Adaptive Control for Time-Varying Systems: Congelation and Interconnection

Kaiwen Chen
M.Sc.

Supervisor: Professor Alessandro Astolfi

A Thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy
of Imperial College London

Department of Electrical and Electronic Engineering
Imperial College London

Statement of Originality

The contents in this thesis, unless stated otherwise and appropriately referenced, are my own work.

Kaiwen Chen

Copyright Declaration

Abstract

This thesis investigates the adaptive control problem for systems with time-varying parameters. Two concepts are developed and exploited throughout the thesis: the congelation of variables, and the active nodes.

The thesis first revisits the classical adaptive schemes and explains the challenges brought by the presence of time-varying parameters. Then, the concept of congelation of variables is introduced and its use in combinations with passivity-based, immersion-and-invariant, and identification-based adaptive schemes are discussed. As the congelation of variables method introduces additional interconnection in the closed-loop system, a framework for small-gain-like control synthesis for interconnected systems is needed.

To this end, the thesis proceeds by introducing the notion of active nodes. This is instrumental to show that as long as a class of node systems that possess adjustable damping parameters, that is the active nodes, satisfy certain graph-theoretic conditions, the desired small-gain-like property for the overall system can be enforced via tuning these adjustable parameters. Such conditions for interconnected systems with quadratic, nonlinear, and linearly parametrized supply rates, respectively, are elaborated from the analysis and control synthesis perspectives. The placement and the computation/adaptation of the damping parameters are also discussed.

Following the introduction of these two fundamental tools, the thesis proceeds by discussing state-feedback designs for a class of lower-triangular nonlinear systems. The backstepping technique and the congelation of variables method are combined for passivity-based, immersion-and-invariance, and identification-based schemes. The notion of active nodes is exploited to yield simple and systematic proofs.

Based on the results established for lower-triangular systems, the thesis continues to investigate output-feedback adaptive control problems. An immersion-and-invariance scheme for single-input single-output linear systems and a passivity-based scheme for nonlinear systems in observer form are proposed. The proof and interpretation of these results are also based on the notion of active nodes. The simulation results show that the adaptive control schemes proposed in the thesis have superior performance when compared with the classical schemes in the presence of time-varying parameters.

Finally, the thesis studies two applications of the theoretical results proposed. The servo control problem for serial elastic actuators, and the disease control problem for interconnected settlements. The discussions show that these problems can be solved efficiently using the framework provided by the thesis.

Acknowledgment

Upon the completion of this thesis, I would like to sincerely thank my PhD supervisor, Prof Alessandro Astolfi, for his consistent patience and help with my research work. Working with him, I have never felt like being pushed or pulled along my path. Instead, I have been receiving trust, inspiration, and encouragement from him, all the time throughout these years, and all of these guided me through my PhD study in a natural and pleasant way. Without him, not a single page of this thesis can be done or, at least not with its current quality.

I would also like to thank my academic brothers and sisters: Adam, Anastasia, Jianli, and Joel. Though we did not work on the same topic, the discussions with you were always interesting and inspiring. The time we spent together with Alessandro, sometimes remotely, has become an unforgettable memory filled with happiness.

I need to particularly express my gratitude to the experts in relevant areas, who have generously provided their comments and suggestions to my work. They are Prof Thomas Parisini, Prof Andrea Serrani, Prof Romeo Ortega, Prof Shuzhi Sam Ge, Dr Giordano Scarciotti, Dr Yangyang Chen, Dr Dongjun Wu, Dr Bowen Yi, and Dr Lei Wang. I also do appreciate the feedback from the people whose name I cannot recall, whether their feedback was constructive or not.

Huge thanks to my friends: Junhong, Guitao, Tianwei, Yijun, Xu, Baoru, Andy... Your company, regardless of the distance, has always been heartwarming.

Love to my family: though we are thousands of miles apart, our hearts have been together all the time.

Last but not least, thanks to the people who selflessly kept the world running and

connected during the pandemic in the past two years, saving it from stalling and falling apart. This was one of the darkest times of humankind, but you made it the brightest time of humanity. I dedicate my thesis to you, and to the people who did not manage to see the dawn.

Contents

List of Figures

Abbreviations

DAG: Directed Acyclic Graph

FVS: Feedback Vertex Set

I&I: Immersion and Invariance

iISS: integral Input-to-State Stable/Stability

ISS: Input-to-State Stable/Stability

K-Filter: Kreisselmeier Filter

SEA: Series Elastic Actuator

SIQS: Susceptible-Infectious-Quarantined-Susceptible

SISO: Single-Input Single-Output

Notation

This thesis uses standard notation unless stated otherwise. Most symbols are locally defined and reused if no confusion is caused in the context. There are however some conventions used throughout the thesis, listed by category as follows.

Vectors and Matrices

For an n-dimensional vector $v \in \mathbb{R}^n$, $\underline{v_i} \in \mathbb{R}^i$, $1 \leq i \leq n$, denotes the vector composed of the first i elements of v; e_i denotes the ith unit vector in $\mathbb{R}^n$, that is, the vector in which the ith element is 1 and the other elements are 0. For an $n \times m$ matrix M, $(M)_i$ denotes the ith column; $(M)_{ij}$ denotes the ith element on the jth column; $\mathrm{tr}(M)$ denotes the trace; $|M|_{\mathrm{F}} \triangleq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}(M)_{ij}^2}$ denotes the Frobenius norm. The symbol $\otimes$ denotes the Kronecker product; I and S denote the identity matrix and the upper-shift matrix, respectively, sometimes with a subscript n indicating that $I_n \in \mathbb{R}^{n\times n}$ or $S_n \in \mathbb{R}^{n\times n}$; $\mathit{1}$ and $\mathit{0}$ denote the all-one matrix and the all-zero matrix, respectively, sometimes with a subscript $n\times m$ indicating that $\mathit{1} \in \mathbb{R}^{n\times m}$ or $\mathit{0} \in \mathbb{R}^{n\times m}$; $M > 0$ means that M is element-wise positive and similarly for other inequality signs. If M is a symmetric matrix, namely $M = M^\top$, $M \succ 0$ means M is positive definite and similarly for other curved inequality signs. $|v|$ denotes the Euclidean 2-norm; $|v|_M \triangleq \sqrt{v^\top M v}$, $M = M^\top \succ 0$, denotes the weighted 2-norm of the vector v with weight M;

Time-Varying Signals

For an n-dimensional signal $\theta : \mathbb{R} \to \mathbb{R}^n$, the image of which is contained in a compact set Θ, $\Delta_\theta : \mathbb{R} \to \mathbb{R}^n$ denotes the deviation of θ from a constant value ℓ_θ, *i.e.* $\Delta_\theta(t) \triangleq \theta(t) - \ell_\theta$;

$\delta_\theta \in \mathbb{R}$ denotes the supremum of the 2-norm of θ, *i.e.* $\delta_\theta \triangleq \sup_{t \geq 0} |\theta(t)| \geq 0$; and $\theta^{(i)}$ denotes the ith time derivative of $\theta(t)$, namely, $\dfrac{\mathrm{d}^i\theta}{\mathrm{d}t^i}$, assuming it exists.

