedness of the variations of $\hat{\mathbf{y}}_u^{(*)}(t)$, $\tilde{\mathbf{y}}_u^{(n)}(t)$, and $\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\tilde{\mathbf{u}}(t)$, we use the circle criterion in a similar way to previous work[32]. To this end, we note that, from the definition of k_a and the singular value decomposition $\mathbf{K}_a(t) =$

$\mathbf{U}(t)\boldsymbol{\Sigma}(t)\mathbf{V}(t)^{\mathrm{T}}$, where $\mathbf{U}(t)$ and $\mathbf{V}(t)$ are matrices with orthogonal columns and $\boldsymbol{\Sigma}(t)$ is a diagonal matrix of singular values,

$$\hat{\mathbf{y}}_u^{(*)}(t)^{\mathrm{T}}\mathbf{K}_a(t)^{\mathrm{T}}\mathbf{K}_a(t)\hat{\mathbf{y}}_u^{(*)}(t) = \hat{\mathbf{y}}_u^{(*)}(t)^{\mathrm{T}}\mathbf{V}(t)\boldsymbol{\Sigma}(t)\boldsymbol{\Sigma}(t)\mathbf{V}(t)^{\mathrm{T}}\hat{\mathbf{y}}_u^{(*)}(t) \leq \hat{\mathbf{y}}_u^{(*)}(t)^{\mathrm{T}}\mathbf{V}(t)\big(k_a^2\mathbf{I}\big)\mathbf{V}(t)^{\mathrm{T}}\hat{\mathbf{y}}_u^{(*)}(t) = k_a^2\hat{\mathbf{y}}_u^{(*)}(t)^{\mathrm{T}}\hat{\mathbf{y}}_u^{(*)}(t). \quad (78)$$

Then, it is possible to show that the negative feedback loop with input $\mathbf{u}_u(t)$ that consists in the connection of the transfer function $G_e(s)\mathbf{I}_{n_c}$ and the time-variant gain $\mathbf{K}_a(t)$ is input-to-state stable if $\|G_e(s)\|_\infty k_a < 1$ [40]. Since the latter condition is satisfied by assumption, the variations of $\hat{\mathbf{y}}_u^{(*)}(t)$, $\tilde{\mathbf{y}}_u^{(n)}(t)$, and $\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\tilde{\mathbf{u}}(t)$ are bounded if the variations of $\mathbf{u}_u(t)$ are bounded.

One can then define

$$\Delta\mathbf{w}_c^{(n)}(t-\Delta t+\tau) = \mathbf{w}_c^{(n)}(t) - \mathbf{w}_c^{(n)}(t-\Delta t+\tau), \quad (79)$$

$$\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau) = \frac{\int_\tau^{\Delta t}\frac{\partial\mathbf{s}_u}{\partial\mathbf{x}}\big(\mathbf{x}(t-\Delta t+\zeta)\big)\dot{\mathbf{x}}(t-\Delta t+\zeta)\mathrm{d}\zeta}{\Delta t}, \quad (80)$$

$$\begin{aligned}\Delta\mathbf{y}_a^{(n)}(t-\Delta t+\tau) = {}& \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t) + \mathbf{K}_a(t)\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\tilde{\mathbf{u}}(t)\\ &- \boldsymbol{\beta}_a\big(\mathbf{y}(t-\Delta t+\tau)\big) + \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t-\Delta t+\tau)\big) - \mathbf{B}_a\big(\mathbf{y}(t-\Delta t+\tau)\big)\mathbf{d}(t-\Delta t+\tau)\\ &- \mathbf{K}_a(t-\Delta t+\tau)\tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t-\Delta t+\tau)\big)\tilde{\mathbf{u}}(t-\Delta t+\tau), \end{aligned} \quad (81)$$

for $\tau \in [0, \Delta t]$. This implies that

$$\begin{aligned}\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau) &= \mathbf{w}_c^{(n)}(t-\Delta t+\tau) + \mathbf{s}_u\big(\mathbf{x}(t-\Delta t+\tau)\big)\\ &\quad + \mathbf{s}_a\big(\mathbf{y}(t-\Delta t+\tau), \mathbf{u}(t-\Delta t+\tau)\big) - \tilde{\mathbf{s}}_a\big(\tilde{\mathbf{y}}(t-\Delta t+\tau), \tilde{\mathbf{u}}(t-\Delta t+\tau)\big)\\ &= \mathbf{w}_c^{(n)}(t) - \Delta\mathbf{w}_c^{(n)}(t-\Delta t+\tau) + \mathbf{s}_u\big(\mathbf{x}(t)\big) - \Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau)\\ &\quad + \mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big) - \tilde{\mathbf{s}}_a\big(\tilde{\mathbf{y}}(t), \tilde{\mathbf{u}}(t)\big) - \Delta\mathbf{y}_a^{(n)}(t-\Delta t+\tau)\\ &= \tilde{\mathbf{y}}_u^{(n)}(t) - \Delta\mathbf{w}_c^{(n)}(t-\Delta t+\tau) - \Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau) - \Delta\mathbf{y}_a^{(n)}(t-\Delta t+\tau), \end{aligned} \quad (82)$$

which shows that, if the control of $\tilde{\mathbf{y}}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0,\dots,n-1$, keeps $\dot{\mathbf{x}}(t)$ bounded and the variations of the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$ and $\mathbf{B}_a\big(\mathbf{y}(t-\Delta t+\tau)\big)\mathbf{d}(t-\Delta t+\tau)$ and of the modeling error in the known function $\boldsymbol{\beta}_a\big(\mathbf{y}(t-\Delta t+\tau)\big)$ are bounded, then there exists a bounded vector $\boldsymbol{\delta}_s$ and a scalar $\Delta\bar{t} > 0$ such that, for any sufficiently small window size $\Delta t \leq \Delta\bar{t}$, the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$, the unknown function $\mathbf{s}_u\big(\mathbf{x}(t-\Delta t+\tau)\big)$, and the modeling error in the known function $\mathbf{s}_a\big(\mathbf{y}(t-\Delta t+\tau), \mathbf{u}(t-\Delta t+\tau)\big)$ do not vary more than $\boldsymbol{\delta}_s$ in the interval $\tau \in [0, \Delta t]$, where these variations are given by $\Delta\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$, $\Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau)$, and $\Delta\mathbf{y}_a^{(n)}(t-\Delta t+\tau)$, respectively.

Moreover, from (49), (50), (82), and the definitions of $\mathcal{V}_{p,n,z}$ and $\mathcal{Z}_{p,n,z}$ in Lemma 4, one can observe that

$$\hat{\mathbf{y}}_c^{(p)}(t) = \tilde{\mathbf{y}}_c^{(p)}(t) - \mathcal{Z}_{p,n,\Delta t}(\Delta\mathbf{w}_c^{(n)} + \Delta t\bar{\mathbf{y}}_u^{(n+1)} + \Delta\mathbf{y}_a^{(n)}, t), \quad p = 0,\dots,n-1, \quad (83)$$

$$\hat{\mathbf{y}}_u^{(n)}(t) = \tilde{\mathbf{y}}_u^{(n)}(t) - \mathcal{Z}_{n,n,\Delta t}(\Delta\mathbf{w}_c^{(n)} + \Delta t\bar{\mathbf{y}}_u^{(n+1)} + \Delta\mathbf{y}_a^{(n)}, t). \quad (84)$$

Consequently,

$$\begin{aligned}\dot{\mathbf{y}}_c^{(n-1)}(t) &= \mathbf{y}_u^{(n)}(t) + \mathbf{y}_a^{(n)}(t)\\ &= \mathbf{s}_u\big(\mathbf{x}(t)\big) + \mathbf{s}_a\big(\mathbf{y}(t), \mathbf{u}(t)\big)\\ &= \mathbf{s}_u\big(\mathbf{x}(t)\big) + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{u}(t)\\ &= \mathbf{s}_u\big(\mathbf{x}(t)\big) + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{u}(t) + \tilde{\mathbf{v}}(t) - \hat{\mathbf{y}}_u^{(n)}(t) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) - \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\tilde{\mathbf{u}}(t)\\ &= \mathbf{s}_u\big(\mathbf{x}(t)\big) + \alpha\tau_c^{-n}\big(\mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)\big) + \textstyle\sum_{p=0}^{n-1}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \hat{\mathbf{y}}_c^{(p)}(t)\big) - \hat{\mathbf{y}}_u^{(n)}(t)\\ &\quad + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t) + \Big(\mathbf{B}_a\big(\mathbf{y}(t)\big) - \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big)\tilde{\mathbf{u}}(t)\\ &= \mathbf{s}_u\big(\mathbf{x}(t)\big) + \textstyle\sum_{p=0}^{n-1}\binom{n}{p}\tau_c^{p-n}\big(\mathbf{r}^{(p)}(t) - \tilde{\mathbf{y}}_c^{(p)}(t)\big) - \tilde{\mathbf{y}}_u^{(n)}(t)\\ &\quad + \boldsymbol{\beta}_a\big(\mathbf{y}(t)\big) - \tilde{\boldsymbol{\beta}}_a\big(\tilde{\mathbf{y}}(t)\big) + \mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t) + \Big(\mathbf{B}_a\big(\mathbf{y}(t)\big) - \tilde{\mathbf{B}}_a\big(\tilde{\mathbf{y}}(t)\big)\Big)\tilde{\mathbf{u}}(t)\\ &\quad + \textstyle\sum_{p=0}^{n}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\mathcal{Z}_{p,n,\Delta t}(\Delta\mathbf{w}_c^{(n)} + \Delta t\bar{\mathbf{y}}_u^{(n+1)} + \Delta\mathbf{y}_a^{(n)}, t)\\ &= \textstyle\sum_{p=0}^{n-1}\binom{n}{p}\tau_c^{p-n}\mathbf{e}_c^{(p)}(t) - \mathbf{w}_c^{(n)}(t) - \mathbf{u}_c(t), \end{aligned} \quad (85)$$

where $\mathbf{e}_c(t) := \mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)$ and, from the definition of $\mathcal{Z}_{p,n,z}$ in Lemma 4,

$$\begin{aligned}\mathbf{u}_c(t) &:= -\textstyle\sum_{p=0}^{n}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\mathcal{Z}_{p,n,\Delta t}(\Delta\mathbf{w}_c^{(n)} + \Delta t\bar{\mathbf{y}}_u^{(n+1)} + \Delta\mathbf{y}_a^{(n)}, t)\\ &= -\textstyle\sum_{p=0}^{n}(1-\alpha 0^p)\binom{n}{p}\tau_c^{p-n}\frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\tau)\right)_{p+1}\frac{\Delta\mathbf{w}_c^{(n)}(t-\Delta t+\tau)+\Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau)+\Delta\mathbf{y}_a^{(n)}(t-\Delta t+\tau)}{(-1)^n}\mathrm{d}\tau\\ &= -\textstyle\sum_{p=0}^{n}(1-\alpha 0^p)\binom{n}{p}\left(\frac{\Delta t}{\tau_c}\right)^{n-p}\int_0^1\sum_{i=0}^{n}\frac{\xi^{i+n}p!i!}{(i+n)!}\left(\mathbf{a}_{p,n,\Delta t}^1\right)_{i+1}\frac{\Delta\mathbf{w}_c^{(n)}(t-\Delta t+\Delta t\xi)+\Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\Delta t\xi)+\Delta\mathbf{y}_a^{(n)}(t-\Delta t+\Delta t\xi)}{(-1)^n}\mathrm{d}\xi,\end{aligned} \quad (86)$$

thus, since $\Delta t > 0$ and $\tau_c > 0$, there exists a nondecreasing scalar function $c_s(\Delta t/\tau_c)$ such that

$$\begin{aligned}\|\mathbf{u}_c(t)\| &\leq \textstyle\sum_{p=0}^{n}(1-\alpha 0^p)\binom{n}{p}\left(\frac{\Delta t}{\tau_c}\right)^{n-p}\int_0^1\sum_{i=0}^{n}\frac{\xi^{i+n}p!i!}{(i+n)!}\left|\left(\mathbf{a}_{p,n,\Delta t}^1\right)_{i+1}\right|\mathrm{d}\xi\|\boldsymbol{\delta}_s\|\\ &= c_s(\Delta t/\tau_c)\|\boldsymbol{\delta}_s\|.\end{aligned} \quad (87)$$

Since $\dot{\mathbf{r}}^{(n-1)}(t) = \mathbf{0}_{n_c}$ between time instants with discontinuity of $\mathbf{r}^{(n-1)}(t)$, this results in the minimal realization

$$\dot{\mathbf{x}}_c(t) = \mathbf{A}_c\mathbf{x}_c(t) + \mathbf{B}_c\mathbf{u}_c(t), \quad (88)$$

$$\left[\mathbf{e}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \mathbf{e}_c(t)^{\mathrm{T}}\right]^{\mathrm{T}} = \mathbf{x}_c(t), \quad (89)$$

with

$$\mathbf{A}_c = \begin{bmatrix}\mathbf{A}_{c,n-1} & \cdots & \mathbf{A}_{c,0}\\ \mathbf{I}_{n_c} & \mathbf{0}_{n_c\times(n-2)n_c} & \mathbf{0}_{n_c\times n_c}\\ \mathbf{0}_{(n-2)n_c\times n_c} & \mathbf{I}_{(n-2)n_c} & \mathbf{0}_{(n-2)n_c\times n_c}\end{bmatrix}, \quad \mathbf{A}_{c,p} = -\binom{n}{p}\tau_c^{p-n}\mathbf{I}_{n_c}, \quad p = 0,\ldots,n-1, \quad (90)$$

$$\mathbf{B}_c = \begin{bmatrix}\mathbf{I}_{n_c}\\ \mathbf{0}_{(n-1)n_c\times n_c}\end{bmatrix}. \quad (91)$$

The corresponding response is

$$\left[\mathbf{e}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \mathbf{e}_c(t)^{\mathrm{T}}\right]^{\mathrm{T}} = \exp\left(\mathbf{A}_c(t-t_0)\right)\left[\mathbf{e}_c^{(n-1)}(t_0)^{\mathrm{T}} \ \cdots \ \mathbf{e}_c(t_0)^{\mathrm{T}}\right]^{\mathrm{T}} + \int_{t_0}^{t}\exp\left(\mathbf{A}_c(t-\zeta)\right)\mathbf{B}_c\mathbf{u}_c(\zeta)\mathrm{d}\zeta, \quad (92)$$

where t_0 is the previous time instant with discontinuity of $\mathbf{r}^{(n-1)}(t)$.

Hence, one constructs the decomposition $\mathbf{A}_c\tau_c = \mathbf{V}_c\mathbf{J}_c\mathbf{V}_c^{-1}$, where $\mathbf{V}_c$ is a matrix with the generalized eigenvectors of $\mathbf{A}_c\tau_c$ in its columns and $\mathbf{J}_c$ is the Jordan normal form of $\mathbf{A}_c\tau_c$, with

$$\mathbf{J}_c = \mathbf{I}_{n_c} \otimes \left(-\mathbf{I}_n + \begin{bmatrix}\mathbf{0}_{n-1} & \mathbf{I}_{n-1}\\ 0 & \mathbf{0}_{n-1}^{\mathrm{T}}\end{bmatrix}\right), \quad (93)$$

which implies that

$$\exp\left(\mathbf{A}_c(t-\zeta)\right) = \exp\left(\mathbf{A}_c\tau_c\frac{t-\zeta}{\tau_c}\right) = \exp\left(-\frac{t-\zeta}{\tau_c}\right)\mathbf{K}_c\left(\frac{t-\zeta}{\tau_c}, \tau_c\right), \quad (94)$$

where

$$\mathbf{K}_c\left(\frac{t-\zeta}{\tau_c}, \tau_c\right) = \mathbf{V}_c\left(\mathbf{I}_{n_c} \otimes \left(\textstyle\sum_{p=0}^{n-1}\frac{1}{p!}\left(\frac{t-\zeta}{\tau_c}\right)^p\begin{bmatrix}\mathbf{0}_{n-1} & \mathbf{I}_{n-1}\\ 0 & \mathbf{0}_{n-1}^{\mathrm{T}}\end{bmatrix}^p\right)\right)\mathbf{V}_c^{-1} \quad (95)$$

is a matrix with elements that depend polynomially on $\frac{t-\zeta}{\tau_c}$.