Graphs

In a directed graph $\mathcal{G}$, $\mathcal{P}_i$ and $\mathcal{S}_i$ denote the index set of the direct predecessors and successors, respectively, of the vertex i. $|\cdot|$ denotes the cardinality of such a set.

Functions and Mappings

All functions and mappings, unless stated otherwise, are smooth. The operator "$\circ$" denotes function composition, for example, for functions $\alpha : \mathbb{R} \to \mathbb{R}$, $\beta : \mathbb{R} \to \mathbb{R}$, and a constant $c \in \mathbb{R}$, the expression "$\alpha \circ c\beta$" means "$\alpha\big(c\beta(\cdot)\big)$".

Chapter 1

Introduction

This thesis discusses adaptive control schemes for systems with time-varying parameters that the author has developed in the past four years, mainly based on two concepts: *congelation of variables* and *active nodes*. These two concepts will be further elaborated on in the remainder of the thesis. As the thesis mainly focuses on adaptive control, it is natural and helpful for understanding the context of the thesis to revisit classical adaptive control methods designed for systems with constant parameters.

1.1 Revisiting Classical Adaptive Control

Adaptive control, as the name suggests, is a control method for coping with unknown and varying plants and environments. The early development of adaptive control (see *e.g.* [98, 113, 136]) was motivated by flight control problems for aircraft to be operated in a large region within their flight envelope and with desirable performance, in which case a fixed-gain controller cannot work well. Although lacking rigorous proofs of stability, these results advocated the idea of dynamically updating the controller gains using the system output, that is, to "adapt" the controller for the present operating condition. Since the 1980s global stability and convergence of adaptive control systems have been established in the seminal works [42, 93], giving a solid foundation to the development of adaptive control theory. The literature on adaptive control has become vast since then and has been supported by rigorous results. The mainstream methods have been summarized and elaborated in books and monographs, see *e.g.* [5, 47, 51, 75, 76, 92, 126] and references therein.

What Is Adaptive Control?

To define what is adaptive control, consider the system

$$\begin{aligned} \dot{x} &= f(x, \theta, u), \\ y &= h(x, \theta, u), \end{aligned} \tag{1.1}$$

where $x(t) \in \mathbb{R}^n$ is the state; $u(t) \in \mathbb{R}^m$ is the input; $y(t) \in \mathbb{R}^p$ is the output; $\theta(t) \in \mathbb{R}^q$ is the vector of system parameters, typically unknown; $f : \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^n$ and $h : \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^n$ are mappings that describe the state evolution and the system output, respectively. The objective of the control problem is to let y track a given reference y_r asymptotically, that is, to achieve reference tracking or set-point regulation, depending on whether y_r is time-varying or constant, respectively. An adaptive controller is a dynamic feedback controller designed to complete this task without requiring the knowledge of θ, as illustrated in Fig. 1.1. The controller can further be expressed by the equations

$$\dot{\xi} = w(\xi, y, d), \tag{1.2}$$

$$u = v(\xi, y, d), \tag{1.3}$$

where $\xi(t) \in \mathbb{R}^{n_C}$ is the controller state; d stands for y_r and its time derivatives; $w : \mathbb{R}^{n_C} \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^{n_C}$ and $v : \mathbb{R}^{n_C} \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^m$ are mappings that describe the controller state update law[1], called "adaptation" in Fig. 1.1, and the control law, respectively. The key part of the adaptive controller is the dynamic "adaptation" part (1.2). The most common and intuitive interpretation of this part is that it provides an "estimate" for

[1]Sometimes (1.2) is said to depend also on u. It is a common misunderstanding that this indicates the need for measuring input signal. In practice measuring the input signal is unnecessary as it is determined by the control law, a function of the controller state, the system output, and the reference. The "dependency" on u means a direct use of the control law to update ξ and does not change the actual dependency at all. Therefore in (1.2), u is not listed among the arguments to avoid confusion.

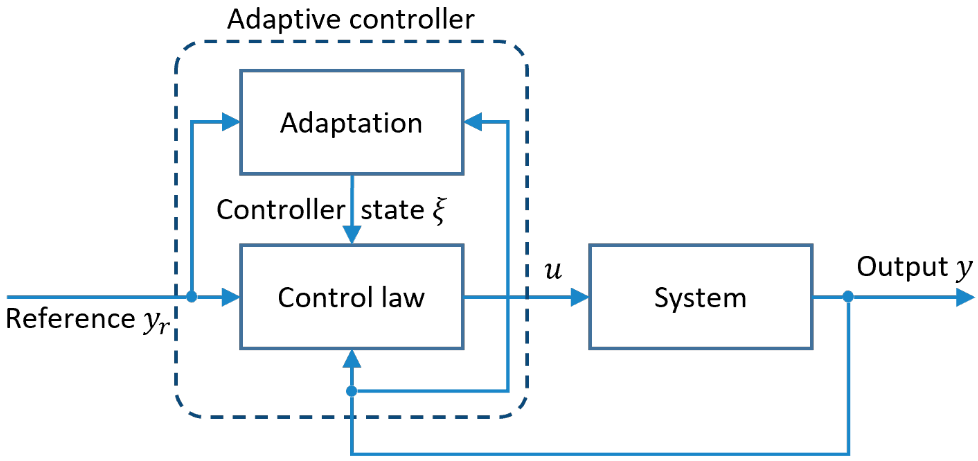

Figure 1.1: Schematic representation of an adaptive controller in a closed-loop system.

the unknown system parameters, which is intuitively correct, but may be misleading in general. The interpretation is partially correct since, in some cases, the controller state ξ (or a part of it) can be used in a non-implementable control law parametrized by unknown parameters, as a substitute for the unknown parameters, therefore making the control law implementable. The motivation behind is that the control law parametrized by the "parameter estimates" can be considered "equivalent" to the control law parametrized by the true parameters, if the parameter update law can let the "parameter estimates" converge to the true parameters and establish "certainty", which yields the well-known *certainty-equivalence principle*[2] . The interpretation is misleading since 1) in most cases these "parameter estimates"[3] do not converge to the true parameters; 2) ξ can include more "estimates" (as in the so-called overparametrization, see *e.g.* [60, 61]) or less "estimates" (as in dynamic high gains, see *e.g.* [79,80]) than the true parameters to overcome structural limitations; and 3) ξ may also contain compensation terms for unmeasured system state, or auxiliary variables facilitating parameter estimation. In the light of this, an adaptive

[2]The term "certainty-equivalence" or "certainty-equivalent" originates from the optimal and stochastic control literature. The first known use of the term is in [116], indicating the substitution of the unconditional expectation of the state for the state measured with uncertainty in the decision-making policy and treating it as equivalent to the policy in the deterministic case.

[3]Though the term "parameter estimate" may not be completely precise, to maintain consistency with existing literature, the controller state variables that are related to system parameters are still denoted with this term in the rest of the thesis. This should, however, not hinder the understanding of the true roles of these variables.

controller is not a controller that controls the system by estimating system parameters, but a controller dynamically adapting its own state based on measured data such that the control law fulfils the control objective without the knowledge of the system parameters.

The key features characterizing adaptive control are "dynamic adaptation" and "without knowledge of parameters", which distinguish adaptive control from other schemes mitigating the effects of parametric uncertainty. In contrast, control schemes like gain scheduling do not require a dynamical system for adaptation (described by (1.2)). They, instead, admit a set of gain-combinations for different operating conditions, and switch the controller gains to the prescribed combination once certain conditions hold. Furthermore, control schemes like robust control require the knowledge of nominal system parameters and the performance is affected by how the nominal values are close to the true values.

Taxonomy

Though adaptive control is not characterized by parameter estimation, the thesis focuses on the schemes that relate at least part of the controller state to the system parameters. Among various methods for categorizing adaptive control, the thesis uses the one based on parameter update laws. In this framework, most adaptive control research can be categorized into three trends: passivity-based schemes, identification-based schemes, and immersion-and-invariance schemes. Passivity-based schemes are also known as the Lyapunov-based schemes. The general idea of these schemes can be demonstrated by the following simple example. Consider the scalar system (that is, $x(t) \in \mathbb{R}$ and $u(t) \in \mathbb{R}$)

$$\dot{x} = \theta x^2 + u, \tag{1.4}$$

and the dynamic feedback controller

$$\dot{\hat{\theta}} = x^3, \tag{1.5}$$

$$u = -x - \hat{\theta} x^2, \tag{1.6}$$

where θ is an unknown constant parameter. Differentiating the functions $V_1 \triangleq \frac{1}{2}x^2$, $V_2 \triangleq \frac{1}{2}(\hat{\theta}-\theta)^2$, and $V \triangleq V_1 + V_2$ along the trajectories of the closed-loop system yields

$$\dot{V} = \dot{V}_1 + \dot{V}_2 = -x^2 - (\hat{\theta}-\theta)x^3 + (\hat{\theta}-\theta)x^3 = -x^2 \leq 0, \tag{1.7}$$

which proves that V is a valid Lyapunov function. Meanwhile, the cancellation of the $(\hat{\theta}-\theta)x^3$ terms can be interpreted as resulting from the interconnection in negative feedback of two passive systems, with storage functions V_1 and V_2, respectively. These two observations also explain why the scheme is "Lyapunov-based" or "passivity-based". In more general situations, the selection of V may not yield a Lyapunov function (therefore sometimes referred to as "Lyapunov-like" functions, see *e.g.* [50]), but the passivity interpretation always works. Therefore the term "passivity-based scheme" will be used throughout the thesis. A key feature of passivity-based schemes is that the parametric model used for designing the parameter update law is a differential model, say, for the aforementioned example the differential parametric model is

$$\dot{x} = -x - (\hat{\theta}-\theta)x^2. \tag{1.8}$$

The schemes in this "family" adopt more complex structures for additional considerations (*e.g.*, to achieve modularity [74]), but the general ideas are essentially similar.

Another trend is given by identification-based (or parameter-estimation-based) schemes, which considers an algebraic parametric model obtained by filtering and adopts linear regression algorithms to design parameter update laws. Consider again (1.4) with the control law

$$u = -k(x)x - \hat{\theta}x^2, \tag{1.9}$$

and the filters

$$\begin{aligned} \dot{\omega}_0 =& -k(x)\omega_0 - \hat{\theta}x^2, \\ \dot{\omega} =& -k(x)\omega + x^2. \end{aligned} \tag{1.10}$$

where k is a strictly positive damping term. The filters allows writing $x = \omega_0 + \omega\theta + \varepsilon$, where ε is such that $\dot{\varepsilon} = -k(x)\varepsilon$, hence converges asymptotically to zero. Then by defining $\hat{x} = \omega_0 + \omega\hat{\theta}$, one can further obtain an algebraic parametric model

$$\tilde{x} \triangleq x - \hat{x} = \omega(\theta - \hat{\theta}) + \varepsilon. \tag{1.11}$$

The algebraic parametric model allows solving the parameter estimation problem as a linear regression problem with methods available in system identification literature (see, *e.g.*, [44] for discrete-time identification schemes and, *e.g.*, [51] for continuous-time counterparts). One of the advantages of identification-based schemes is that the "estimates" of the parameters can converge to the true parameters under a condition known as *persistence of excitation*[4] (see, *e.g.*, [92, Chapter 6]). The condition of *persistence of excitation* can be enforced in some reference tracking problems by designing specific reference signals [10, 11], but it cannot be guaranteed nor verified *a priori* in general. It should also be emphasized that though the initial motivation of adaptive control was based on the "intuition" that the convergence of estimates to the true parameters makes the adaptive controller *certainty-equivalent* to a stabilizing true-parameter controller, this is in general not true, see the simple counterexample provided in [75, Section 1.2.1]. Most rigorous proofs for identification-based schemes exploit the so-called *swapping* lemma [90], in which the convergence of $\tilde{x}$ rather than that of $\theta - \hat{\theta}$ is required, and therefore do not require *persistence of excitation.* The general idea is to first establish $\mathcal{L}_2$-stability for subsystems in the analysis and then propagate this property to the overall system by exploiting the cascaded interconnection of the subsystems.