From (87), (92), (94), and the Cauchy-Schwarz inequality, one can then show that (72) holds. This proves that $\tilde{\mathbf{y}}_c^{(p)}(t)$ converge exponentially to a distance linear in $\|\boldsymbol{\delta}_s\|$ from $\mathbf{r}^{(p)}(t)$, for $p = 0,\ldots,n-1$, with time constants τ_c if the window size Δt is sufficiently small such that the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$, the unknown function $\mathbf{s}_u\left(\mathbf{x}(t-\Delta t+\tau)\right)$, and the modeling error in the known function $\mathbf{s}_a\left(\mathbf{y}(t-\Delta t+\tau), \mathbf{u}(t-\Delta t+\tau)\right)$ do not vary more than $\boldsymbol{\delta}_s$ in the interval $\tau \in [0, \Delta t]$. □

Corollary 9. Suppose that the conditions of Theorem 8 are satisfied and the disturbances $\mathbf{w}_c^{(n)}(t)$ and $\mathbf{B}_a\left(\mathbf{y}(t)\right)\mathbf{d}(t)$ and the modeling error in the known function $\boldsymbol{\beta}_a\left(\mathbf{y}(t)\right)$ are absolutely continuous functions of t. Then, as $\Delta t \to 0$, the controller given by (13), (14) with (6), (15), (47), (48) for the system (2)–(5) forces $\tilde{\mathbf{y}}_c^{(p)}(t)$ to converge exponentially to $\mathbf{r}^{(p)}(t)$, for $p = 0,\ldots,n-1$, with time constants τ_c, that is, there exists a matrix $\mathbf{K}_c\left(\frac{t-\zeta}{\tau_c}, \tau_c\right)$ with elements that depend polynomially on $\frac{t-\zeta}{\tau_c}$ such that

$$\left[\mathbf{e}_c^{(n-1)}(t)^{\mathrm{T}} \ \cdots \ \mathbf{e}_c(t)^{\mathrm{T}}\right]^{\mathrm{T}} = \exp\left(-\frac{t-t_0}{\tau_c}\right)\mathbf{K}_c\left(\frac{t-t_0}{\tau_c}, \tau_c\right)\left[\mathbf{e}_c^{(n-1)}(t_0)^{\mathrm{T}} \ \cdots \ \mathbf{e}_c(t_0)^{\mathrm{T}}\right]^{\mathrm{T}}, \quad (96)$$

where $\mathbf{e}_c(t) := \mathbf{r}(t) - \tilde{\mathbf{y}}_c(t)$ and t_0 is the previous time instant with discontinuity of $\mathbf{r}^{(n-1)}(t)$.

Proof. The proof proceeds as for Theorem 8, but this time we note that, since the disturbances $\mathbf{w}_c^{(n)}(t)$ and $\mathbf{B}_a\big(\mathbf{y}(t)\big)\mathbf{d}(t)$ and the modeling error in the known function $\boldsymbol{\beta}_a\big(\mathbf{y}(t)\big)$ are absolutely continuous functions of t, one can define

$$\bar{\mathbf{w}}_c^{(n+1)}(t-\Delta t+\tau) = \frac{\Delta \mathbf{w}_c^{(n)}(t-\Delta t+\tau)}{\Delta t} = \frac{\int_\tau^{\Delta t} \dot{\mathbf{w}}_c^{(n)}(t-\Delta t+\zeta)\mathrm{d}\zeta}{\Delta t}, \tag{97}$$

as well as $\bar{\mathbf{y}}_a^{(n+1)}(t-\Delta t+\tau)$ in a similar way, where $\bar{\mathbf{w}}_c^{(n+1)}(t-\Delta t+\tau)$ and $\bar{\mathbf{y}}_a^{(n+1)}(t-\Delta t+\tau)$ are bounded. Then, one can observe that the variations of the disturbances $\mathbf{w}_c^{(n)}(t-\Delta t+\tau)$, of the unknown function $\mathbf{s}_u\big(\mathbf{x}(t-\Delta t+\tau)\big)$, and of the modeling error in the known function $\mathbf{s}_a\big(\mathbf{y}(t-\Delta t+\tau), \mathbf{u}(t-\Delta t+\tau)\big)$ are given by $\Delta t\bar{\mathbf{w}}_c^{(n+1)}(t-\Delta t+\tau)$, $\Delta t\bar{\mathbf{y}}_u^{(n+1)}(t-\Delta t+\tau)$, and $\Delta t\bar{\mathbf{y}}_a^{(n+1)}(t-\Delta t+\tau)$, respectively. Hence, there exists a vector $\boldsymbol{\delta}_s \to \mathbf{0}_{n_c}$ that bounds these variations as $\Delta t \to 0$, which concludes the proof. ☐

Theorem 8 and Corollary 9 imply that the exponential convergence of $\tilde{\mathbf{y}}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0, \ldots, n-1$, with time constants τ_c is a general feature of the controller given by (13), (14) with (6), (15), (47), (48) for some sufficiently small Δt. However, note that, to ensure this convergence, one must verify that controlling $\tilde{\mathbf{y}}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0, \ldots, n-1$, keeps $\dot{\mathbf{x}}(t)$ bounded. This is similar to the condition of stable zero dynamics that is typically required for the successful implementation of input-output feedback linearization in the case of fully known dynamic models. However, with the proposed approach, this successful implementation of feedback linearization is possible under the same condition despite the fact that only an incomplete dynamic model is used.

Since the values of Δt and τ_c needed to guarantee fast exponential convergence to the setpoints according to Theorem 8 and Corollary 9 are typically too small to be practically useful due to the consequent amplification of measurement noise, the question becomes how much one can increase the values of Δt and τ_c without compromising the stability and performance of the closed-loop system. Moreover, in general it may be difficult to verify that controlling $\tilde{\mathbf{y}}_c^{(p)}(t)$ to $\mathbf{r}^{(p)}(t)$, for $p = 0, \ldots, n-1$, keeps $\dot{\mathbf{x}}(t)$ bounded. These challenges are dealt with in the next subsection.

4.2 | Stability and performance

We now address another goal of this paper, which is the assessment of the local stability and performance of the control scheme. Although several other criteria exist, the performance is assessed here only in terms of the speed of the closed-loop response. The concept of eigenvalues is used for this analysis. To this end, the plant and the controller are linearized to allow expressing the open-loop and closed-loop systems as linear state-space representations, which are accurate approximations of the true nonlinear systems if the deviations are not too large.

Although local stability and performance analysis via eigenvalues is a basic and standard concept in control theory, it is appropriate to present it in this section because this analysis is not straightforward for the proposed control approach and is not available in the literature since the proposed control approach has not been proposed before, to the best of our knowledge. The presented conditions for closed-loop stability and performance analysis can then be readily applied for any equilibrium point of a system. In contrast, a Lyapunov-based analysis, which is a particular type of time-domain stability and performance analysis, would require the verification of existence of a Lyapunov function for each equilibrium point of the closed-loop system[10]. However, this verification would have to be performed for all the trajectories in some invariant set, which may be more difficult.

While it was assumed for control design that a part of the model is fully unknown, it is assumed for stability and performance analysis that the uncertainty in the knowledge of $\mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ is described by n_θ unknown plant parameters $\boldsymbol{\theta} := (\theta_1, \ldots, \theta_{n_\theta})$ that can take any value in a compact set $\boldsymbol{\Theta}$. Although it may seem that only parametric uncertainty in the knowledge of $\mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ can be handled, in fact structural mismatch between the true nonlinear system and its model can be handled if we suppose that the plant parameters $\boldsymbol{\theta}$ correspond to the uncertain elements of the Jacobian matrices $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\big(\bar{\mathbf{x}}, \bar{\mathbf{u}}\big)$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\big(\bar{\mathbf{x}}, \bar{\mathbf{u}}\big)$.

The state-space representation of the plant with the system outputs $\mathbf{y}(t)$ as outputs, the system inputs $\mathbf{u}(t)$ as inputs, and the states $\mathbf{x}(t)$ is given by the matrices $\mathbf{A}_p, \mathbf{B}_p, \mathbf{C}_p$. The state-space representation of the controller with the actuator inputs $\tilde{\mathbf{u}}(t)$ as outputs, the sensor outputs $\tilde{\mathbf{y}}(t)$ and the setpoints $\mathbf{r}(t)$ as inputs, and the states $\mathbf{z}(t)$ is given by the matrices $\mathbf{A}_k, \mathbf{B}_k^y, \mathbf{B}_k^{r,p}, \mathbf{C}_k, \mathbf{D}_k^y, \mathbf{D}_k^{r,p}$, for $p = 0, \ldots, n-1$. These state-space representations depend on the unknown plant parameters $\boldsymbol{\theta}$ and the control design parameters Δt and τ_c. The open-loop transfer function of the plant is computed as $\mathbf{G}_p(s)$, and the open-loop transfer functions of the controller with respect to the sensor outputs $\tilde{\mathbf{y}}(t)$ (superscript y) and the setpoints $\mathbf{r}(t)$ (r) are computed as $\mathbf{G}_k^y(s)$ and $\mathbf{G}_k^r(s)$. Appendix G details the state-space representations and the open-loop transfer functions of the plant and the controller.

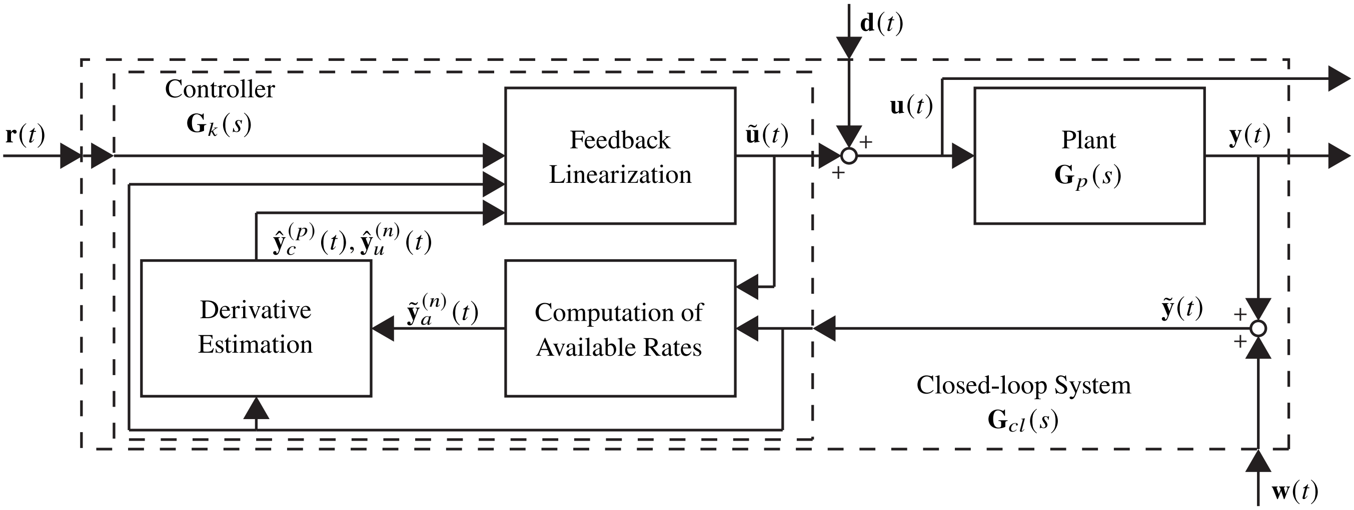

FIGURE 1 Schematic of the closed-loop system that results from multivariable control based on incomplete models via feedback linearization and continuous-time derivative estimation.

Also, we need to consider the following relations between the inputs and outputs of the plant and controller:

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{d}(t), \tag{98a}$$

$$\tilde{\mathbf{y}}(t) = \mathbf{y}(t) + \mathbf{w}(t). \tag{98b}$$

Figure 1 shows a schematic of the closed-loop system. For the stability and performance analysis, we aim to explicitly obtain the state-space representation of the closed-loop system with the outputs $\mathbf{y}_{cl}(t) := \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix}$, the inputs $\mathbf{u}_{cl}(t) := \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{d}(t) \\ \mathbf{w}(t) \end{bmatrix}$, and the states $\mathbf{x}_{cl}(t) := \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{z}(t) \end{bmatrix}$, given by the matrices $\mathbf{A}_{cl}$, $\mathbf{B}_{cl}^{p}$, $\mathbf{C}_{cl}$, $\mathbf{D}_{cl}^{p}$, for $p = 0, \ldots, n-1$. This results in the closed-loop transfer functions of the plant outputs $\mathbf{y}(t)$ (superscript y) and inputs $\mathbf{u}(t)$ (u) with respect to the setpoints $\mathbf{r}(t)$ (r), the input disturbances $\mathbf{d}(t)$ (d), and the output disturbances $\mathbf{w}(t)$ (w) as $\mathbf{G}_{cl}^{r\to y}(s)$, $\mathbf{G}_{cl}^{d\to y}(s)$, $\mathbf{G}_{cl}^{w\to y}(s)$, $\mathbf{G}_{cl}^{r\to u}(s)$, $\mathbf{G}_{cl}^{d\to u}(s)$, $\mathbf{G}_{cl}^{w\to u}(s)$. These transfer functions are the so-called "gang of six" that fully specifies the closed-loop behavior and are also detailed in Appendix G.

To analyze the stability and performance of the closed-loop system for an equilibrium point, it is necessary to note that $\mathbf{A}_{cl}$, $\mathbf{B}_{cl}^{p}$, $\mathbf{C}_{cl}$, $\mathbf{D}_{cl}^{p}$, for $p = 0, \ldots, n-1$ depend not only on the design parameters Δt and τ_c but also on the unknown plant parameters $\boldsymbol{\theta}$ in the set $\boldsymbol{\Theta}$. Then, one can use the eigenvalues of $\mathbf{A}_{cl}$ to assess the robust local stability and some of those eigenvalues to assess the robust local performance of the closed-loop system for an equilibrium point, as stated by the following theorem.

Theorem 10. Consider the matrices

$$\mathbf{A}_{cl} = \begin{bmatrix} \mathbf{A}_p + \mathbf{B}_p\mathbf{D}_k^{y}\mathbf{C}_p & \mathbf{B}_p\mathbf{C}_k \\ \mathbf{B}_k^{y}\mathbf{C}_p & \mathbf{A}_k \end{bmatrix}, \tag{99a}$$

$$\mathbf{B}_{cl}^{p} = \begin{bmatrix} \mathbf{B}_{cl}^{r,p} & \mathbf{B}_{cl}^{d,p} & \mathbf{B}_{cl}^{w,p} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_p\mathbf{D}_k^{r,p} & 0^p\mathbf{B}_p & 0^p\mathbf{B}_p\mathbf{D}_k^{y} \\ \mathbf{B}_k^{r,p} & \mathbf{0}_{n_z\times n_u} & 0^p\mathbf{B}_k^{y} \end{bmatrix}, \quad p = 0, \ldots, n-1, \tag{99b}$$

$$\mathbf{C}_{cl} = \begin{bmatrix} \mathbf{C}_{cl}^{y} \\ \mathbf{C}_{cl}^{u} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_p & \mathbf{0}_{n_y\times n_z} \\ \mathbf{D}_k^{y}\mathbf{C}_p & \mathbf{C}_k \end{bmatrix}, \tag{99c}$$

$$\mathbf{D}_{cl}^{p} = \begin{bmatrix} \mathbf{0}_{n_y\times n_u} & \mathbf{0}_{n_y\times n_u} & \mathbf{0}_{n_y\times n_y} \\ \mathbf{D}_k^{r,p} & 0^p\mathbf{I}_{n_u} & 0^p\mathbf{D}_k^{y} \end{bmatrix}, \quad p = 0, \ldots, n-1, \tag{99d}$$

with $\mathbf{A}_p$, $\mathbf{B}_p$, $\mathbf{C}_p$ in (G51) and $\mathbf{A}_k$, $\mathbf{B}_k^{y}$, $\mathbf{B}_k^{r,p}$, $\mathbf{C}_k$, $\mathbf{D}_k^{y}$, $\mathbf{D}_k^{r,p}$ in (G66), and construct the decomposition $\mathbf{A}_{cl} = \mathbf{V}_{cl}\mathbf{J}_{cl}\mathbf{V}_{cl}^{-1}$, where $\mathbf{V}_{cl}$ is a matrix with the generalized eigenvectors of $\mathbf{A}_{cl}$ in its columns and $\mathbf{J}_{cl}$ is the Jordan normal form of $\mathbf{A}_{cl}$.

For all the plant states $\mathbf{x}(t)$ in a neighborhood of the equilibrium point $\bar{\mathbf{x}}$ and for all the unknown plant parameters $\boldsymbol{\theta}$ in the set $\boldsymbol{\Theta}$, the closed-loop system is locally asymptotically stable if and only if

$$\lambda_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta}) < 0, \qquad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, \tag{100}$$

where $\lambda_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ is the maximum real part of the eigenvalues of the matrix $\mathbf{A}_{cl}$, or equivalently the maximum real part of the diagonal elements of $\mathbf{J}_{cl}$, and all the closed-loop time constants are less than γ^{-1} if and only if

$$s_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta}) < -\gamma, \qquad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, \tag{101}$$

with $\gamma > 0$, where $s_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ is the maximum real part of the diagonal elements of $\mathbf{J}_{cl}$ that correspond to nonzero columns of $\mathbf{C}_{cl}\mathbf{V}_{cl}$, since the maximum closed-loop time constant $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta}) = -s_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})^{-1} < \gamma^{-1}$.

Proof. From (98), (G50), and (G65), the state-space representation of the closed-loop system in a neighborhood of $\bar{\mathbf{x}}$ is

$$\delta \dot{\mathbf{x}}_{cl}(t) = \mathbf{A}_{cl}\delta \mathbf{x}_{cl}(t) + \textstyle\sum_{p=0}^{n-1} \mathbf{B}_{cl}^{p} \delta \mathbf{u}_{cl}^{(p)}(t), \tag{102a}$$

$$\delta \mathbf{y}_{cl}(t) = \mathbf{C}_{cl}\delta \mathbf{x}_{cl}(t) + \textstyle\sum_{p=0}^{n-1} \mathbf{D}_{cl}^{p} \delta \mathbf{u}_{cl}^{(p)}(t), \tag{102b}$$

with $\mathbf{A}_{cl}$, $\mathbf{B}_{cl}^{p}$, $\mathbf{C}_{cl}$, $\mathbf{D}_{cl}^{p}$ in (99).