The *immersion-and-invariance* (I&I) schemes [6, 61, 62, 64, 65], which form the third trend, adopt parameter estimates inspired by state estimates of reduced-order observers, see, *e.g.*, [124]. The parameter estimate is not directly updated, and instead, it is composed of a dynamically updated part $\hat{\theta}$ and a state-dependent part $\beta(x)$. The resulting controller

[4] Such a condition can be further relaxed using the so-called *Dynamic Regressor Extension and Mixing* (DREM) method [4, 97].

for (1.4) is

$$\dot{\hat{\theta}} = -x^2\Big(\big(\hat{\theta}+\beta(x)\big)x^2+u\Big), \tag{1.12}$$

$$u = -x-\big(\hat{\theta}+\beta(x)\big)x^2, \tag{1.13}$$

where $\beta(x) \triangleq \frac{1}{3}x^3$. Then if one defines the parameter estimation error as $z_\theta = \hat{\theta}-\theta+\beta(x)$, the closed-loop system dynamics can be described by

$$\dot{x} = -x-x^2 z_\theta, \tag{1.14}$$

$$\dot{z}_\theta = -x^4 z_\theta. \tag{1.15}$$

It is not difficult to see that $x^2(t)z_\theta(t) \in \mathcal{L}_2$ and the subsystem (1.14) is finite-gain $\mathcal{L}_2$-stable with input $x^2 z_\theta$ and output x, from which stability properties of the overall system can be concluded due to the cascaded interconnection. Although the I&I scheme adopts a differential parametric model, the parameter estimation error dynamics described by (1.15) possess certain stability properties under conditions similar to *persistence of excitation*. Similarly to the case in identification-based schemes, such conditions can be enforced in some reference-tracking tasks by properly selecting the reference signal [133, 134], but in general cannot be guaranteed *a priori* in stabilization problems.

It is worth noting that there are other categorizations that are even more popular than the one stated above. Categorized based on the control law design/synthesize method, there are the so-called *model reference adaptive control* (see, *e.g.*, [36,89,121]) and *adaptive pole placement control* (also known as *self-tuning control*, see, *e.g.*, [13,34,94]). The former one is an adaptive version of the tracking controller designed by pole-zero cancellation and therefore only applicable to minimum-phase systems. It has, however, some advantages as the output of the system tracks the output of a reference model, the transient behaviours of which are known and can be tuned. *Adaptive pole placement control* schemes can be used for nonminimum-phase systems though there is no reference model to provide a preview of the transient behaviours and an intermediate model for tuning. Categorized based on the computation of controller parameters, there are *direct* adaptive control and *indirect* adaptive control. In the *direct* adaptive control schemes [40,49], a parametrized control law is first derived and the controller parameters are directly updated by (1.2). In the *indirect* schemes [135, 140], the estimates for the system parameters are updated by (1.2). The controller parameters are solved from an algebraic equation parametrized by the system parameter estimates, which only allow implicit expression of the controller parameters and are therefore called "indirect". These categorization methods are common in the adaptive control literature and comprehensive surveys can be found in [51, 127]. The reason why these popular categorization methods are not considered in this thesis is that the thesis uses a standard nonlinear framework to study adaptive control problems. Notions like transfer functions, poles, and zeros that rely on linearity and time-invariance are not exploited throughout the thesis (although these terms are used to facilitate the discussion wherever convenient). In addition, in a nonlinear framework, the design of the control laws and of the adaptation laws are typically coupled and performed in a recursive manner. Therefore typically one cannot design a parametrized controller and directly set its parameters using a separate update law, nor indirectly compute the controller parameters from an algebraic equation. All of these make the categorization by parameter update laws more suitable in the context of the thesis.

1.2 Adaptive Control for Time-Varying Systems

An important caveat in classical adaptive control schemes is that most of the supporting proofs are only valid when the system parameters are constant. In the passivity-based, identification-based, or I&I schemes, the parameter estimation error terms, which contain the system parameter θ, are used in the candidate Lyapunov functions or Lyapunov-like functions. As a result, when differentiating the candidate Lyapunov function, the time derivatives of θ are injected into the system dynamics, depriving them of either passivity or $\mathcal{L}_2$-stability properties that hold in the constant-parameter scenarios and therefore voiding the stability proofs. The use of adaptive control is concisely summarized in *Astrom's flowchart* [7, Section 1.6, Fig. 1.22], which can be interpreted as follows: for time-invariant plant/process, one should consider controllers with constant parameters; for time-varying systems with "predictable" (known) parameters, one should consider gain-scheduling schemes; and only when the parameters are both time-varying and "unpredictable" (unknown), one should consider adaptive control. In this sense, the assumption of constant parameters is favourable to stability proof, however, somehow making adaptive control deviate from its initial aspiration and guidelines of use.

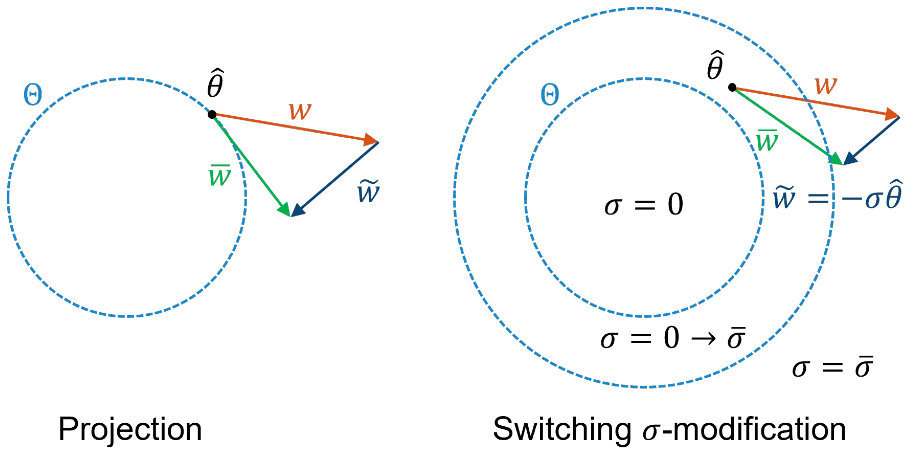

Figure 1.2: Illustration of the projection operation and of the switching σ-modification.

Several attempts have been made to solve the time-varying parameter issue since the 1980s. Some pioneering works on adaptive control for time-varying systems (see, *e.g.*, [41]) exploit *persistence of excitation* to guarantee stability by ensuring that the parameter estimates converge to the true parameters[5]. Subsequent works (see, *e.g.*, [73], [88]) have removed the restriction of *persistence of excitation* by requiring bounded and slow (in an average sense) parameter variations. Since the time derivative of θ is typically coupled with the parameter estimation error $\hat{\theta}-\theta$, when the system parameters are varying with bounded rates, if one can guarantee boundedness of the parameter estimates $\hat{\theta}$, the effects of time-varying parameters can be viewed as that of bounded disturbances. To this end, two widely used techniques are the *projection operation* (see, *e.g.*, [43, 102]) and the *switching σ-modification* (see, *e.g.*, [49, 51]). These methods assume that the true

[5]Methods that require identification of true parameters as a prerequisite, however, are considered to be more identification-oriented due to the need for excitation conditions and less pursued in control-oriented works, due to the reason explained in Section 1.1, though there are indeed identification methods suitable for estimating time-varying parameters and can achieve convergence in finite time (see *e.g.* [106, 109]).

parameter vector θ is confined to a convex set Θ (this set does not have to be tight so the simplest choice is a ball centred at 0). The projection operator is an identity operator when $\hat{\theta}$ belongs to the interior of Θ, otherwise, it alters the update law $\dot{\hat{\theta}} = w$ by an additional term $\tilde{w}$ once the parameter estimate vector $\hat{\theta}$ hits the boundary of Θ from inside and "projects" w into a new direction $\bar{w}$ tangent to the boundary of Θ. The *switching σ-modification* introduces a leakage term $-\sigma(\hat{\theta})\hat{\theta}$ to the update law w. The leakage rate $\sigma(\hat{\theta})$ is a continuous function that remains 0 when $\hat{\theta}$ is in Θ, yet continuously "switches" from 0 to $\bar{\sigma}$ as $\hat{\theta}$ moves away from Θ. A graphic illustration of the two update law modifications is given in Fig. 1.2. These modifications guarantee boundedness of the parameter estimation error, and only add non-positive terms to the derivative of the Lyapunov function, which does not undermine the results guaranteed by the unmodified update law $\dot{\hat{\theta}} = w$.[6] Exploiting these modified update laws, if the parameter variations are slow in some senses, bounded tracking error can be achieved (see, *e.g.*, [73,88,129,132,145]), and if the rates of parameter variations satisfy some integrability conditions, asymptotic tracking can be achieved (see, *e.g.*, [84–86]). However, these methods in general cannot achieve asymptotic results when the parameters are unknown and persistently varying, and fast varying parameters can further reduce the guaranteed performance, as the time derivatives of the parameters are taken into account in the analysis.

It is difficult to avoid differentiating the system parameters in schemes that exploit system parameter estimates for controller adaptation, which inevitably causes the aforementioned issues. Therefore, in some works, instead of using the classical controller structure developed in the spirit of the *certainty-equivalence principle*, the controller structure is modified in a such a way that a nominal constant-parameter controller can robustly dominate the parametric uncertainties, regardless if the parameters are constant or time-varying. It only remains to substitute adaptively updated estimates for these nominal constant parameters, using classical adaptive control techniques. These methods exploit either dynamically updated control gains (see, *e.g.*, [79, 80, 131]), or sliding-mode-like "switching" terms multiplied by dynamically updated amplitudes (see, *e.g.*, [48, 143]).

[6] In contrast, the original non-switching *σ-modification* [50] adopts a constant $\sigma = \bar{\sigma}$, which guarantees boundedness of estimates yet causes loss of asymptotic properties.

These methods can guarantee bounded tracking error if using continuously "switching" terms. Asymptotic results can be achieved by making the "switching" terms discontinuous or "asymptotically" discontinuous [48, 143], but in this case the classical concerns about discontinuous control law appears, such as "chattering" and lack of differentiability for recursive design, which undermines the significance of the asymptotic results. It is worth noting that the recent work [100] considers the technique that is an adaptive counterpart of the second-order sliding mode control [137], and therefore alleviates the effects associated with a discontinuous control law.

Another trend dealing with partial-state feedback nonlinear control is also worth mentioning. Although these results do not mention "adaptive control" or "parameters" explicitly, the unmeasured state variables are *de facto* time-varying parameters, and therefore the resulting nonlinear control designs can be viewed as adaptive control designs. In these works the time-varying parameters θ, instead of being generated by the equation $\dot{\theta} = 0$, are generated by a dynamical system that is referred to as unmeasured dynamics or zero dynamics and compensated by update laws that contain the information of the parameter model. The designs in [37, 38] are essentially generalizations of the passivity-based adaptive design, and the results in [63, 64] are generalizations of the I&I adaptive scheme. Although such schemes require additional information on the parameters (that is, the dynamical model), the parameters themselves are still unknown. Therefore such a scheme does not violate the definition of adaptive control considered by the thesis. These methods can achieve asymptotic results and are especially useful in dealing with systems with periodic parameters, as these parameters can be either modelled or approximated by sinusoidal signals that can be generated by oscillator-type dynamical systems.