If the real parts of the eigenvalues of $\mathbf{A}_{cl}$ are less than a value $-\gamma$, with $\gamma > 0$, then the closed-loop system is locally asymptotically stable and all the closed-loop time constants are less than γ^{-1}. However, this might be more conservative than needed because some eigenvalues of $\mathbf{A}_{cl}$ might be related to unobservable closed-loop time constants. In that case, one constructs the decomposition $\mathbf{A}_{cl} = \mathbf{V}_{cl}\mathbf{J}_{cl}\mathbf{V}_{cl}^{-1}$, where $\mathbf{V}_{cl}$ is a matrix with the generalized eigenvectors of $\mathbf{A}_{cl}$ in its columns and $\mathbf{J}_{cl}$ is the Jordan normal form of $\mathbf{A}_{cl}$. The observable closed-loop time constants are related to the diagonal elements of $\mathbf{J}_{cl}$ that correspond to nonzero columns of $\mathbf{C}_{cl}\mathbf{V}_{cl}$. The theorem then follows from standard control theory [10]. □

Remark 6. Note that the conditions (100) and (101) in Theorem 10 are nonconvex in general but can be verified easily if the number of parameters $\boldsymbol{\theta}$ is not too large. Moreover, results from robust control theory can be applied to transform (100) and (101) into conditions that can be verified in a more tractable way. However, this reformulation is considered out of the scope of this paper and is not further analyzed here.

4.3 | Practical considerations and challenges

As shown throughout this paper, the proposed control strategy via feedback linearization and derivative estimation achieves satisfactory closed-loop performance in the presence of structural mismatch between true nonlinear systems and their models. In addition, the proposed control strategy is explicitly constructed for multivariable control and is computationally efficient since it does not require solving an optimization problem at each sampling point. These are important practical advantages of the proposed control strategy. On the other hand, due to the absence of any optimization-based component, the proposed control strategy cannot guarantee constraint satisfaction. Furthermore, the main practical challenge faced by the proposed control strategy is the handling of measurement noise, even though the proposed method for continuous-time derivative estimation minimizes the effect of measurement noise. In fact, the proposed control strategy is expected to perform best when frequent and precise output measurements $\tilde{\mathbf{y}}(t)$ are available. This is not surprising since the measured data are used, in a certain sense, to compensate the absence of an accurate model. From this perspective, the proposed control strategy can be regarded as a data-driven control strategy. Since the proposed control strategy is particularly useful for fast sampling applications in which it makes sense to use continuous-time derivative estimation, the requirement of frequent measurements does not seem to be a major issue.

Another important practical consideration concerns the information that the user needs to supply before applying the proposed control strategy to a given control problem. Upon analyzing the controller given by (13), (14) with (6), (15), (47), (48), it is possible to observe that, besides the output data $\tilde{\mathbf{y}}(t)$ and the derivatives $\mathbf{r}^{(p)}(t)$ of the setpoints, for $p = 0, \ldots, n-1$, the user needs to supply the partially known functions $\tilde{\boldsymbol{\beta}}_a(\mathbf{y}(t))$ and $\tilde{\mathbf{B}}_a(\mathbf{y}(t))$ and the two design parameters Δt and τ_c. Regarding the partially known functions $\tilde{\boldsymbol{\beta}}_a(\mathbf{y}(t))$ and $\tilde{\mathbf{B}}_a(\mathbf{y}(t))$, they can be obtained from a coarse control-relevant model of the form in (6) for $\dot{\mathbf{y}}_c^{(n-1)}(t)$. Note that this coarse model does not need to satisfy strict accuracy requirements. Indeed, according to Theorem 8, the only condition that needs to be satisfied is $k_a < \frac{1}{\|G_e(s)\|_\infty}$, with $k_a := \sup_t \bar{\sigma}\left(\left(\mathbf{B}_a(\mathbf{y}(t)) - \tilde{\mathbf{B}}_a(\tilde{\mathbf{y}}(t))\right)\tilde{\mathbf{B}}_a(\tilde{\mathbf{y}}(t))^{-1}\right)$ and the definition of $\|G_e(s)\|_\infty$ in Theorem 8. Hence, $\tilde{\mathbf{B}}_a(\tilde{\mathbf{y}}(t))$ only needs to be a coarse approximation of the matrix $\mathbf{B}_a(\mathbf{y}(t))$ that satisfies (2)–(5) for the true dynamical system.

Regarding the two design parameters Δt and τ_c, as mentioned in Section 4.1, they should be chosen so as to represent a sufficiently small window size in which the unknown function $\mathbf{s}_u(\mathbf{x}(t))$ is approximately constant and a sufficiently small dominant closed-loop time constant, respectively. This apparently suggests that it is always advantageous to choose Δt and τ_c as small as possible. However, as also mentioned in Section 4.1, if these parameters are too small, this may amplify the effect of measurement noise. Hence, as a practical guideline, if the effect of measurement noise is deemed to be excessive after choosing

Δt and τ_c and applying the proposed control strategy to the given control problem, the user may assess which term in the control law (13) causes that effect: if it is $\hat{\mathbf{y}}_u^{(n)}(t)$, it may be necessary to increase Δt; if it is $\tilde{\mathbf{v}}(t)$, it may be necessary to increase τ_c. In this design process, the stability and performance analysis in Section 4.2 can be performed to obtain additional guarantees that the chosen design parameters are acceptable. This design process is potentially more intuitive than the design of multi-loop PID controllers for multivariable systems, in which the choice of design parameters may not be obvious without using dedicated computational tools even in the case of relatively simple models, due to the lack of clear meaning of the design parameters[41].

5 | ILLUSTRATIVE EXAMPLE

In models of chemical and biological process systems, it is known that (i) one part is well known owing to its macroscopic nature, for instance the flows of material and energy between different units, and (ii) another part is unknown due to the fact that it results from microscopic processes, such as the reaction kinetics and transport phenomena. For example, regarding chemical and biological reaction systems, efficient control typically requires good kinetic models to predict the dynamic effects, namely the reaction rates. In the particular case of continuous stirred-tank reactors, that is, perfectly mixed reactors with constant volume, various control structures based on reaction variants, inventories, and extensive variables have been proposed[35,37,42,43,44]. However, there does not exist a systematic way of taking advantage of partial model knowledge and multiple measurements of variables that are not directly controlled to simplify the design of multivariable controllers, in particular without the use of a kinetic model.

Hence, in this section, the possibility of controlling chemical and biological reactors without the use of kinetic models is investigated as an illustrative example of the theory in the previous sections. The reaction rates are estimated from concentration and temperature measurements and then used via a feedback-linearization scheme to control the reactor temperature and reactant concentrations by manipulating the amount of heat that is exchanged with the environment and the inlet flowrates in a continuous stirred-tank reactor. This particular example concerns a nonisothermal reactor with 4 species (A, B, C, D) and 2 reactions (A+B→C and 2B→D).

5.1 | System description

The $n_x = 5$ states $\mathbf{x}(t)$ are the heat $Q(t) := V\rho c_p\big(T(t) - T_{ref}\big)$ and the numbers of moles $n_s(t) := Vc_s(t)$ of the species $s = \text{A}, \ldots, \text{D}$. Here, V is the constant volume, $c_s(t)$ is the concentration of the species s, $T(t)$ is the reactor temperature, T_{ref} is the reference temperature, ρ is the constant density, and c_p is the constant specific heat capacity. The $n_u = 3$ inputs $\mathbf{u}(t)$ are the exchanged heat power $q_{ex}(t)$ and the volumetric flowrates of two inlets at the temperature T_{ref}. One inlet is fed with the concentration of A $c_{in,\text{A}}$ and the flowrate $F_\text{A}(t)$ and the other inlet is fed with the concentration of B $c_{in,\text{B}}$ and the flowrate $F_\text{B}(t)$. The outlet flowrate is the sum of the inlet flowrates. Since the $n_y = 3$ outputs and the $n_c = 3$ controlled outputs are $\mathbf{y}(t) = \mathbf{y}_c(t) := [\, Q(t) \;\; n_\text{A}(t) \;\; n_\text{B}(t) \,]^\text{T}$, the relative degree of which is $n = 1$, one can define $\mathbf{g}\big(\mathbf{x}(t)\big) := [\, \mathbf{I}_3 \;\; \mathbf{0}_{3\times 2} \,]\mathbf{x}(t)$ and $\mathbf{S} := \mathbf{I}_3$.

Hence, the dynamic model of the plant can be written as

$$\dot{Q}(t) = -\Delta H_{r,1} r_{v,1}(t) - \Delta H_{r,2} r_{v,2}(t) + q_{ex}(t) - \omega(t)Q(t), \tag{103}$$

$$\dot{n}_\text{A}(t) = -r_{v,1}(t) + c_{in,\text{A}} F_\text{A}(t) - \omega(t) n_\text{A}(t), \tag{104}$$

$$\dot{n}_\text{B}(t) = -r_{v,1}(t) - 2r_{v,2}(t) + c_{in,\text{B}} F_\text{B}(t) - \omega(t) n_\text{B}(t), \tag{105}$$

$$\dot{n}_\text{C}(t) = r_{v,1}(t) - \omega(t) n_\text{C}(t), \tag{106}$$

$$\dot{n}_\text{D}(t) = r_{v,2}(t) - \omega(t) n_\text{D}(t), \tag{107}$$

where $r_{v,i}(t) := V r_i\big(\mathbf{x}(t)\big)$, with $r_i\big(\mathbf{x}(t)\big)$ the rate of the ith reaction, $\Delta H_{r,i}$ is the enthalpy of the ith reaction at T_{ref}, and $\omega(t) := \frac{F_\text{A}(t)+F_\text{B}(t)}{V}$ is the inverse of the residence time.

The part of the dynamic model shown in (103)–(107) is known, which is typically the case. However, the part of the model that concerns the reaction kinetics, that is, the functional relationships $r_i\big(\mathbf{x}(t)\big)$, is unknown, which is common in reactor models. In other words, it is assumed that the kinetic model of the 2 unknown rates

$$r_1(\mathbf{x}(t)) = k_{1,ref} \exp\left(-\frac{E_{a,1}}{RT(t)} + \frac{E_{a,1}}{RT_{ref}}\right) \frac{n_\text{A}(t)}{V} \frac{n_\text{B}(t)}{V}, \tag{108a}$$

$$r_2(\mathbf{x}(t)) = k_{2,ref} \exp\left(-\frac{E_{a,2}}{RT(t)} + \frac{E_{a,2}}{RT_{ref}}\right) \left(\frac{n_\text{B}(t)}{V}\right)^2, \tag{108b}$$

TABLE 1 Plant parameters and operating conditions.

Variable	Value	Units
$k_{1,ref}$	0.53	L mol^{-1} min^{-1}
$k_{2,ref}$	1.28	L mol^{-1} min^{-1}
$E_{a,1}$	20000	J mol^{-1}
$E_{a,2}$	10000	J mol^{-1}
R	8.314	J mol^{-1} K^{-1}
T_{ref}	298.15	K
$-\Delta H_{r,1}$	70	kJ mol^{-1}
$-\Delta H_{r,2}$	100	kJ mol^{-1}
$c_{in,\mathrm{A}}$	2.0	mol L^{-1}
$c_{in,\mathrm{B}}$	1.5	mol L^{-1}
V	500	L
$V\rho c_p$	736.3	kJ K^{-1}

is used only to simulate the system, but not to implement control. In (108), $k_{i,ref}$ is the rate constant of the ith reaction at T_{ref}, and $E_{a,i}$ is the activation energy of the ith reaction, which correspond to the unknown plant parameters $\boldsymbol{\theta}$.

Hence, the terms in (103)–(107) that include $r_{v,i}(t)$ correspond to $\mathbf{s}_u\big(\mathbf{x}(t)\big)$, while the other terms correspond to $\mathbf{s}_a\big(\mathbf{y}(t),\mathbf{u}(t)\big)$. From (103)–(107), one can then observe that

$$\begin{bmatrix} \dot{Q}(t) \\ \dot{n}_\mathrm{A}(t) \\ \dot{n}_\mathrm{B}(t) \end{bmatrix} = \mathbf{s}_u\big(\mathbf{x}(t)\big) + \mathbf{s}_a\big(\mathbf{y}(t),\mathbf{u}(t)\big), \tag{109}$$

where

$$\mathbf{s}_u\big(\mathbf{x}(t)\big) := \begin{bmatrix} -\Delta H_{r,1} & -\Delta H_{r,2} \\ -1 & 0 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} V r_1(\mathbf{x}(t)) \\ V r_2(\mathbf{x}(t)) \end{bmatrix}, \tag{110}$$

$$\mathbf{s}_a\big(\mathbf{y}(t),\mathbf{u}(t)\big) := \begin{bmatrix} 1 & -\frac{Q(t)}{V} & -\frac{Q(t)}{V} \\ 0 & c_{in,\mathrm{A}}-\frac{n_\mathrm{A}(t)}{V} & -\frac{n_\mathrm{A}(t)}{V} \\ 0 & -\frac{n_\mathrm{B}(t)}{V} & c_{in,\mathrm{B}}-\frac{n_\mathrm{B}(t)}{V} \end{bmatrix} \mathbf{u}(t). \tag{111}$$

Hence, since the function $\mathbf{s}_u\big(\mathbf{x}(t)\big)$ on the right-hand side of (109) is unknown, the only way to estimate its value is via estimation of the output derivatives on the left-hand side.

The numerical values used for the plant parameters and operating conditions are summarized in Table 1.

5.2 | Stability and performance

First of all, the stability and performance of the closed-loop system are analyzed by computing the closed-loop response for the nominal values $\tau_c = 3$ min and $\Delta t = 3$ min of the design parameters. These parameters are chosen such that their sum is 5 times less than the desired settling time of 30 min, based on the reasonable assumption that the unknown reaction rates remain approximately constant in a window of $\Delta t = 3$ min. Since the closed-loop system has 9 inputs and 6 outputs, it is clearly inconvenient to show here the closed-loop response to all the inputs of the closed-loop system, but this analysis is performed for the setpoints and input disturbances. Then, the stability and performance are analyzed by computing $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ for values of the design parameters τ_c and Δt around their nominal values, as well as for values of the unknown plant parameters $k_{1,ref}$ and $k_{2,ref}$ around their true values. Note that the values of these plant parameters are not used for controller design.

Figure 2, Figure 3, and Figure 4 show the closed-loop response to step increases of 0.5% of $Q_0 - V\rho c_p(273.15 - T_{ref}) = 18407$ kJ in the setpoint $Q^r(t)$, 0.5% of $n_{\mathrm{A},0} = 204.40$ mol in the setpoint $n_\mathrm{A}^r(t)$, and 0.5% of $n_{\mathrm{B},0} = 51.09$ mol in the setpoint $n_\mathrm{B}^r(t)$. Figure 5, Figure 6, and Figure 7 show the closed-loop response to step increases of 0.5% of $q_{ex,0} = -1443$ kJ min^{-1} in the input disturbance $q_{ex}^d(t)$, 0.5% of $F_{\mathrm{A},0} = 11.67$ L min^{-1} in the input disturbance $F_\mathrm{A}^d(t)$, and 0.5% of $F_{\mathrm{B},0} = 18.33$ L min^{-1} in the input disturbance $F_\mathrm{B}^d(t)$. The initial conditions for this simulation correspond to the steady state in which $\bar{Q} = 0$ kJ, $\bar{F}_\mathrm{A} = 11.67$ L min^{-1}, and $\bar{F}_\mathrm{B} = 18.33$ L min^{-1}. In Figure 4, one can observe that $n_\mathrm{B}(t)$ converges to the new value of $n_\mathrm{B}^r(t)$ and

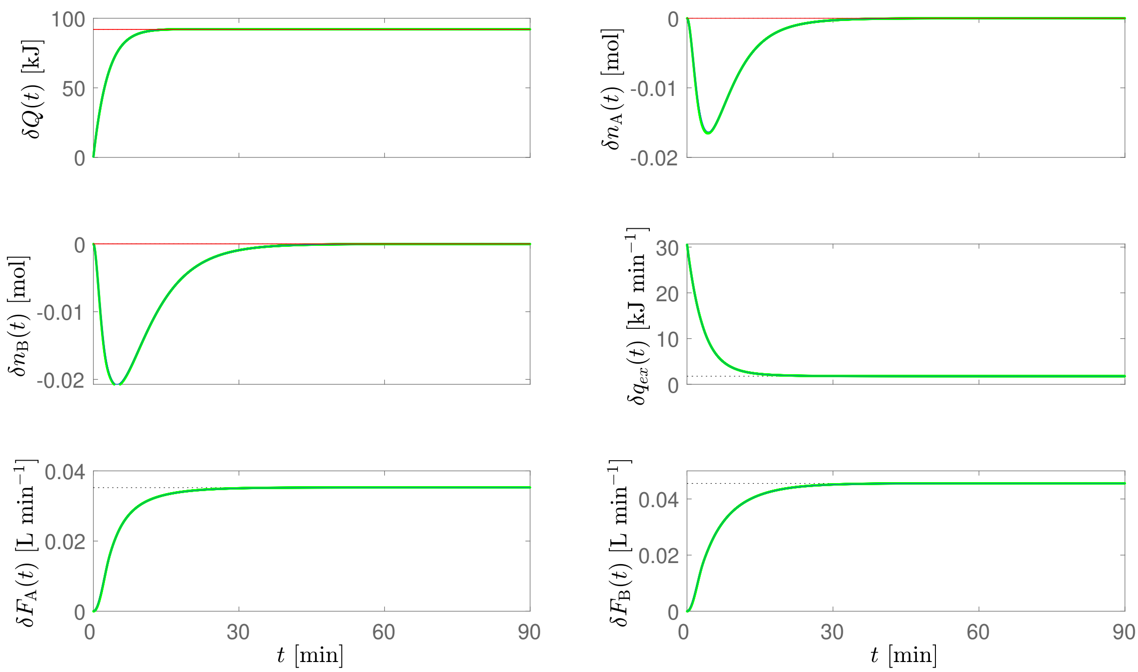

FIGURE 2 Closed-loop response to a step increase in the setpoint $Q^r(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

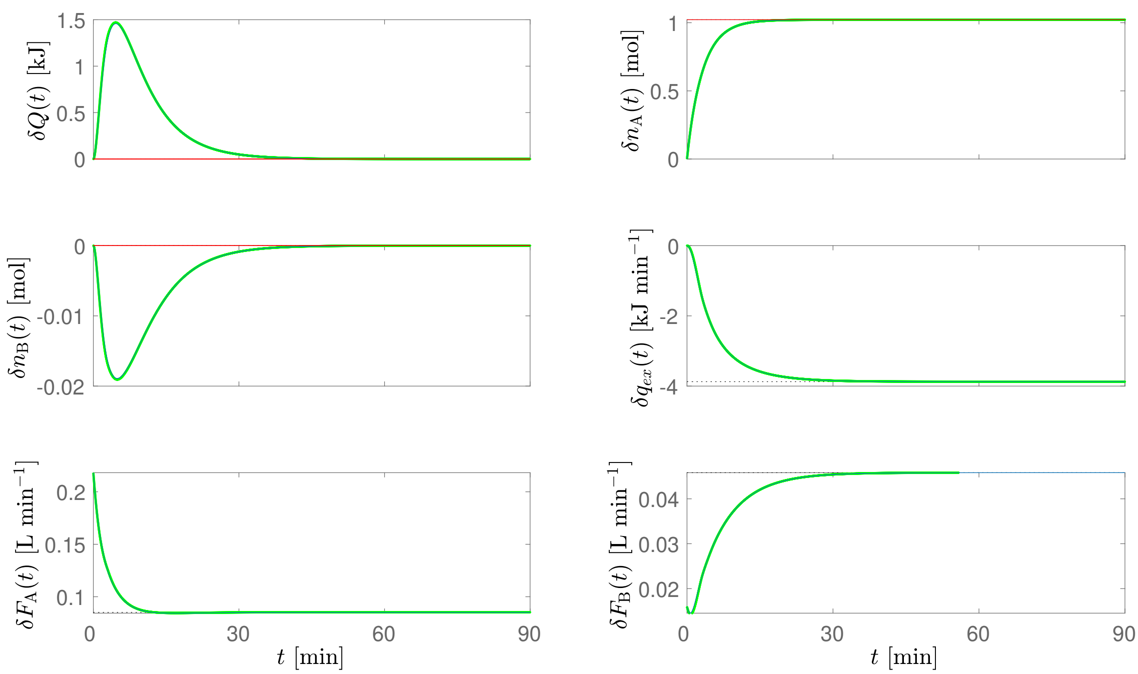

FIGURE 3 Closed-loop response to a step increase in the setpoint $n_{\mathrm{A}}^r(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

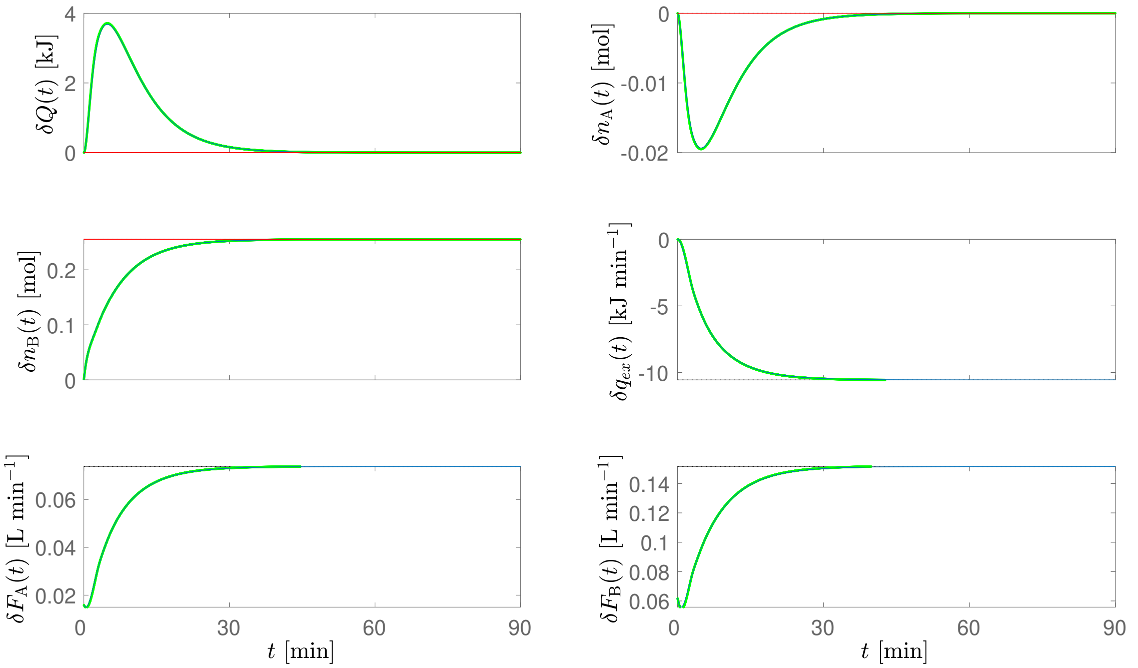

FIGURE 4 Closed-loop response to a step increase in the setpoint $n_{\mathrm{B}}^{r}(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

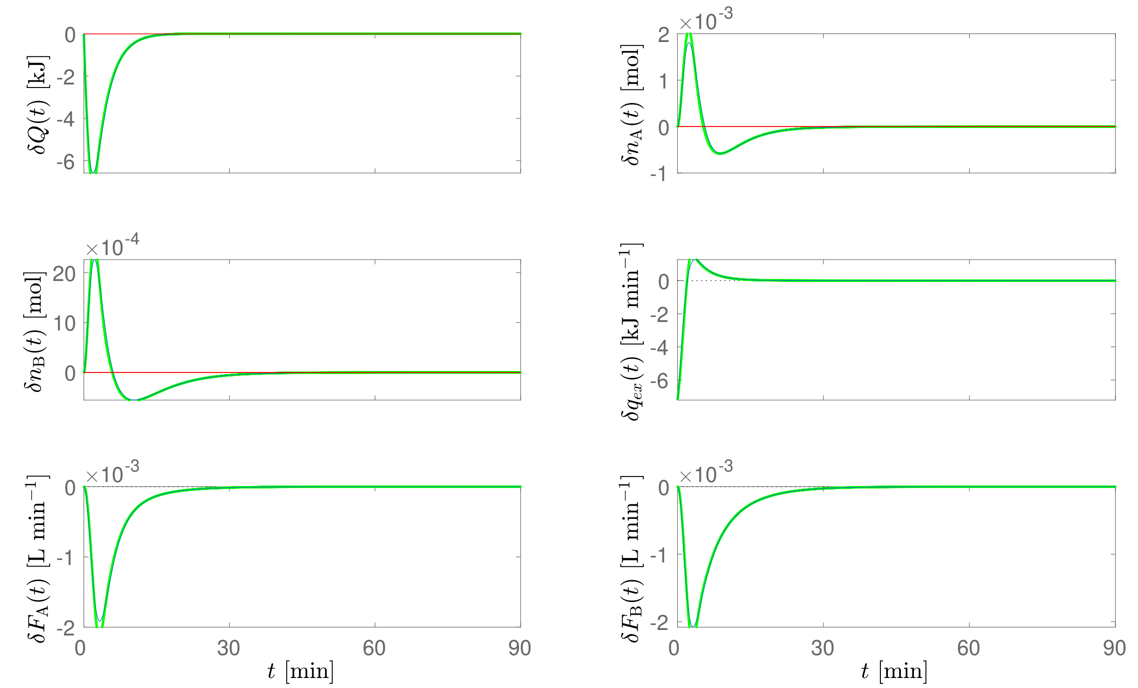

FIGURE 5 Closed-loop response to a step increase in the input disturbance $q_{ex}^{d}(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

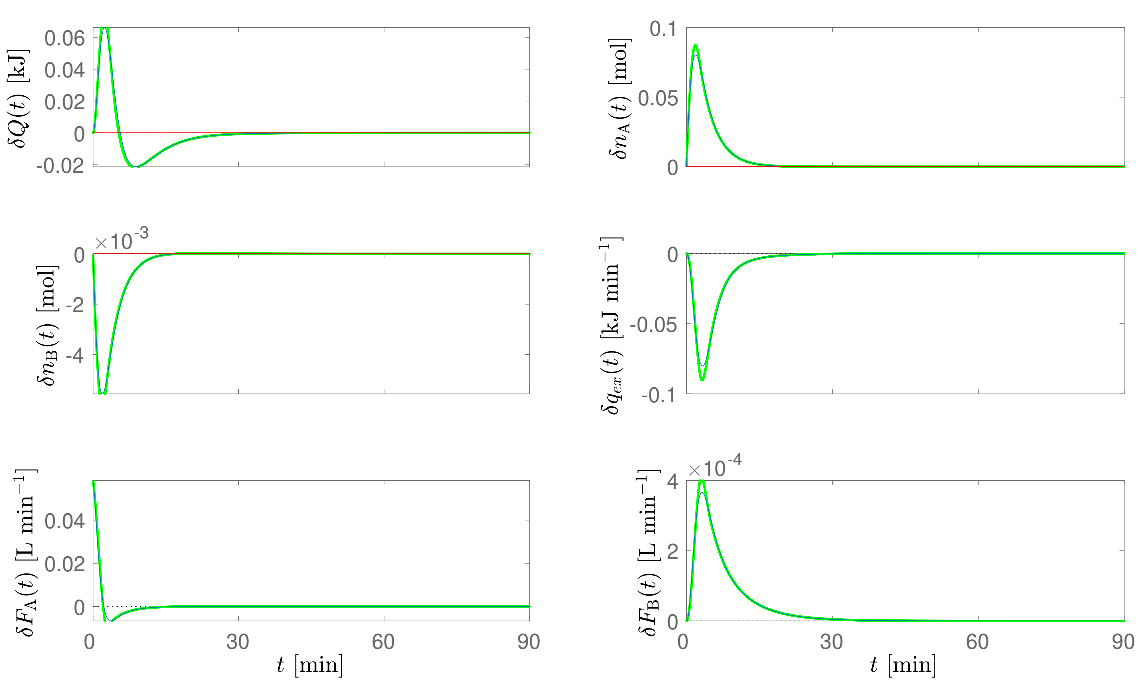

FIGURE 6 Closed-loop response to a step increase in the input disturbance $F_{\rm A}^d(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

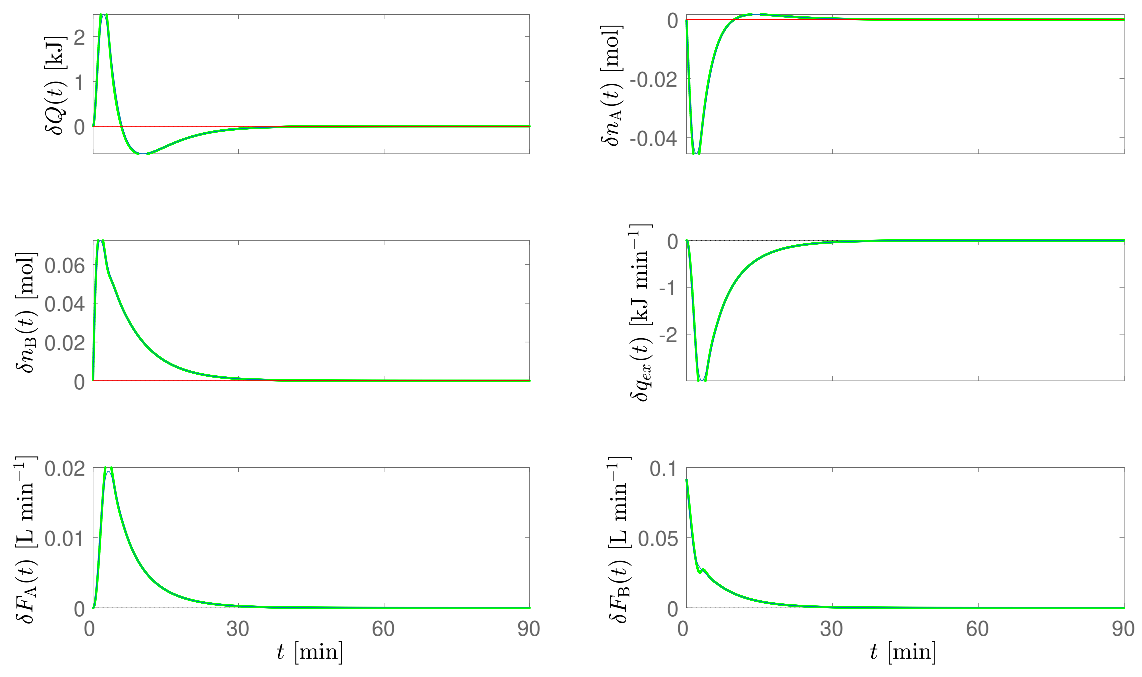

FIGURE 7 Closed-loop response to a step increase in the input disturbance $F_{\rm B}^d(t)$, with the setpoints represented by the red lines. The blue lines represent the response obtained via linearization and the green lines represent the response obtained via numerical simulation of the nonlinear system.

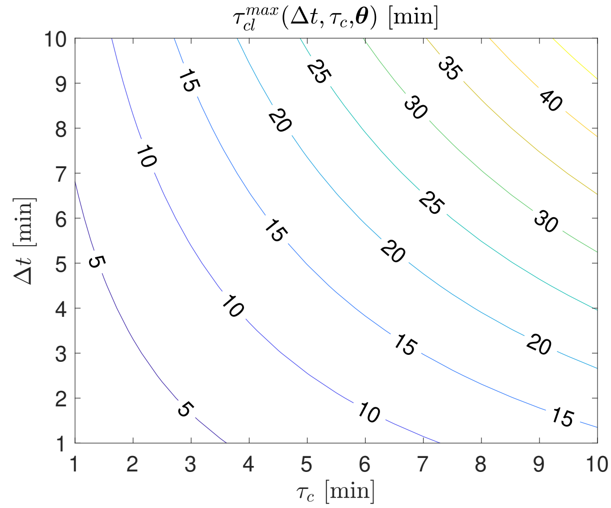

FIGURE 8 Contour plot of the maximum closed-loop time constant as a function of the design parameters τ_c and Δt.

both $Q(t)$ and $n_\mathrm{A}(t)$ return to their initial values within the desired settling time of 30 min, and similar conclusions can be drawn from Figure 2 and Figure 3. In Figure 7, one can observe that $Q(t)$, $n_\mathrm{A}(t)$, and $n_\mathrm{B}(t)$ as well as $q_{ex}(t)$, $F_\mathrm{A}(t)$, and $F_\mathrm{B}(t)$ return to their initial values within the desired settling time of 30 min despite the step increase in $F_\mathrm{B}^d(t)$, and similar conclusions can be drawn from Figure 5 and Figure 6. This elimination of steady-state error and this rejection of input disturbances are obtained despite the fact that the controller does not include any integral term and uses an incomplete plant model. This closed-loop behavior also agrees well with the corresponding maximum time constant $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta}) = 6.8$ min, which is considerably smaller than the maximum time constant of the open-loop system $\tau_p^{max}(\boldsymbol{\theta}) = 16.7$ min. The aforementioned figures also compare the responses obtained via numerical simulation of the nonlinear system and via linearization, which shows that the linearization approximates well the dynamics of the nonlinear system.

Figure 8 shows $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ as a function of values of the design parameters around their nominal values, that is, for $\tau_c \in [1, 10]$ min and $\Delta t \in [1, 10]$ min. Although the performance of the closed-loop system in terms of its maximum time constant becomes worse with increasing values of τ_c and Δt, as expected, the stability of the closed-loop system is not compromised for any values of these parameters. Note that the states $n_s(t)$ for $s = \mathrm{A}, \ldots, \mathrm{D}$ are necessarily bounded owing to the fact that the numbers of moles in a reactor are bounded and the state $Q(t)$ that could become unbounded is controlled in this example, which implies that all the states $\mathbf{x}(t)$ are bounded. This, in turn, entails the boundedness of $\dot{\mathbf{x}}(t)$ and guarantees stability for sufficiently small Δt according to Theorem 8. In fact, one can observe in Figure 8 that, for small values of Δt, the maximum time constant approaches the value of τ_c, as predicted by Theorem 8. Moreover, Figure 9 shows $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ as a function of values of the unknown plant parameters around their true values, that is, for $k_{1,ref} \in [0.2, 2]$ L mol^{-1} min^{-1} and $k_{2,ref} \in [0.4, 4]$ L mol^{-1} min^{-1}. The effect of the unknown plant parameters $E_{a,1}$ and $E_{a,2}$ on $\tau_{cl}^{max}(\Delta t, \tau_c, \boldsymbol{\theta})$ is not shown here since it is much weaker than the effect of $k_{1,ref}$ and $k_{2,ref}$. Again, the performance of the closed-loop system becomes worse with increasing values of $k_{1,ref}$ and $k_{2,ref}$, but its stability is not compromised.