The method coping with time-varying parameters that is more relevant to the thesis is the so-called *congelation of variables* method, originally proposed in the author's M.Sc. thesis [15]. The general idea is to represent the time-varying parameters as the sum of constant unknown parameters and time-varying perturbations. The design problem is then divided into an adaptive control design coping with the constant unknown parameters and a robust control design coping with the time-varying perturbations. A significant advantage of this method is that it does not require the system parameter variations to be slow or vanishing to achieve the convergence of the state/output to 0. At the same time this scheme retains compatibility with most classical adaptive control schemes, in the sense that when the parameters are constant, the modified controller reduces to the associated classical controller. The *congelation of variables* method has been combined with passivity-based schemes [16,20] and the I&I schemes [17] using state feedback, in which the damping design can be conducted with the help of a recursive change of coordinates commonly known as *backstepping* (see, *e.g.*, [75]). For systems in which full-state measurement is not available, an output-feedback scheme has to be considered and auxiliary filters have to be used. As a result, interconnections among multiple subsystems have to be dealt with in both the design and the analysis steps. In contrast to the classical time-invariant scenarios, in which the subsystems are cascaded or isolated, the time-varying perturbation terms can connect subsystems in feedback loops. This requires a comprehensive small-gain-like design and analysis for the network-structured overall system, which turns out to be more complicated than the standalone design and analysis performed for classical schemes. The *congelation of variables* method has been applied to the output-feedback passivity-based schemes [18,22] and the I&I scheme [19]. Similarly, as identification-based schemes use auxiliary filters to obtain an algebraic parametric model, the time-varying perturbations also create cyclic interconnections between the subsystems involved in the design and analysis. The combination of the *congelation of variables* method with an identification-based scheme is discussed in [23]. The relaxation of the constant parameter assumption improves the flexibility of adaptive control: the adaptive control schemes developed in the spirit of *congelation of variables* have been applied to practical scenarios such as formation control over time-varying flowfield [28,29] and cyber-physical systems under time-varying sensor and actuator attacks [25].

Although it has been realized that time-varying perturbations create cyclic interconnections and require a small-gain-like framework for network systems to efficiently design the controller and to analyze stability properties, most of the aforementioned schemes developed using the *congelation of variables* method exploit constructive controller designs and stability analysis, which are complicated and not sufficiently flexible in selecting design parameters. The need for a more concise and more flexible small-gain-like control synthesis tool has further motivated the other cornerstone of the thesis to be discussed next.

1.3 Network Small-Gain Results

Dynamical systems interconnected in a network structure have seen extensive research since the second half of the last century, under the names of dynamical network systems and large-scale interconnected systems. In particular, stability analysis (see, *e.g.*, [57, 68, 115,122,130]) has been one of the main research areas since, as it is well known, even if each subsystem (node) possesses some stability properties when disconnected from the overall system (network), such properties are not necessarily preserved under interconnection.

The majority of the works that guarantee network stability properties via the preservation of node stability properties (see, *e.g.*, *input-to-state* stability (ISS) [119], *integral input-to-state* stability (iISS) [118], and *input-to-output* stability (IOS) [120]) under interconnection are based on small-gain theorems (see [56] for results based on trajectories and [58] for an equivalent Lyapunov interpretation). These results can be intuitively understood, from a signal perspective, as requiring that the signals be not amplified while flowing through the interconnection; and, from an energy perspective, as requiring that energy do not accumulate via the interaction between the node subsystems.

Compared to the early results focusing on single-cycle interconnection, many subsequent works have taken generic network structures into account to extend the small-gain results to large-scale interconnected systems. These works can be categorized into two trends. The works in the first trend (see, e.g., [30, 32, 33]) exploit a nonlinear counterpart of Perron-Frobenius Theorem [39] and conclude matrix-operator-based small-gain conditions. These results have recently been extended to dynamical networks with infinitely many nodes in [31, 67]. The matrix-operator-condition proposed in these works can be viewed as a generalization of the spectral radius condition in the linear case, and it is equivalent to the *simple contraction* condition (that is, the loop gain operator is smaller than identity) in the classical two-node case. It should however be noted that the matrix operator is composed of nonlinear functions and the conditions based on such an operator are in general difficult to check [57].

The other trend adopts graph-based small-gain conditions based on the loop gain along every cycle path in the network, which is referred to as the *cyclic* small-gain theorem [59]. In the second trend, [81, 82] use a max-type Lyapunov function construction and consider a max-type aggregation of the neighbours' inputs, which yields a *simple contraction* condition similar to the classical two-node case. [53] uses a sum-type Lyapunov function construction and considers sum-type supply rates, which also leads to a similar contraction condition but requires a decomposition of the gain functions in the node dissipation inequalities. This is because the gain functions in the sum-type dissipation inequality are associated with the edges of the underlying graph and each edge may be shared by multiple cycles. As a result, the loop-gain condition uses decomposed components of gain functions to avoid counting the effect of a gain function repeatedly. Although the decomposition increases the complexity, the sum-type formulation is preferred in some senses as it allows the smooth construction of the Lyapunov function and its algebraic form allows direct application of linear algebra theory when the gains are linear.

Most of the aforementioned works (and of course many works not listed here) first assume a specific stability property of the node system and then find the conditions that preserve the same stability property for the network system. As the stability properties need rigorous conditions to be established, introducing the conditions that particularly serve for stability properties in the beginning may add unwanted restrictions and further complicate the final results that are already complicated due to the network structure of the overall system. As pointed out in [54, 68], one can change the assumed stability properties by changing the assumptions on the gain functions in the dissipation inequality. One can also conclude, for example, asymptotic stability of the interconnected system without requiring ISS of the node systems [2], provided the node dissipation inequalities are in a certain algebraic form. Moreover, there are cases in which no stability property is required at all, and one may only need a relaxed condition to allow invariance-like convergence analysis [111, 112].

Motivated by the need for a small-gain-like analysis and control synthesis in adaptive control problems for time-varying systems, the notion of the *active nodes* has been proposed in [21]. This framework focuses on the feasibility to construct an overall dissipation inequality for the network with negative supply rates by means of linear scaling. Though the results in [21] are limited to linearly parametrized supply rates, the same notion has been extended to a more general class of sum-type nonlinear supply rates in [26] by exploiting the nonlinear scaling techniques introduced in [52, 69]. These results provide feasibility conditions on the structure of the network interconnection such that once these are verified, the network dissipation inequality can be tuned to possess desired algebraic properties, without referring to stability properties explicitly, though adding proper restrictions to node dissipation inequalities can imply stability properties. The notion of *active nodes* has been used to interpret the output-feedback passivity-based adaptive control scheme in [24], the identification-based adaptive control scheme in [23], and the control synthesis and analysis steps for the adaptive coordinated attack rejection scheme in [25] for cyber-physical systems.