6 | CONCLUSION

This paper has presented a control scheme based on incomplete models, which is implemented without full knowledge of the model owing to the use of continuous-time derivative estimation and takes advantage of the knowledge about part of the model owing to the use of feedback linearization. Although only incomplete knowledge of the model is used and the controller possesses only two design parameters that are rather simple to tune thanks to their clear meaning, the control scheme converges quickly to its setpoints and can eliminate steady-state error and reject constant input disturbances without any integral term. The two design parameters are the time constant of exponential convergence of each controlled output that determines the controller gain and the parameter of the FIR differentiation filter used for estimation of the unknown derivatives without a model. Instead of linearizing the system around a given steady state, this controller implements feedback linearization, which simplifies control

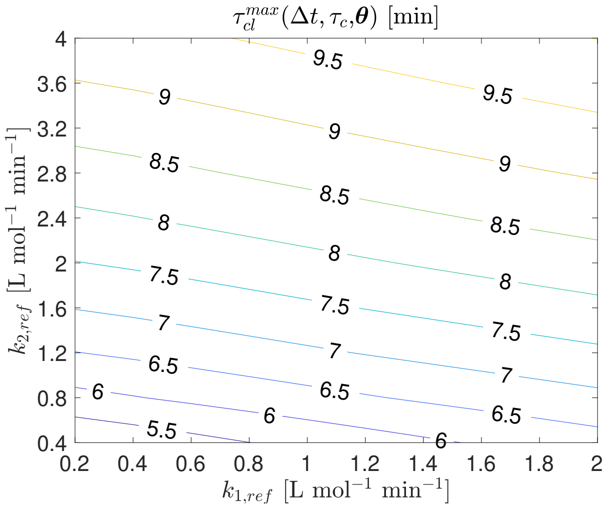

FIGURE 9 Contour plot of the maximum closed-loop time constant as a function of the unknown plant parameters $k_{1,ref}$ and $k_{2,ref}$.

design significantly and enables the use of the controller even when the states are not similar to the corresponding steady-state values. The feedback linearization sets a rate of variation for the controlled outputs and forces the control error to converge exponentially to zero. The resulting multivariable controller shows good performance for the case of frequent measurements of several outputs. Furthermore, quantitative criteria to analyze the robust stability and performance of this control strategy have been provided for the continuous-time case. These features have been illustrated by an example of a realistic control problem, that is, control of temperature and reactant concentrations in a homogeneous reactor without kinetic models for the reaction rates. Besides this illustrative example, the proposed control strategy is broadly applicable to a large variety of systems.

Future work could provide new criteria to assess the stability and performance in a robust sense and extend these results to the discrete-time case. It would also be useful to formulate conditions that express the robust stability and performance criteria in a way that is as explicit and computationally cheap as possible. Finally, it would be rather interesting to investigate how much one can improve the closed-loop performance by using complete models that result from system identification and how this control scheme compares with the use of techniques for robust control design in terms of stability and performance.

ACKNOWLEDGEMENTS

Dominique Bonvin and Håkan Hjalmarsson are gratefully acknowledged for their comments and suggestions, which helped clarify and improve previous versions of the methods in this paper. This work was supported by the Swiss National Science Foundation, project number 184521.

Conflict of interest

The authors declare no potential conflict of interests.

How to cite this article: D. Rodrigues and A. Mesbah (2021), Multivariable Control Based on Incomplete Models via Feedback Linearization and Continuous-Time Derivative Estimation, *Int. J. of Rob. and Non. Contr.*, *?*.

APPENDIX

A PROOF OF LEMMA 1

Firstly, we show that $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p=0,\ldots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ minimize the mean squared error $J_{n,z}(t)$. Upon defining the predicted parameter matrix $\hat{\boldsymbol{\theta}}_{n,z}$ with $n+1$ columns $\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_1,\ldots,\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{n+1}$ as

$$\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{l+1}=\begin{cases}\frac{\Delta t^l\hat{\mathbf{y}}_c^{(l)}(t-\Delta t+z)}{l!}, & l=0,\ldots,n-1\\ \frac{\Delta t^n\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{n!}, & l=n\end{cases}, \tag{A1}$$

and the true parameter matrix $\boldsymbol{\theta}_{n,z}$ with $n+1$ columns $\left(\boldsymbol{\theta}_{n,z}\right)_1,\ldots,\left(\boldsymbol{\theta}_{n,z}\right)_{n+1}$ as

$$\left(\boldsymbol{\theta}_{n,z}\right)_{l+1}=\begin{cases}\frac{\Delta t^l\mathbf{y}_c^{(l)}(t-\Delta t+z)}{l!}, & l=0,\ldots,n-1\\ \frac{\Delta t^n\mathbf{y}_u^{(n)}(t-\Delta t+z)}{n!}, & l=n\end{cases}, \tag{A2}$$

and noticing that (18) holds, it follows that

$$\begin{aligned}\hat{\boldsymbol{\epsilon}}_{n,z}(t-\Delta t+\tau)&=\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)-\textstyle\sum_{l=0}^{n-1}\frac{(\tau-z)^l}{l!}\hat{\mathbf{y}}_c^{(l)}(t-\Delta t+z)-\frac{(\tau-z)^n}{n!}\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)\\&=\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)-\textstyle\sum_{l=0}^{n}\left(\frac{\tau-z}{\Delta t}\right)^l\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{l+1},\end{aligned} \tag{A3}$$

$$\begin{aligned}J_{n,z}(t)&=\textstyle\sum_{l=0}^{n}\left(\sum_{i=0}^{n}\int_0^{\Delta t}\frac{1}{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^{i+l}\mathrm{d}\tau\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{i+1}^{\mathrm{T}}-\int_0^{\Delta t}\frac{2}{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)^{\mathrm{T}}\mathrm{d}\tau\right)\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{l+1}\\&+\frac{1}{\Delta t}\int_0^{\Delta t}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)^{\mathrm{T}}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)\mathrm{d}\tau,\end{aligned} \tag{A4}$$

$$\frac{1}{2}\frac{\partial J_{n,z}(t)}{\partial(\hat{\boldsymbol{\theta}}_{n,z})_{l+1}}=\textstyle\sum_{i=0}^{n}\int_{-\frac{z}{\Delta t}}^{1-\frac{z}{\Delta t}}x^{i+l}\mathrm{d}x\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{i+1}^{\mathrm{T}}-\int_0^{\Delta t}\frac{1}{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)^{\mathrm{T}}\mathrm{d}\tau,\qquad l=0,\ldots,n. \tag{A5}$$

One can also define the $(n+1)$-dimensional vector $\mathbf{d}_{n,z}(\tau)$ as

$$\left(\mathbf{d}_{n,z}(\tau)\right)_{l+1}=\frac{1}{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l,\qquad l=0,\ldots,n, \tag{A6}$$

which implies that

$$\frac{1}{2}\frac{\partial J_{n,z}(t)}{\partial\hat{\boldsymbol{\theta}}_{n,z}}=\mathbf{G}_n^r\hat{\boldsymbol{\theta}}_{n,z}^{\mathrm{T}}-\int_0^{\Delta t}\mathbf{d}_{n,z}(\tau)\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)^{\mathrm{T}}\mathrm{d}\tau, \tag{A7}$$

with $r=1-\frac{z}{\Delta t}$. Since the parameters $\hat{\boldsymbol{\theta}}_{n,z}$ that minimize $J_{n,z}(t)$ are such that $\frac{\partial J_{n,z}(t)}{\partial\hat{\boldsymbol{\theta}}_{n,z}}=\mathbf{0}$ and

$$\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}=\Delta t\textstyle\sum_{i=0}^{n}\left(\mathbf{d}_{n,z}(\tau)\right)_{i+1}\left(\mathbf{A}_n^r\right)_{i+1,p+1},\qquad p=0,\ldots,n, \tag{A8}$$

it follows that

$$\hat{\boldsymbol{\theta}}_{n,z}=\frac{1}{\Delta t}\int_0^{\Delta t}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)\mathbf{c}_{n,z}(\tau)^{\mathrm{T}}\mathrm{d}\tau, \tag{A9}$$

which implies that $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p=0,\ldots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ minimize the mean squared error $J_{n,z}(t)$.

Secondly, we show that $\hat{\mathbf{y}}_c^{(p)}(t-\Delta t+z)$, for $p=0,\ldots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t-\Delta t+z)$ are the linear unbiased estimators of $\mathbf{y}_c^{(p)}(t-\Delta t+z)$, for $p=0,\ldots,n-1$, and $\mathbf{y}_u^{(n)}(t-\Delta t+z)$ that provide minimal variance if Assumptions 1-4 hold. To this end, we assess whether there exists any linear unbiased estimator $\bar{\boldsymbol{\theta}}_{n,z}$ of $\boldsymbol{\theta}_{n,z}$ with smaller variance than $\hat{\boldsymbol{\theta}}_{n,z}$ by expressing $\bar{\boldsymbol{\theta}}_{n,z}$ in terms of $\hat{\boldsymbol{\theta}}_{n,z}$ as

$$\bar{\boldsymbol{\theta}}_{n,z}=\hat{\boldsymbol{\theta}}_{n,z}+\frac{1}{\Delta t}\int_0^{\Delta t}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)\delta\mathbf{c}_{n,z}(\tau)^{\mathrm{T}}\mathrm{d}\tau=\frac{1}{\Delta t}\int_0^{\Delta t}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)\left(\mathbf{c}_{n,z}(\tau)+\delta\mathbf{c}_{n,z}(\tau)\right)^{\mathrm{T}}\mathrm{d}\tau, \tag{A10}$$

and assessing whether any L^2 vector function $\delta\mathbf{c}_{n,z}(\tau)$ improves $\bar{\boldsymbol{\theta}}_{n,z}$ with respect to $\hat{\boldsymbol{\theta}}_{n,z}$. Upon defining the true errors as

$$\boldsymbol{\epsilon}_c(t-\Delta t+\tau)=\mathbf{w}_c(t-\Delta t+\tau)+\int_z^{\tau}\frac{\mathbf{y}_u^{(n)}(t-\Delta t+\zeta)-\mathbf{y}_u^{(n)}(t-\Delta t+z)+\mathbf{y}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta, \tag{A11}$$

and noticing that Assumptions 1 and 2 imply that $\boldsymbol{\epsilon}_c(t-\Delta t+\tau)=\mathbf{w}_c(t-\Delta t+\tau)$, it follows that

$$\begin{aligned}\boldsymbol{\upsilon}_{n,z}(t-\Delta t+\tau)&=\boldsymbol{\epsilon}_c(t-\Delta t+\tau)+\textstyle\sum_{l=0}^{n-1}\frac{(\tau-z)^l}{l!}\mathbf{y}_c^{(l)}(t-\Delta t+z)+\frac{(\tau-z)^n}{n!}\mathbf{y}_u^{(n)}(t-\Delta t+z)\\&=\boldsymbol{\epsilon}_c(t-\Delta t+\tau)+\textstyle\sum_{l=0}^{n}\left(\frac{\tau-z}{\Delta t}\right)^l\left(\boldsymbol{\theta}_{n,z}\right)_{l+1}\\&=\mathbf{w}_c(t-\Delta t+\tau)+\Delta t\boldsymbol{\theta}_{n,z}\mathbf{d}_{n,z}(\tau).\end{aligned} \tag{A12}$$

Moreover, since the definition of $\mathbf{c}_n(\tau)$ implies that

$$\int_0^{\Delta t} \left(\mathbf{d}_{n,z}(\tau)\right)_{l+1} \left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} \mathrm{d}\tau = \sum_{i=0}^{n} \frac{1}{\Delta t}\int_0^{\Delta t} \left(\frac{\tau - z}{\Delta t}\right)^{i+l} \left(\mathbf{A}_n^r\right)_{i+1,p+1} \mathrm{d}\tau = \sum_{i=0}^{n} \left(\mathbf{G}_n^r\right)_{l+1,i+1} \left(\mathbf{A}_n^r\right)_{i+1,p+1}$$
$$= \begin{cases} 0, & l \neq p \\ 1, & l = p \end{cases}, \qquad l = 0,\dots,n, \qquad p = 0,\dots,n, \tag{A13}$$

and Assumption 3 implies that $\mathrm{E}\left[\mathbf{w}_c(t - \Delta t + \tau)\right] = \mathbf{0}_{n_c}$,

$$\begin{aligned}\mathrm{E}\left[\left(\bar{\boldsymbol{\theta}}_{n,z}\right)_{l+1}\right] &= \frac{1}{\Delta t}\int_0^{\Delta t} \left(\mathrm{E}\left[\mathbf{w}_c(t - \Delta t + \tau)\right] + \Delta t \boldsymbol{\theta}_{n,z} \mathbf{d}_{n,z}(\tau)\right) \left(\left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} + \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}\right) \mathrm{d}\tau \\ &= \boldsymbol{\theta}_{n,z}\int_0^{\Delta t} \mathbf{d}_{n,z}(\tau) \left(\left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} + \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}\right) \mathrm{d}\tau \\ &= \left(\boldsymbol{\theta}_{n,z}\right)_{l+1} + \boldsymbol{\theta}_{n,z}\int_0^{\Delta t} \mathbf{d}_{n,z}(\tau) \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1} \mathrm{d}\tau, \qquad l = 0,\dots,n,\end{aligned} \tag{A14}$$

thus one concludes that $\left(\bar{\boldsymbol{\theta}}_{n,z}\right)_{l+1}$ is an linear unbiased estimator of $\left(\boldsymbol{\theta}_{n,z}\right)_{l+1}$ if and only if $\int_0^{\Delta t} \mathbf{d}_{n,z}(\tau) \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1} \mathrm{d}\tau = \mathbf{0}_{n+1}$, which is satisfied by $\left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{l+1}$. Since (24) holds according to Assumption 4 and $\left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} + \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}$ is an L^2 function,

$$\begin{aligned}\mathrm{Var}\left[\left(\bar{\boldsymbol{\theta}}_{n,z}\right)_{l+1}\right] &= \mathrm{Var}\left[\frac{1}{\Delta t}\int_0^{\Delta t} \mathbf{w}_c(t - \Delta t + \tau) \left(\left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} + \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}\right) \mathrm{d}\tau\right] \\ &= \frac{1}{\Delta t}\int_0^{\Delta t} \left(\left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} + \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}\right)^2 \mathrm{d}\tau \boldsymbol{\Sigma}_{\mathbf{w}_c} \\ &= \sum_{i=0}^{n}\int_0^{\Delta t} \left(\mathbf{d}_{n,z}(\tau)\right)_{i+1} \left(\mathbf{c}_{n,z}(\tau)\right)_{l+1} \mathrm{d}\tau \left(\mathbf{A}_n^r\right)_{i+1,l+1} \boldsymbol{\Sigma}_{\mathbf{w}_c} \\ &\quad + 2\sum_{i=0}^{n}\int_0^{\Delta t} \left(\mathbf{d}_{n,z}(\tau)\right)_{i+1} \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1} \mathrm{d}\tau \left(\mathbf{A}_n^r\right)_{i+1,l+1} \boldsymbol{\Sigma}_{\mathbf{w}_c} \\ &\quad + \frac{1}{\Delta t}\int_0^{\Delta t} \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}^2 \mathrm{d}\tau \boldsymbol{\Sigma}_{\mathbf{w}_c} \\ &= \left(\mathbf{A}_n^r\right)_{l+1,l+1} \boldsymbol{\Sigma}_{\mathbf{w}_c} + \frac{1}{\Delta t}\int_0^{\Delta t} \left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1}^2 \mathrm{d}\tau \boldsymbol{\Sigma}_{\mathbf{w}_c}, \qquad l = 0,\dots,n,\end{aligned} \tag{A15}$$

thus one concludes that $\left(\bar{\boldsymbol{\theta}}_{n,z}\right)_{l+1}$ is the linear unbiased estimator of $\left(\boldsymbol{\theta}_{n,z}\right)_{l+1}$ that provides minimal variance if and only if $\left(\delta \mathbf{c}_{n,z}(\tau)\right)_{l+1} = 0$ and $\left(\bar{\boldsymbol{\theta}}_{n,z}\right)_{l+1} = \left(\hat{\boldsymbol{\theta}}_{n,z}\right)_{l+1}$. This implies that $\hat{\mathbf{y}}_c^{(p)}(t - \Delta t + z)$, for $p = 0,\dots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t - \Delta t + z)$ are the linear unbiased estimators of $\mathbf{y}_c^{(p)}(t - \Delta t + z)$, for $p = 0,\dots,n-1$, and $\mathbf{y}_u^{(n)}(t - \Delta t + z)$ that provide minimal variance.

B PROOF OF LEMMA 2

The shifted Legendre polynomials with parameter r and orders k are defined as

$$P_k^r(x) = \sum_{i=0}^{k} \left(\mathbf{P}_n^r\right)_{i+1,k+1} x^i, \qquad k = 0,\dots,n, \tag{B16}$$

for some $(n+1)\times(n+1)$ upper triangular matrix $\mathbf{P}_n^r$, such that they satisfy the conditions

$$\int_{r-1}^{r} P_k^r(x) P_l^r(x) \mathrm{d}x = \begin{cases} \frac{1}{2k+1}, & k = l \\ 0, & k \neq l \end{cases}, \qquad k = 0,\dots,n, \qquad l = 0,\dots,n. \tag{B17}$$

From the definition of shifted Legendre polynomials, the $(n+1)\times(n+1)$ matrix $\mathbf{D}_n$ defined as

$$\left(\mathbf{D}_n\right)_{k+1,l+1} = \int_{r-1}^{r} P_k^r(x) P_l^r(x) \mathrm{d}x, \qquad k = 0,\dots,n, \qquad l = 0,\dots,n, \tag{B18}$$

is diagonal. Some particular examples of shifted Legendre polynomials are

$$P_k^1(x) = \sum_{i=0}^{k} (-1)^{k+i} \binom{k}{i}\binom{k+i}{k} x^i, \qquad k = 0,\dots,n, \tag{B19}$$

$$P_k^{1/2}(x) = \sum_{i=0}^{k} 2^{k+i} \binom{k}{i}\binom{\frac{k+i-1}{2}}{k} x^i, \qquad k = 0,\dots,n, \tag{B20}$$

$$P_k^0(x) = \sum_{i=0}^{k} \binom{k}{i}\binom{k+i}{k} x^i, \qquad k = 0,\dots,n. \tag{B21}$$

In general, for $r \in \{0, 1/2, 1\}$ and all $0 \le s \le 1$,

$$\left(\mathbf{P}_n^s\right)_{i+1,k+1} = \sum_{j=i}^{k} \left(\mathbf{P}_n^r\right)_{j+1,k+1} \binom{j}{i} (r-s)^{j-i}, \qquad i = 0,\dots,n, \qquad k = 0,\dots,n, \tag{B22}$$

since

$$P_k^s(x) = P_k^r(x + r - s) = \sum_{i=0}^{k} \left(\mathbf{P}_n^r\right)_{i+1,k+1} \sum_{j=0}^{i} \binom{i}{j} x^j \, (r - s)^{i-j}$$
$$= \sum_{i=0}^{k} \left(\sum_{j=i}^{k} \left(\mathbf{P}_n^r\right)_{j+1,k+1} \binom{j}{i} (r - s)^{j-i} \right) x^i, \qquad k = 0, \ldots, n. \tag{B23}$$

From the definition of $\mathbf{D}_n$,

$$(\mathbf{D}_n)_{k+1,l+1} = \int_{r-1}^{r} \left(\sum_{j=0}^{n} \left(\mathbf{P}_n^r\right)_{j+1,k+1} x^j \right) \left(\sum_{i=0}^{n} \left(\mathbf{P}_n^r\right)_{i+1,l+1} x^i \right) \mathrm{d}x$$
$$= \left(\sum_{j=0}^{n} \sum_{i=0}^{n} \left(\mathbf{P}_n^r\right)_{j+1,k+1} \left(\mathbf{G}_n^r\right)_{j+1,i+1} \left(\mathbf{P}_n^r\right)_{i+1,l+1} \right), \tag{B24}$$

which results in

$$\mathbf{D}_n = \mathbf{P}_n^{r\,\mathrm{T}} \mathbf{G}_n^r \mathbf{P}_n^r. \tag{B25}$$

This implies that

$$\mathbf{G}_n^r = \left(\mathbf{P}_n^{r\,\mathrm{T}} \right)^{-1} \mathbf{D}_n \left(\mathbf{P}_n^r\right)^{-1}, \tag{B26}$$
$$\mathbf{A}_n^r = \mathbf{P}_n^r \left(\mathbf{D}_n\right)^{-1} \mathbf{P}_n^{r\,\mathrm{T}}. \tag{B27}$$

Consequently,

$$\left(\mathbf{A}_n^r\right)_{i+1,m+1} = \sum_{k=0}^{n} \frac{\left(\mathbf{P}_n^r\right)_{i+1,k+1} \left(\mathbf{P}_n^r\right)_{m+1,k+1}}{(\mathbf{D}_n)_{k+1,k+1}}$$
$$= \sum_{k=\max(i,m)}^{n} (2k + 1) \left(\mathbf{P}_n^r\right)_{i+1,k+1} \left(\mathbf{P}_n^r\right)_{m+1,k+1}, \qquad i = 0, \ldots, n, \qquad m = 0, \ldots, n, \tag{B28}$$

and, in particular,

$$\left(\mathbf{A}_n^1\right)_{i+1,m+1} = \sum_{k=\max(i,m)}^{n} (2k + 1) (-1)^{k+i} \binom{k}{i} \binom{k+i}{k} (-1)^{k+m} \binom{k}{m} \binom{k+m}{k}$$
$$= (-1)^{i+m} \frac{\prod_{k=0}^{n} (n+i+1-k)\,(n+m+1-k)}{(i+m+1)\,i!\,(n-i)!\,m!\,(n-m)!}, \tag{B29}$$
$$\left(\mathbf{A}_n^{1/2}\right)_{i+1,m+1} = \sum_{k=\max(i,m)}^{n} (2k + 1)\, 2^{k+i} \binom{k}{i} \binom{\frac{k+i-1}{2}}{k} 2^{k+m} \binom{k}{m} \binom{\frac{k+m-1}{2}}{k}$$
$$= 2^{i+m} \frac{\prod_{k=0}^{n} (n+i+1-2k)\,(n+m+1-2k) + \prod_{k=0}^{n} (n+i-2k)\,(n+m-2k)}{(i+m+1)\,i!\,(n-i)!\,m!\,(n-m)!}, \tag{B30}$$
$$\left(\mathbf{A}_n^0\right)_{i+1,m+1} = \sum_{k=\max(i,m)}^{n} (2k + 1) \binom{k}{i} \binom{k+i}{k} \binom{k}{m} \binom{k+m}{k}$$
$$= \frac{\prod_{k=0}^{n} (n+i+1-k)\,(n+m+1-k)}{(i+m+1)\,i!\,(n-i)!\,m!\,(n-m)!}, \tag{B31}$$

for $r \in \{0, 1/2, 1\}$. In addition, for all $0 \le s \le 1$,

$$\left(\mathbf{A}_n^s\right)_{i+1,p+1} = \sum_{k=0}^{n} \frac{\left(\mathbf{P}_n^s\right)_{i+1,k+1} \left(\mathbf{P}_n^s\right)_{p+1,k+1}}{(\mathbf{D}_n)_{k+1,k+1}} = \sum_{j=i}^{n} \sum_{m=p}^{n} \binom{j}{i} \binom{m}{p} (r - s)^{j-i+m-p} \sum_{k=0}^{n} \frac{\left(\mathbf{P}_n^r\right)_{j+1,k+1} \left(\mathbf{P}_n^r\right)_{m+1,k+1}}{(\mathbf{D}_n)_{k+1,k+1}}$$
$$= \sum_{j=i}^{n} \sum_{m=p}^{n} \binom{j}{i} \binom{m}{p} (r - s)^{j-i+m-p} \left(\mathbf{A}_n^r\right)_{j+1,m+1}, \qquad i = 0, \ldots, n, \qquad p = 0, \ldots, n. \tag{B32}$$

C PROOF OF LEMMA 3

One can prove that, for $r \in \{0, 1/2, 1\}$ and all $0 \le s \le 1$,

$$(\mathbf{c}_n(\tau))_{m+1} = \sum_{j=0}^{n} \sum_{i=0}^{j} \binom{j}{i} \left(\frac{\tau}{\Delta t} - 1 + s \right)^i (r - s)^{j-i} \left(\mathbf{A}_n^r\right)_{j+1,m+1}$$
$$= \sum_{i=0}^{n} \left(\frac{\tau}{\Delta t} - 1 + r \right)^i \left(\mathbf{A}_n^r\right)_{i+1,m+1}, \qquad m = 0, \ldots, n, \tag{C33}$$

which implies that

$$\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1} = \sum_{i=0}^{n} \left(\frac{\tau - z}{\Delta t} \right)^i \sum_{j=i}^{n} \sum_{m=p}^{n} \binom{j}{i} \binom{m}{p} \left(r - 1 + \frac{z}{\Delta t} \right)^{j-i+m-p} \left(\mathbf{A}_n^r\right)_{j+1,m+1}$$
$$= \sum_{m=p}^{n} \binom{m}{p} \left(\frac{z}{\Delta t} - 1 + r \right)^{m-p} \sum_{i=0}^{n} \sum_{j=i}^{n} \binom{j}{i} \left(\frac{\tau - z}{\Delta t} \right)^i \left(\frac{z}{\Delta t} - 1 + r \right)^{j-i} \left(\mathbf{A}_n^r\right)_{j+1,m+1}$$
$$= \sum_{m=p}^{n} \binom{m}{p} \left(\frac{z}{\Delta t} - 1 + r \right)^{m-p} \sum_{i=0}^{n} \left(\frac{\tau}{\Delta t} - 1 + r \right)^i \left(\mathbf{A}_n^r\right)_{i+1,m+1}$$
$$= \sum_{m=p}^{n} \binom{m}{p} \left(\frac{z}{\Delta t} - 1 + r \right)^{m-p} (\mathbf{c}_n(\tau))_{m+1}, \qquad p = 0, \ldots, n. \tag{C34}$$

Note that, upon defining

$$(\mathbf{d}_n(\tau))_{l+1} = \frac{1}{\Delta t}\left(\frac{\tau}{\Delta t} - 1 + r\right)^l, \qquad l = 0,\dots,n, \tag{C35}$$

the definition of $\mathbf{c}_n(\tau)$ implies that

$$\begin{aligned}\int_0^{\Delta t} (\mathbf{d}_n(\tau))_{l+1}\,(\mathbf{c}_n(\tau))_{m+1}\,\mathrm{d}\tau &= \textstyle\sum_{i=0}^n \frac{1}{\Delta t}\int_0^{\Delta t}\left(\frac{\tau}{\Delta t} - 1 + r\right)^{i+l}\left(\mathbf{A}_n^r\right)_{i+1,m+1}\mathrm{d}\tau = \sum_{i=0}^n \left(\mathbf{G}_n^r\right)_{l+1,i+1}\left(\mathbf{A}_n^r\right)_{i+1,m+1}\\ &= \begin{cases}0, & l \neq m\\ 1, & l = m\end{cases}, \qquad l = 0,\dots,n, \qquad m = 0,\dots,n.\end{aligned} \tag{C36}$$

The previous equality results in

$$\begin{aligned}\frac{1}{\Delta t}\int_0^{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\mathrm{d}\tau &= \frac{1}{\Delta t}\int_0^{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l \textstyle\sum_{m=p}^n \binom{m}{p}\left(\frac{z}{\Delta t} - 1 + r\right)^{m-p}(\mathbf{c}_n(\tau))_{m+1}\,\mathrm{d}\tau\\
&= \frac{1}{\Delta t}\int_0^{\Delta t}\textstyle\sum_{m=p}^n \binom{m}{p}\left(\frac{z}{\Delta t} - 1 + r\right)^{m-p}(\mathbf{c}_n(\tau))_{m+1}\sum_{j=0}^l\binom{l}{j}\left(\frac{\tau}{\Delta t} - 1 + r\right)^j\left(1 - r - \frac{z}{\Delta t}\right)^{l-j}\mathrm{d}\tau\\
&= \textstyle\sum_{m=p}^n \binom{m}{p}\left(\frac{z}{\Delta t} - 1 + r\right)^{m-p}\sum_{j=0}^l\binom{l}{j}\left(1 - r - \frac{z}{\Delta t}\right)^{l-j}\int_0^{\Delta t}(\mathbf{d}_n(\tau))_{j+1}(\mathbf{c}_n(\tau))_{m+1}\,\mathrm{d}\tau\\
&= \begin{cases}0, & l = 0,\dots,p-1\\ \sum_{m=p}^l \binom{m}{p}\left(\frac{z}{\Delta t} - 1 + r\right)^{m-p}\binom{l}{m}\left(1 - r - \frac{z}{\Delta t}\right)^{l-m}, & l = p,\dots,n\end{cases}\\
&= \begin{cases}0, & l = 0,\dots,p-1\\ \sum_{m=0}^{l-p} \binom{m+p}{p}\left(\frac{z}{\Delta t} - 1 + r\right)^{m}\binom{l}{m+p}\left(1 - r - \frac{z}{\Delta t}\right)^{l-p-m}, & l = p,\dots,n\end{cases}\\
&= \begin{cases}0, & l = 0,\dots,p-1\\ \frac{l!}{p!(l-p)!}\sum_{m=0}^{l-p}\frac{(l-p)!}{m!(l-p-m)!}\left(\frac{z}{\Delta t} - 1 + r\right)^{m}\left(1 - r - \frac{z}{\Delta t}\right)^{l-p-m}, & l = p,\dots,n\end{cases}\\
&= \begin{cases}0, & l = 0,\dots,p-1\\ \frac{l!}{p!(l-p)!}0^{l-p}, & l = p,\dots,n\end{cases}\\
&= \begin{cases}0, & l \neq p\\ 1, & l = p\end{cases}, \qquad l = 0,\dots,n.\end{aligned} \tag{C37}$$

Consequently,

$$\begin{aligned}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau \frac{p!}{\Delta t^{n+1}}\frac{(\tau-\zeta)^{n-1}}{(n-1)!}\mathrm{d}\zeta\mathrm{d}\tau &= \frac{p!}{(n-1)!\Delta t}\int_0^{\Delta t}\frac{\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}}{\Delta t^n}\int_z^\tau(\tau-\zeta)^{n-1}\,\mathrm{d}\zeta\mathrm{d}\tau\\
&= \frac{p!}{n!\Delta t}\int_0^{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^n\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\mathrm{d}\tau\\
&= \begin{cases}0, & p = 0,\dots,n-1\\ 1, & p = n\end{cases}.\end{aligned} \tag{C38}$$

From the fact that $\tilde{\mathbf{y}}_c^{(n-1)}$ is absolutely continuous,

$$\begin{aligned}\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) &= \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\tilde{\mathbf{y}}_c(t - \Delta t + \tau)\mathrm{d}\tau\\
&= \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\left(\textstyle\sum_{l=0}^{n-1}\frac{\tilde{\mathbf{y}}_c^{(l)}(t-\Delta t+z)}{l!}(\tau - z)^l\right)\mathrm{d}\tau\\
&\quad + \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)}{(n-1)!}(\tau-\zeta)^{n-1}\,\mathrm{d}\zeta\mathrm{d}\tau\\
&= \textstyle\sum_{l=0}^{n-1}\tilde{\mathbf{y}}_c^{(l)}(t - \Delta t + z)\frac{p!\Delta t^{l-p}}{l!\Delta t}\int_0^{\Delta t}\left(\frac{\tau-z}{\Delta t}\right)^l\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\mathrm{d}\tau\\
&\quad + \Delta t^{-p}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{p!}{\Delta t^{n+1}}\tilde{\mathbf{y}}_c^{(n)}(t - \Delta t + \zeta)\Delta t^n\frac{(\tau-\zeta)^{n-1}}{(n-1)!}\mathrm{d}\zeta\mathrm{d}\tau,\end{aligned} \tag{C39}$$

which implies that, for $p = 0,\dots,n-1$,

$$\begin{aligned}\mathcal{H}_{p,n,z}(\tilde{\mathbf{y}}_c, t) &= \tilde{\mathbf{y}}_c^{(p)}(t - \Delta t + z) + \Delta t^{-p}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{p!}{\Delta t^{n+1}}\tilde{\mathbf{y}}_c^{(n)}(t - \Delta t + \zeta)\Delta t^n\frac{(\tau-\zeta)^{n-1}}{(n-1)!}\mathrm{d}\zeta\mathrm{d}\tau\\
&= \tilde{\mathbf{y}}_c^{(p)}(t - \Delta t + z) + \frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^\tau\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta) - \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\,\mathrm{d}\zeta\mathrm{d}\tau,\end{aligned} \tag{C40}$$

and

$$\begin{aligned}\mathcal{H}_{n,n,z}(\tilde{\mathbf{y}}_c,t) &= \Delta t^{-n}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{n+1}\int_z^{\tau}\frac{n!}{\Delta t^{n+1}}\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)\Delta t^n\frac{(\tau-\zeta)^{n-1}}{(n-1)!}\mathrm{d}\zeta\mathrm{d}\tau\\ &= \tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)+\frac{n!}{\Delta t^{n+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{n+1}\int_z^{\tau}\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau.\end{aligned} \tag{C41}$$

D PROOF OF LEMMA 4

Since $\tilde{\mathbf{y}}_c^{(n)}(t)=\tilde{\mathbf{y}}_u^{(n)}(t)+\tilde{\mathbf{y}}_a^{(n)}(t)$ from (29) and, according to the Cauchy formula for repeated integration,

$$\begin{aligned}&\frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^{\tau}\frac{\tilde{\mathbf{y}}_c^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\
&=\frac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^{\tau}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\
&=-\frac{p!}{\Delta t^{p+1}}\int_0^{z}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_{\tau}^{z}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\
&\quad+\frac{p!}{\Delta t^{p+1}}\int_z^{\Delta t}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\int_z^{\tau}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(n-1)!}(\tau-\zeta)^{n-1}\mathrm{d}\zeta\mathrm{d}\tau\\
&=-\frac{p!}{\Delta t^{p+1}}\int_0^{z}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^{n-1}}\int_0^{\zeta}\frac{(\zeta-\tau)^{n-1}}{(n-1)!}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\mathrm{d}\tau\mathrm{d}\zeta\\
&\quad-\frac{p!}{\Delta t^{p+1}}\int_z^{\Delta t}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^{n-1}}\int_{\Delta t}^{\zeta}\frac{(\zeta-\tau)^{n-1}}{(n-1)!}\left(\mathbf{c}_{n,z}(\tau)\right)_{p+1}\mathrm{d}\tau\mathrm{d}\zeta\\
&=\frac{p!}{\Delta t^{p+1}}\int_0^{z}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^{n}}\left(\mathbf{c}_{n,z,1}^{(-n)}(\zeta)\right)_{p+1}\mathrm{d}\zeta\\
&\quad+\frac{p!}{\Delta t^{p+1}}\int_z^{\Delta t}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\zeta)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\zeta)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)}{(-1)^{n}}\left(\mathbf{c}_{n,z,0}^{(-n)}(\zeta)\right)_{p+1}\mathrm{d}\zeta\\
&=\frac{p!}{\Delta t^{p+1}}\int_0^{z}\left(\mathbf{c}_{n,z,1}^{(-n)}(\tau)\right)_{p+1}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)}{(-1)^{n}}\mathrm{d}\tau\\
&\quad+\frac{p!}{\Delta t^{p+1}}\int_z^{\Delta t}\left(\mathbf{c}_{n,z,0}^{(-n)}(\tau)\right)_{p+1}\frac{\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u^{(n)}(t-\Delta t+z)+\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)}{(-1)^{n}}\mathrm{d}\tau,\end{aligned} \tag{D42}$$

this lemma is a consequence of Lemma 3.