1.4 Organization and Contributions

The rest of the thesis is organized as follows. In Chapter 2, the challenges brought by time-varying system parameters are explained and the method of the *congelation of variables* is elaborated with a scalar nonlinear system. The combination of the proposed method with the passivity-based scheme, the I&I scheme, and the identification-based scheme are discussed. The fact that the time-varying perturbations cause cyclic interconnection between subsystems is demonstrated, and the need for a small-gain-like analysis/synthesis tool is discussed. The chapter covers the motivating ideas in the author's publications [16, 17, 20, 23].

In Chapter 3, the notion of *active nodes* is introduced motivated by the need for a small-gain-like tool. An analysis condition and a synthesis condition are discussed for three classes of network systems: the ones with quadratic supply rates, the ones with generic sum-type supply rates, and the ones with linearly parametrized supply rates. The chapter also discusses the control synthesis methods derived from the aforementioned results, which relate to the placement of the *active nodes* and the computation/adaptation of the *active node* parameters. The chapter covers the results in the author's publications [21, 22].

In Chapter 4, state-feedback adaptive control problems for a class of *lower-triangular* systems are considered to demonstrate the combination of the *congelation of variables* method with adaptive backstepping techniques. The passivity-base scheme, the I&I scheme, and the identification-based scheme are discussed, with a focus on how to overcome the structural limitation posed by the *unmatched* parametric uncertainty and the additional difficulties caused by parameter variations. The proofs based on *active nodes* are provided to facilitate the understanding of the proposed scheme. The chapter mainly covers the technical results in [17, 22, 23].

In Chapter 5, output-feedback adaptive control schemes are discussed. First, an I&I scheme for a class of single-input single-output (SISO) linear systems are discussed. Then a more complex SISO nonlinear system is considered with a passivity-based scheme. For both systems, the control schemes are explained using both classical constructive proofs and the small-gain-like framework based on the *active nodes*. The chapter mainly covers the technical results in [19, 22], and the conceptual proof in [24].

Chapter 6 discusses some potential applications of the proposed theoretical results. A robotic actuator is discussed to show how to model state-dependent nonlinearities with time-varying parameters and solve the new control problem using the *congelation of variables* method. Then, a disease control scheme is proposed to demonstrate how to control the spread of infectious diseases among interconnected settlements with minimum intervention by exploiting the notion of *active nodes*. The real-life examples in this chapter are based on the results in [17, 26].

The main contributions of the thesis are summarized as follows.

- The method called *congelation of variables* is proposed: this solves the long-standing constant-parameter restriction of classical adaptive control and is compatible with most existing adaptive control schemes.

- A small-gain-like framework based on the notion of *active nodes* is proposed. It provides flexible and simple tools to enforce a small-gain-like algebraic condition from a control synthesis perspective for networks of dynamical systems, without restricting itself to specific stability properties.
- The use of the *active-node* framework provides a standard routine for the analysis and the control synthesis for adaptive control problems for time-varying systems, the underlying networks of which possess cyclic and complex interconnections. The resulting interpretations yield better conciseness and readability when compared to classical constructive methods.

1.5 Fundamental Assumptions

To keep consistency throughout the thesis, several assumptions should be made before the discussion of the technical content. The time-varying parameters considered in this paper are assumed to satisfy one or more of the following conditions.

Assumption 1.1 (Bounded parameters)**.** *The parameter θ is piecewise continuous and $\theta(t) \in \Theta_0$, for all $t \geq 0$, where Θ_0 is a compact set. The "radius" of Θ_0,* i.e. *δ_{Δ_θ}, is assumed to be known, while Θ_0 can be unknown (see Fig. 1.3).*

Assumption 1.2 (Smooth bounded parameters)**.** *The parameter θ is smooth, that is, $\theta^{(i)}(t) \in \Theta_i$, for $i \geq 0$, for all $t \geq 0$, respectively, where Θ_i's are compact sets possibly unknown. δ_{Δ_θ} is assumed to be known.*

Assumption 1.3 (Sign-definite parameter)**.** *The parameter $b_m(t)$ is bounded away from 0 in the sense that there exists a constant ℓ_{b_m} such that $\operatorname{sgn}(\ell_{b_m}) = \operatorname{sgn}\big(b_m(t)\big) \neq 0$ and $0 < |\ell_{b_m}| \leq |b_m(t)|$, for all $t \geq 0$. The sign of ℓ_{b_m} and $b_m(t)$, for all $t \geq 0$, is known and does not change.*

The regressor functions, if not stated otherwise, satisfy the following condition.

Assumption 1.4. *The mapping $\phi : \mathbb{R}^n \to \mathbb{R}^q$ is a smooth mapping satisfying $\phi(0) = 0$.*

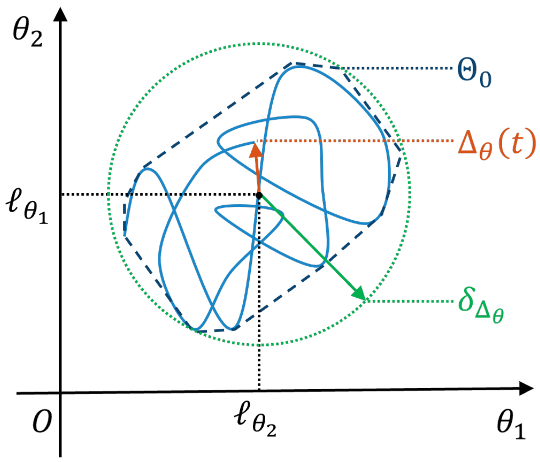

Figure 1.3: Graphical illustration of the role of Θ_0, ℓ_θ, $\Delta_\theta(t)$, and δ_{Δ_θ}.

Chapter 2

Congelation of Variables

In this chapter we discuss some simple scalar examples to understand the basic mechanism of the so-called *congelation of variables* method and how this helps to circumvent the restrictions due to the presence of time-varying system parameters. Though the spirit of the *congelation of variables* method is unique, due to some differences in aspects of implementation, in the following sections, the discussions are made separately in the context of three main-stream adaptive schemes: the passivity-based scheme, the immersion-and-invariance (I&I) scheme, and the identification-based scheme.

2.1 Passivity-Based Scheme

The passivity-based scheme is commonly known as Lyapunov-based scheme. This is partially due to the fact that the parameter estimator is designed by cancelling the parameter estimation error term that appears in the time derivative of a Lyapunov-like function along the system trajectories. Since such a Lyapunov-like function is not necessarily a Lyapunov function, whereas such cancellation can be precisely described by the interconnection of passive systems, we refer to such an adaptive scheme as a passivity-based scheme to highlight its nature.