E PROOF OF LEMMA 5

The following equation is a consequence of Lemma 4:

$$\begin{aligned}\mathcal{D}(\tilde{\mathbf{y}}_c,t)-\tilde{\mathbf{y}}_u(t) &= \frac{1}{\Delta t^2}\int_0^{\Delta t}\left(\mathbf{c}_{1,\Delta t,1}^{(-1)}(\tau)\right)_2\frac{\tilde{\mathbf{y}}_u(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u(t)+\tilde{\mathbf{y}}_a(t-\Delta t+\tau)}{-1}\mathrm{d}\tau\\
&=\int_0^{\Delta t}\left(\left(\frac{\tau}{\Delta t}\right)\left(\mathbf{a}_{1,1,\Delta t}^{1}\right)_1+\left(\frac{\tau}{\Delta t}\right)^2\frac{\left(\mathbf{a}_{1,1,\Delta t}^{1}\right)_2}{2}\right)\frac{\tilde{\mathbf{y}}_u(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u(t)}{-\Delta t}\mathrm{d}\tau\\
&\quad+\int_0^{\Delta t}\left(\left(\frac{\tau}{\Delta t}\right)\left(\mathbf{a}_{1,1,\Delta t}^{1}\right)_1+\left(\frac{\tau}{\Delta t}\right)^2\frac{\left(\mathbf{a}_{1,1,\Delta t}^{1}\right)_2}{2}\right)\frac{\tilde{\mathbf{y}}_a(t-\Delta t+\tau)}{-\Delta t}\mathrm{d}\tau\\
&=\int_0^{\Delta t}6\left(\frac{\tau}{\Delta t}-\left(\frac{\tau}{\Delta t}\right)^2\right)\frac{\tilde{\mathbf{y}}_u(t-\Delta t+\tau)-\tilde{\mathbf{y}}_u(t)}{\Delta t}\mathrm{d}\tau+\int_0^{\Delta t}6\left(\frac{\tau}{\Delta t}-\left(\frac{\tau}{\Delta t}\right)^2\right)\frac{\tilde{\mathbf{y}}_a(t-\Delta t+\tau)}{\Delta t}\mathrm{d}\tau.\end{aligned} \tag{E43}$$

F PROOF OF LEMMA 6

The estimation of $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p=0,\dots,n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$ in (47) and (48) uses the inputs $\tilde{\mathbf{y}}_c(t)$ and $\tilde{\mathbf{y}}_a^{(n)}(t)$. Then, one applies the Laplace transform to (47) and (48). Considering that $\delta\tilde{\mathbf{y}}_c(t)=\mathbf{0}_{n_c}$ and $\delta\tilde{\mathbf{y}}_a^{(n)}(t)=\mathbf{0}_{n_c}$ for $t<0$, that is, $\delta\tilde{\mathbf{y}}_c(t)=H(t)\delta\tilde{\mathbf{y}}_c(t)$

and $\delta\tilde{\mathbf{y}}_a^{(n)}(t) = H(t)\delta\tilde{\mathbf{y}}_a^{(n)}(t)$, where $H(t)$ is the Heaviside step function, it holds that

$$\begin{aligned}
\hat{\mathbf{Y}}_c^{(p)}(s) &= \int_0^\infty \delta\hat{\mathbf{y}}_c^{(p)}(t)\exp(-st)\,\mathrm{d}t \\
&= \int_0^\infty \left(\mathcal{H}_{p,n,\Delta t}(\delta\tilde{\mathbf{y}}_c, t) - \mathcal{Z}_{p,n,\Delta t}(\delta\tilde{\mathbf{y}}_a^{(n)}, t)\right)\exp(-st)\,\mathrm{d}t \\
&= \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,\Delta t}(\tau)\right)_{p+1}\delta\tilde{\mathbf{y}}_c(t-\Delta t+\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&\quad - \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\tau)\right)_{p+1}\tfrac{\delta\tilde{\mathbf{y}}_a^{(n)}(t-\Delta t+\tau)}{(-1)^n}\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&= \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_{t-\Delta t}^{t}\left(\mathbf{c}_{n,\Delta t}(\tau-t+\Delta t)\right)_{p+1}H(\tau)\delta\tilde{\mathbf{y}}_c(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&\quad - \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_{t-\Delta t}^{t}(-1)^n\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\tau-t+\Delta t)\right)_{p+1}H(\tau)\delta\tilde{\mathbf{y}}_a^{(n)}(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&= \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_0^{t}H(\tau-t+\Delta t)\left(\mathbf{c}_{n,\Delta t}(\tau-t+\Delta t)\right)_{p+1}\delta\tilde{\mathbf{y}}_c(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&\quad - \int_0^\infty \tfrac{p!}{\Delta t^{p+1}}\int_0^{t}H(\tau-t+\Delta t)(-1)^n\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\tau-t+\Delta t)\right)_{p+1}\delta\tilde{\mathbf{y}}_a^{(n)}(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t,
\end{aligned} \tag{F44}$$

with $\hat{\mathbf{Y}}_c^{(p)}(s)$ replaced by $\hat{\mathbf{Y}}_u^{(n)}(s)$ if $p = n$. Upon defining

$$g_{p,n,\Delta t}(t) = \tfrac{p!}{\Delta t^{p+1}}H(\Delta t - t)\left(\mathbf{c}_{n,\Delta t}(\Delta t - t)\right)_{p+1}, \tag{F45}$$

$$g_{p,n,\Delta t,1}^{(-n)}(t) = -\tfrac{p!}{\Delta t^{p+1}}H(\Delta t - t)(-1)^n\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\Delta t - t)\right)_{p+1}, \tag{F46}$$

and noting that

$$\left(\mathbf{c}_{n,\Delta t}(\tau)\right)_{p+1} = \textstyle\sum_{i=0}^n\left(\tfrac{\tau}{\Delta t}\right)^i\left(\mathbf{a}_{p,n,\Delta t}^1\right)_{i+1} = \sum_{i=0}^n\left(\tfrac{\tau}{\Delta t}\right)^i\sum_{j=i}^n\binom{j}{i}(-1)^{j-i}\left(\mathbf{A}_n^0\right)_{j+1,p+1}, \tag{F47}$$

$$\left(\mathbf{c}_{n,\Delta t,1}^{(-n)}(\tau)\right)_{p+1} = \textstyle\sum_{i=0}^n\left(\tfrac{\tau}{\Delta t}\right)^{i+n}\tfrac{\Delta t^n i!}{(i+n)!}\left(\mathbf{a}_{p,n,\Delta t}^1\right)_{i+1} = \sum_{i=0}^n\left(\tfrac{\tau}{\Delta t}\right)^{i+n}\tfrac{\Delta t^n i!}{(i+n)!}\sum_{j=i}^n\binom{j}{i}(-1)^{j-i}\left(\mathbf{A}_n^0\right)_{j+1,p+1}, \tag{F48}$$

it follows that, according to the properties of Laplace transforms with respect to convolution,

$$\begin{aligned}
\hat{\mathbf{Y}}_c^{(p)}(s) &= \int_0^\infty\int_0^t g_{p,n,\Delta t}(t-\tau)\delta\tilde{\mathbf{y}}_c(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t + \int_0^\infty\int_0^t g_{p,n,\Delta t,1}^{(-n)}(t-\tau)\delta\tilde{\mathbf{y}}_a^{(n)}(\tau)\mathrm{d}\tau\exp(-st)\,\mathrm{d}t \\
&= \left(\int_0^\infty g_{p,n,\Delta t}(t)\exp(-st)\,\mathrm{d}t\right)\left(\int_0^\infty\delta\tilde{\mathbf{y}}_c(t)\exp(-st)\,\mathrm{d}t\right) \\
&\quad + \left(\int_0^\infty g_{p,n,\Delta t,1}^{(-n)}(t)\exp(-st)\,\mathrm{d}t\right)\left(\int_0^\infty\delta\tilde{\mathbf{y}}_a^{(n)}(\tau)\exp(-st)\,\mathrm{d}t\right) \\
&= \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\textstyle\sum_{i=0}^n\left(1-\tfrac{t}{\Delta t}\right)^i\sum_{j=i}^n\binom{j}{i}(-1)^{j-i}\left(\mathbf{A}_n^0\right)_{j+1,p+1}\exp(-st)\,\mathrm{d}t\mathbf{S}\tilde{\mathbf{Y}}(s) \\
&\quad - \tfrac{p!}{\Delta t^{p+1}}\int_0^{\Delta t}\textstyle\sum_{i=0}^n\left(1-\tfrac{t}{\Delta t}\right)^{i+n}\tfrac{\Delta t^n i!}{(i+n)!}\sum_{j=i}^n\binom{j}{i}(-1)^{n+j-i}\left(\mathbf{A}_n^0\right)_{j+1,p+1}\exp(-st)\,\mathrm{d}t\tilde{\mathbf{Y}}_a^{(n)}(s),
\end{aligned} \tag{F49}$$

with $\hat{\mathbf{Y}}_c^{(p)}(s)$ replaced by $\hat{\mathbf{Y}}_u^{(n)}(s)$ if $p = n$. The proof is similar for (49) and (50) and yields the second equality of the lemma, with $\tilde{\mathbf{Y}}_c^{(p)}(s)$ removed if $p = n$.

G OPEN-LOOP AND CLOSED-LOOP TRANSFER FUNCTIONS

We start by describing the plant in (1) in terms of deviation variables, which yields

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}_p\delta\mathbf{x}(t) + \mathbf{B}_p\delta\mathbf{u}(t), \tag{G50a}$$

$$\delta\mathbf{y}(t) = \mathbf{C}_p\delta\mathbf{x}(t), \tag{G50b}$$

with

$$\mathbf{A}_p = \tfrac{\partial\mathbf{f}}{\partial\mathbf{x}}(\bar{\mathbf{x}}, \bar{\mathbf{u}}), \tag{G51a}$$

$$\mathbf{B}_p = \tfrac{\partial\mathbf{f}}{\partial\mathbf{u}}(\bar{\mathbf{x}}, \bar{\mathbf{u}}), \tag{G51b}$$

$$\mathbf{C}_p = \tfrac{\partial\mathbf{g}}{\partial\mathbf{x}}(\bar{\mathbf{x}}). \tag{G51c}$$

The transfer function of the plant is computed as $\mathbf{G}_p(s)$:

$$\mathbf{G}_p(s) = \mathbf{C}_p\left(s\mathbf{I}_{n_x} - \mathbf{A}_p\right)^{-1}\mathbf{B}_p. \tag{G52}$$

The computation of $\tilde{\mathbf{y}}_a^{(n)}(t)$ from $\tilde{\mathbf{y}}(t)$ and $\tilde{\mathbf{u}}(t)$ implies

$$\delta\tilde{\mathbf{y}}_a^{(n)}(t) = \mathbf{D}_a^y \delta\tilde{\mathbf{y}}(t) + \mathbf{D}_a^u \delta\tilde{\mathbf{u}}(t), \tag{G53}$$

with

$$\mathbf{D}_a^y = \frac{\partial \bar{\mathbf{s}}_a}{\partial \mathbf{y}}\left(\bar{\mathbf{y}}, \bar{\mathbf{u}}\right), \tag{G54a}$$

$$\mathbf{D}_a^u = \frac{\partial \bar{\mathbf{s}}_a}{\partial \mathbf{u}}\left(\bar{\mathbf{y}}, \bar{\mathbf{u}}\right). \tag{G54b}$$

As shown in Lemma 6, the Laplace transform of $\hat{\mathbf{y}}_c^{(p)}(t)$, for $p = 0, \ldots, n-1$, and $\hat{\mathbf{y}}_u^{(n)}(t)$ in (47) and (48) is given by (53) and (54). Since this Laplace transform includes $\exp(-s\Delta t)$, one can use the Padé approximation of order $n+1$ of $\exp(-s\Delta t)$ about $s = 0$ to express (53) and (54) as the rational function

$$\hat{\mathbf{Y}}_c^{(p)}(s) = \frac{\sum_{k=0}^{n} b_{p,n+1-k}^y s^k}{s^{n+1} + \sum_{k=0}^{n} a_{p,n+1-k} s^k} \mathbf{S}\tilde{\mathbf{Y}}(s) + \frac{\sum_{k=0}^{n} b_{p,n+1-k}^a s^k}{s^{n+1} + \sum_{k=0}^{n} a_{p,n+1-k} s^k} \tilde{\mathbf{Y}}_a^{(n)}(s), \tag{G55}$$

with $\hat{\mathbf{Y}}_c^{(p)}(s)$ replaced by $\hat{\mathbf{Y}}_u^{(n)}(s)$ if $p = n$.

Hence, the derivative estimation can be described by the minimal realization of the transfer function (G55), which yields the following linear system with $n_z = (n + 1 - \lfloor \alpha \rfloor)(n+1) n_c$ states:

$$\delta\dot{\mathbf{z}}(t) = \mathbf{A}_e \delta\mathbf{z}(t) + \mathbf{B}_e^y \delta\tilde{\mathbf{y}}(t) + \mathbf{B}_e^a \delta\tilde{\mathbf{y}}_a^{(n)}(t), \tag{G56a}$$

$$\delta\hat{\mathbf{y}}_c^{(p)}(t) = \mathbf{C}_{e,p} \delta\mathbf{z}(t), \quad p = \lfloor \alpha \rfloor, \ldots, n-1, \tag{G56b}$$

$$\delta\hat{\mathbf{y}}_u^{(n)}(t) = \mathbf{C}_{e,n} \delta\mathbf{z}(t), \tag{G56c}$$

with

$$\mathbf{A}_e = \operatorname{diag}\begin{bmatrix} \mathbf{A}_{e,n} \\ \vdots \\ \mathbf{A}_{e,\lfloor \alpha \rfloor} \end{bmatrix}, \quad \mathbf{A}_{e,p} = \begin{bmatrix} -a_{p,1}\mathbf{I}_{n_c} & \mathbf{I}_{n_c} & \mathbf{0}_{n_c \times (n-1)n_c} \\ \vdots & \mathbf{0}_{(n-1)n_c \times n_c} & \mathbf{I}_{(n-1)n_c} \\ -a_{p,n+1}\mathbf{I}_{n_c} & \mathbf{0}_{n_c \times n_c} & \mathbf{0}_{n_c \times (n-1)n_c} \end{bmatrix}, \quad p = \lfloor \alpha \rfloor, \ldots, n, \tag{G57a}$$

$$\mathbf{B}_e^y = \begin{bmatrix} \mathbf{B}_{e,n}^y \\ \vdots \\ \mathbf{B}_{e,\lfloor \alpha \rfloor}^y \end{bmatrix}, \quad \mathbf{B}_{e,p}^y = \begin{bmatrix} b_{p,1}^y \mathbf{S} \\ \vdots \\ b_{p,n+1}^y \mathbf{S} \end{bmatrix}, \quad p = \lfloor \alpha \rfloor, \ldots, n, \tag{G57b}$$

$$\mathbf{B}_e^a = \begin{bmatrix} \mathbf{B}_{e,n}^a \\ \vdots \\ \mathbf{B}_{e,\lfloor \alpha \rfloor}^a \end{bmatrix}, \quad \mathbf{B}_{e,p}^a = \begin{bmatrix} b_{p,1}^a \mathbf{I}_{n_c} \\ \vdots \\ b_{p,n+1}^a \mathbf{I}_{n_c} \end{bmatrix}, \quad p = \lfloor \alpha \rfloor, \ldots, n, \tag{G57c}$$

$$\mathbf{C}_{e,p} = \begin{bmatrix} \mathbf{0}_{n_c \times (n-p)(n+1)n_c} & \mathbf{I}_{n_c} & \mathbf{0}_{n_c \times n n_c} & \mathbf{0}_{n_c \times (p - \lfloor \alpha \rfloor)(n+1)n_c} \end{bmatrix}, \quad p = \lfloor \alpha \rfloor, \ldots, n. \tag{G57d}$$

In particular, if $n = 1$, the Laplace transform of $\hat{\mathbf{y}}_u(t)$ in (51) is given by

$$\hat{\mathbf{Y}}_u(s) = 6\frac{(s\Delta t+2)\exp(-s\Delta t) + (s\Delta t - 2)}{s^2 \Delta t^3} \mathbf{S}\tilde{\mathbf{Y}}(s) - 6\frac{(s\Delta t+2)\exp(-s\Delta t) + (s\Delta t - 2)}{s^3 \Delta t^3} \tilde{\mathbf{Y}}_a^{(n)}(s). \tag{G58}$$

Since this Laplace transform includes $\exp(-s\Delta t)$, one can use the Padé approximation

$$\exp(-s\Delta t) = \frac{1 - \frac{s\Delta t}{2} + \frac{s^2 \Delta t^2}{12}}{1 + \frac{s\Delta t}{2} + \frac{s^2 \Delta t^2}{12}} \tag{G59}$$

to express (G58) as the rational function

$$\hat{\mathbf{Y}}_u(s) = \frac{\frac{12}{\Delta t^2} s\mathbf{S}\tilde{\mathbf{Y}}(s) - \frac{12}{\Delta t^2}\tilde{\mathbf{Y}}_a^{(n)}(s)}{s^2 + \frac{6}{\Delta t}s + \frac{12}{\Delta t^2}}. \tag{G60}$$

Hence, the derivative estimation can be described by the minimal realization of the transfer function (G60), which yields the following linear system with $n_z = 2n_c$ states:

$$\delta\dot{\mathbf{z}}(t) = \mathbf{A}_e \delta\mathbf{z}(t) + \mathbf{B}_e^y \delta\tilde{\mathbf{y}}(t) + \mathbf{B}_e^a \delta\tilde{\mathbf{y}}_a^{(n)}(t), \tag{G61a}$$

$$\delta\hat{\mathbf{y}}_u(t) = \mathbf{C}_{e,1} \delta\mathbf{z}(t), \tag{G61b}$$

with

$$\mathbf{A}_e = \begin{bmatrix} -\frac{6}{\Delta t}\mathbf{I}_{n_c} & \mathbf{I}_{n_c} \\ -\frac{12}{\Delta t^2}\mathbf{I}_{n_c} & \mathbf{0}_{n_c\times n_c} \end{bmatrix}, \tag{G62a}$$

$$\mathbf{B}_e^y = \begin{bmatrix} \frac{12}{\Delta t^2}\mathbf{S} \\ \mathbf{0}_{n_c\times n_y} \end{bmatrix}, \tag{G62b}$$

$$\mathbf{B}_e^a = \begin{bmatrix} \mathbf{0}_{n_c\times n_c} \\ -\frac{12}{\Delta t^2}\mathbf{I}_{n_c} \end{bmatrix}, \tag{G62c}$$

$$\mathbf{C}_{e,1} = \begin{bmatrix} \mathbf{I}_{n_c} & \mathbf{0}_{n_c\times n_c} \end{bmatrix}. \tag{G62d}$$

Furthermore, the feedback-linearizing controller in (13) and (14) can be written as

$$\delta\tilde{\mathbf{u}}(t) = \sum_{p=\lfloor\alpha\rfloor}^{n-1}\mathbf{D}_u^{e,p}\delta\hat{\mathbf{y}}_c^{(p)}(t) + \mathbf{D}_u^{e,n}\delta\hat{\mathbf{y}}_u^{(n)}(t) + \mathbf{D}_u^y\delta\tilde{\mathbf{y}}(t) + \sum_{p=0}^{n-1}\mathbf{D}_u^{r,p}\delta\mathbf{r}^{(p)}(t), \tag{G63}$$

with

$$\mathbf{D}_u^{e,p} = -\tilde{\mathbf{B}}_a(\bar{\mathbf{y}})^{-1}\left(1-\alpha 0^p\right)\binom{n}{p}\tau_c^{p-n}, \quad p = \lfloor\alpha\rfloor,\ldots,n, \tag{G64a}$$

$$\mathbf{D}_u^y = -\tilde{\mathbf{B}}_a(\bar{\mathbf{y}})^{-1}\left(\mathbf{S}\alpha\tau_c^{-n} + \frac{\partial\tilde{\mathbf{s}}_a}{\partial\mathbf{y}}(\bar{\mathbf{y}},\bar{\mathbf{u}})\right), \tag{G64b}$$

$$\mathbf{D}_u^{r,p} = \tilde{\mathbf{B}}_a(\bar{\mathbf{y}})^{-1}\binom{n}{p}\tau_c^{p-n}, \quad p = 0,\ldots,n-1. \tag{G64c}$$

The combination of (G53), (G56), and (G63) leads to the state-space representation of the controller as

$$\delta\dot{\mathbf{z}}(t) = \mathbf{A}_k\delta\mathbf{z}(t) + \mathbf{B}_k^y\delta\tilde{\mathbf{y}}(t) + \sum_{p=0}^{n-1}\mathbf{B}_k^{r,p}\delta\mathbf{r}^{(p)}(t), \tag{G65a}$$

$$\delta\tilde{\mathbf{u}}(t) = \mathbf{C}_k\delta\mathbf{z}(t) + \mathbf{D}_k^y\delta\tilde{\mathbf{y}}(t) + \sum_{p=0}^{n-1}\mathbf{D}_k^{r,p}\delta\mathbf{r}^{(p)}(t), \tag{G65b}$$

with

$$\mathbf{A}_k = \mathbf{A}_e + \mathbf{B}_e^a\mathbf{D}_a^u\left(\sum_{p=\lfloor\alpha\rfloor}^{n}\mathbf{D}_u^{e,p}\mathbf{C}_{e,p}\right), \tag{G66a}$$

$$\mathbf{B}_k^y = \mathbf{B}_e^y + \mathbf{B}_e^a\mathbf{D}_a^y + \mathbf{B}_e^a\mathbf{D}_a^u\mathbf{D}_u^y, \tag{G66b}$$

$$\mathbf{B}_k^{r,p} = \mathbf{B}_e^a\mathbf{D}_a^u\mathbf{D}_u^{r,p}, \quad p = 0,\ldots,n-1, \tag{G66c}$$

$$\mathbf{C}_k = \sum_{p=\lfloor\alpha\rfloor}^{n}\mathbf{D}_u^{e,p}\mathbf{C}_{e,p}, \tag{G66d}$$

$$\mathbf{D}_k^y = \mathbf{D}_u^y, \tag{G66e}$$

$$\mathbf{D}_k^{r,p} = \mathbf{D}_u^{r,p}, \quad p = 0,\ldots,n-1. \tag{G66f}$$

and $\mathbf{D}_a^y$, $\mathbf{D}_a^u$ in (G54), $\mathbf{A}_e$, $\mathbf{B}_e^y$, $\mathbf{B}_e^a$, $\mathbf{C}_{e,p}$ in (G57), and $\mathbf{D}_u^{e,p}$, $\mathbf{D}_u^y$, $\mathbf{D}_u^{r,p}$ in (G64).

The transfer functions of the actuator inputs $\tilde{\mathbf{u}}(t)$ with respect to the sensor outputs $\tilde{\mathbf{y}}(t)$ (superscript y) and the setpoints $\mathbf{r}(t)$ (r) are computed as $\mathbf{G}_k^y(s)$ and $\mathbf{G}_k^r(s)$:

$$\mathbf{G}_k^y(s) = \mathbf{C}_k\left(s\mathbf{I}_{n_z} - \mathbf{A}_k\right)^{-1}\mathbf{B}_k^y + \mathbf{D}_k^y, \tag{G67a}$$

$$\mathbf{G}_k^r(s) = \sum_{p=0}^{n-1}\mathbf{C}_k\left(s\mathbf{I}_{n_z} - \mathbf{A}_k\right)^{-1}\mathbf{B}_k^{r,p}s^p + \sum_{p=0}^{n-1}\mathbf{D}_k^{r,p}s^p. \tag{G67b}$$

From these open-loop state-space representations and transfer functions, one can derive the closed-loop state-space representation (102) with the matrices (99) and the closed-loop transfer functions

$$\mathbf{G}_{cl}^{r\to y}(s) = \sum_{p=0}^{n-1}\mathbf{C}_{cl}^y\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{r,p}s^p, \tag{G68a}$$

$$\mathbf{G}_{cl}^{d\to y}(s) = \mathbf{C}_{cl}^y\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{d,0}, \tag{G68b}$$

$$\mathbf{G}_{cl}^{w\to y}(s) = \mathbf{C}_{cl}^y\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{w,0}, \tag{G68c}$$

$$\mathbf{G}_{cl}^{r\to u}(s) = \sum_{p=0}^{n-1}\mathbf{C}_{cl}^u\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{r,p}s^p + \sum_{p=0}^{n-1}\mathbf{D}_k^{r,p}s^p, \tag{G68d}$$

$$\mathbf{G}_{cl}^{d\to u}(s) = \mathbf{C}_{cl}^u\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{d,0} + \mathbf{I}_{n_u}, \tag{G68e}$$

$$\mathbf{G}_{cl}^{w\to u}(s) = \mathbf{C}_{cl}^u\left(s\mathbf{I}_{n_x+n_z} - \mathbf{A}_{cl}\right)^{-1}\mathbf{B}_{cl}^{w,0} + \mathbf{D}_k^y. \tag{G68f}$$

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Github at https://github.com/dfmrodrigues/SNSF-project-P2ELP2_184521.

References

1. Gevers M. Towards a Joint Design of Identification and Control?. In: Trentelman HL, Willems JC., eds. *Essays on Control: Perspectives in the Theory and its Applications*Birkhäuser, Boston, MA. 1993 (pp. 111-151).

2. Schoukens J, Ljung L. Nonlinear System Identification: A User-Oriented Road Map. *IEEE Control Syst. Mag.* 2019; 39(6): 28-99.

3. Pannocchia G, Rawlings JB. Disturbance models for offset-free model-predictive control. *AIChE J.* 2003; 49(2): 426-437.

4. Koryakovskiy I, Kudruss M, Vallery H, Babuška R, Caarls W. Model-plant mismatch compensation using reinforcement learning. *IEEE Robot. Autom. Lett.* 2018; 3(3): 2471-2477.

5. Paulson JA, Santos TLM, Mesbah A. Mixed stochastic-deterministic tube MPC for offset-free tracking in the presence of plant-model mismatch. *J. Process Control* 2019; 83: 102-120.

6. Bonvin D, Georgakis C, Pantelides CC, et al. Linking Models and Experiments. *Ind. Eng. Chem. Res.* 2016; 55(25): 6891-6903.

7. Kravaris C, Chung CB. Nonlinear state feedback synthesis by global input/output linearization. *AIChE J.* 1987; 33(4): 592-603.

8. Nijmeijer H, Schaft v. dAJ. *Nonlinear Dynamical Control Systems*. New York: Springer-Verlag . 1990.

9. Isidori A. *Nonlinear Control Systems*. London: Springer-Verlag. 3rd ed. 1995.

10. Khalil HK. *Nonlinear Systems*. Englewood Cliffs, NJ: Prentice Hall. 3rd ed. 2000.

11. Henson MA, Seborg DE. Critique of exact linearization strategies for process control. *J. Process Control* 1991; 1(3): 122-139.

12. Sastry SS, Isidori A. Adaptive control of linearizable systems. *IEEE Trans. Autom. Contr.* 1989; 34(11): 1123-1131.

13. Kanellakopoulos I, Kokotovic PV, Morse AS. Systematic design of adaptive controllers for feedback linearizable systems. *IEEE Trans. Autom. Contr.* 1989; 36(11): 1241-1253.

14. Kravaris C, Palanki S. Robust nonlinear state feedback under structured uncertainty. *AIChE J.* 1988; 34(7): 1119-1127.

15. Arkun Y, Calvet JP. Robust stabilization of input/output linearizable systems under uncertainty and disturbances. *AIChE J.* 1992; 38(8): 1145-1156.

16. Liao TL, Fu LC, Hsu CF. Output tracking control of nonlinear systems with mismatched uncertainties. *Syst. Control Lett.* 1992; 18(1): 39-47.

17. Chou YS, Wu W. Robust controller design for uncertain nonlinear systems via feedback linearization. *Chem. Eng. Sci.* 1995; 50(9): 1429-1439.

18. Hahn J, Mönnigmann M, Marquardt W. On the use of bifurcation analysis for robust controller tuning for nonlinear systems. *J. Process Control* 2008; 18(3-4): 408-420.

19. Morari M, Lee JH. Model predictive control: Past, present, and future. *Comput. Chem. Eng.* 1999; 23(4-5): 667-682.

20. Qin SJ, Badgwell TA. A survey of industrial model predictive control technology. *Control Eng. Practice* 2003; 11(7): 733-764.

21. Gros S, Zanon M, Quirynen R, Bemporad A, Diehl M. From linear to nonlinear MPC: bridging the gap via the real-time iteration. *Int. J. Control* 2020; 93(1): 62-80.

22. Nielsen C, Fulford C, Maggiore M. Path following using transverse feedback linearization: Application to a maglev positioning system. *Automatica* 2010; 46(3): 585-590.

23. Akhrif O, Okou FA, Dessaint LA, Champagne R. Application of a multivariable feedback linearization scheme for rotor angle stability and voltage regulation of power systems. *IEEE Trans. Power Syst.* 1999; 14(2): 620-628.

24. Bidram A, Davoudi A, Lewis FL, Guerrero JM. Distributed cooperative secondary control of microgrids using feedback linearization. *IEEE Trans. Power Syst.* 2013; 28(3): 3462-3470.

25. Chien TL, Chen CC, Huang CJ. Feedback linearization control and Its application to MIMO cancer immunotherapy. *IEEE Trans. Control Syst. Technol.* 2010; 18(4): 953-961.

26. Levant A. Higher-order sliding modes, differentiation and output-feedback control. *Int. J. Control* 2003; 76(9-10): 924-941.

27. Tikhonov AN, Goncharsky A, Stepanov VV, Yagola AG. *Numerical Methods for the Solution of Ill-Posed Problems*. Dordrecht: Springer . 1995.

28. Mhamdi A, Marquardt W. Estimation of reaction rates by nonlinear system inversion. *IFAC Proc. Vol.* 2004; 37(1): 167-172.

29. Mboup M, Join C, Fliess M. Numerical differentiation with annihilators in noisy environment. *Numer. Algor.* 2009; 50: 439-467.

30. Reger J, Jouffroy J. On algebraic time-derivative estimation and deadbeat state reconstruction. In: 48th IEEE CDC. ; 2009; Shanghai, P.R. China.

31. Perrier M, Azevedo F. dS, Ferreira EC, Dochain D. Tuning of observer-based estimators: theory and application to the on-line estimation of kinetic parameters. *Control Eng. Practice* 2000; 8(4): 377-388.

32. Freidovich LB, Khalil HK. Performance Recovery of Feedback-Linearization-Based Designs. *IEEE Trans. Autom. Contr.* 2008; 53(10): 2324-2334.

33. Corless M, Tu J. State and input estimation for a class of uncertain systems. *Automatica* 1998; 34(6): 757-764.

34. Rodrigues D, Billeter J, Bonvin D. Control of Reaction Systems via Rate Estimation and Feedback Linearization. *Comput. Aided Chem. Eng.* 2015; 37: 137-142.

35. Zhao Z, Wassick JM, Ferrio J, Ydstie BE. Reaction variants and invariants based observer and controller design for CSTRs. *IFAC-PapersOnLine* 2016; 49(7): 1091-1096.

36. Rodrigues D, Amrhein M, Billeter J, Bonvin D. Fast Estimation of Plant Steady State for Imperfectly Known Dynamic Systems, with Application to Real-Time Optimization. *Ind. Eng. Chem. Res.* 2018; 57(10): 3699-3716.

37. Farschman CA, Viswanath KP, Ydstie BE. Process Systems and Inventory Control. *AIChE J.* 1998; 44(8): 1841-1857.

38. Rodrigues D, Hjalmarsson H. Stability and Performance Analysis of Control Based on Incomplete Models. *IFAC-PapersOnLine* 2019; 52(1): 874-879.

39. Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 1964; 36(8): 1627-1639.

40. Jayawardhana B, Logemann H, Ryan EP. The Circle Criterion and Input-to-State Stability. *IEEE Control Syst. Mag.* 2011; 31(4): 32-67.

41. Vu TNL, Lee M. Independent design of multi-loop PI/PID controllers for interacting multivariable processes. *J. Process Control* 2010; 20(8): 922-933.

42. Hammarström LG. Control of chemical reactors in the subspace of reaction and control variants. *Chem. Eng. Sci.* 1979; 34(6): 891-899.

43. Georgakis C. On the Use of Extensive Variables in Process Dynamics and Control. *Chem. Eng. Sci.* 1986; 41(6): 1471-1484.

44. Hoang NH, Dochain D, Ydstie BE. Partial inventory control of the CSTR via reaction-dependent generalized inventories. *IFAC Proc. Vol.* 2014; 47(3): 9123-9128.