2.1.1 Parameter in the Feedback Path

To begin with, consider a scalar nonlinear system described by the equation

$$\dot{x} = \phi(x)\theta(t) + u, \tag{2.1}$$

where $x(t) \in \mathbb{R}$ is the state, $u(t) \in \mathbb{R}$ is the input, $\theta(t) \in \mathbb{R}$ is an unknown time-varying parameter satisfying Assumption 1.1, and $\phi : \mathbb{R} \to \mathbb{R}$ satisfies Assumption 1.4 (with $n = q = 1$). To avoid using the unknown θ explicitly we introduce an "estimate" $\hat{\theta}$ of the parameter θ, and we rewrite (2.1) as

$$\dot{x} = \phi(x)\hat{\theta} + u + \phi(x)(\theta - \hat{\theta}). \tag{2.2}$$

In the passivity-based scheme the parameter update law for $\hat{\theta}$ is designed by considering a storage function of the form

$$V(x, \hat{\theta}, \theta) = \frac{1}{2}x^2 + \frac{1}{2\gamma_\theta}(\theta - \hat{\theta})^2. \tag{2.3}$$

Assuming that θ is differentiable with respect to time, for the time being, and taking the time derivative of V along the solutions of (2.2) yields

$$\dot{V} = x\phi(x)\hat{\theta} + xu + x\phi(x)(\theta - \hat{\theta}) - (\theta - \hat{\theta})\frac{\dot{\hat{\theta}}}{\gamma_\theta} + (\theta - \hat{\theta})\frac{\dot{\theta}}{\gamma_\theta}, \tag{2.4}$$

which implies that the selection of the parameter update law

$$\dot{\hat{\theta}} = \gamma_\theta x\phi(x) \tag{2.5}$$

cancels the effect of the $x\phi(x)(\theta - \hat{\theta})$ term that contains the unknown parameter. The constant $\gamma_\theta > 0$ is known as adaptation gain: it is not essential for stability analysis but plays an important role in controller tuning and it is therefore included hereafter. In classical adaptive control problems one assumes that θ is constant, that is $\dot{\theta}(t) = 0$ for all

$t \geq 0$, and selects the control law

$$u = -kx - x\phi(x)\hat{\theta}, \tag{2.6}$$

with $k > 0$, which yields $\dot{V} = -kx^2 \leq 0$. This leads to the classical adaptive regulation (to the origin) result stated as follows.

Proposition 2.1. *If the system parameter θ is constant, all trajectories of the closed-loop system described by* (2.1), (2.5), *and* (2.6), *are bounded and* $\lim_{t\to+\infty} x(t) = 0$.

Proof. To prove boundedness, note that V in (2.3) is positive definite and radially unbounded in $(x, \theta - \hat{\theta})$, and this, along with the fact that $\dot{V} \leq 0$, yields that V is a proper Lyapunov function and both $x(t)$ and $\hat{\theta}(t)$ are bounded for all $t \geq 0$.

To prove convergence, note that since $\dot{V} \leq 0$, the following inequality holds along the trajectories (with a slight abuse of notation that treats V as a function of time t):

$$\int_0^{+\infty} x^2(t)\mathrm{d}t \leq \frac{1}{k}\big(V(0) - V(+\infty)\big) \leq \frac{1}{k}V(0). \tag{2.7}$$

Note now that $\frac{\mathrm{d}}{\mathrm{d}t}x^2(t) = 2x(t)\dot{x}(t)$ and that both x and $\dot{x}$ are bounded. Hence by invoking Lemma A.4 (Barbalat's lemma), x^2 converges to 0 and so does x. The convergence part of the proof can also be completed by directly invoking Lemma A.5. □

When $\dot{\theta} \neq 0$, one has to deal with the indefinite term $(\theta - \hat{\theta})\frac{\dot{\theta}}{\gamma_\theta}$. One way to do this is to modify (2.5) with the so-called *projection operation* (see, *e.g.*, [43], [102]), which confines the parameter estimate $\hat{\theta}$ inside a convex compact set and therefore guarantees boundedness of $(\theta - \hat{\theta})$. It follows that boundedness of $\dot{\theta}$ guarantees boundedness of x (either exact boundedness, *e.g.* in [144] or boundedness in an average sense, *e.g.* in [88]), and $\dot{\theta} \in \mathcal{L}_1$ guarantees convergence of x to 0 (*e.g.* in [84], [85], [86]). In some other works (*e.g.* in [129], [142], [141]), boundedness of $\hat{\theta}$ is guaranteed by the so-called *switching σ-modification*, which adds some leakage to the integrator (2.5) if the parameter estimate drifts outside a reasonable region: this is often referred to as soft projection. All these schemes share the similarity that they treat $\dot{\theta}$ as a disturbance. As a result some disturbance attenuation effort is made to guarantee that bounded $\dot{\theta}$ yields bounded state/output regulation/tracking error, and sufficiently fast converging $\dot{\theta}$, which means that θ becomes constant sufficiently fast, guarantees the convergence of the error to 0. As a result, none of these methods can guarantee zero-error regulation/tracking when the unknown parameter is persistently time-varying, in which case $\dot{\theta}$ is non-vanishing.

The reason why we cannot avoid $\dot{\theta}$ in the analysis is the $\theta - \hat{\theta}$ term in (2.3). This term is included only to guarantee boundedness of $\hat{\theta}$, yet by no means guaranteeing convergence of $\hat{\theta}$ to θ, no matter whether θ is time-varying or constant, thus replacing θ with an unknown constant ℓ_θ that can guarantee the same properties (as we will see later, ℓ_θ is not required for implementation). In the light of this, consider the modified Lyapunov function candidate

$$V_\ell(x, \hat{\theta}, \ell_\theta) = \frac{1}{2}x^2 + \frac{1}{2\gamma_\theta}(\ell_\theta - \hat{\theta})^2. \tag{2.8}$$

Taking the time derivative of V_ℓ along the trajectories of (2.2) yields

$$\dot{V}_\ell = x\phi(x)\hat{\theta} + xu + x\phi(x)(\ell_\theta - \hat{\theta}) - (\ell_\theta - \hat{\theta})\frac{\dot{\hat{\theta}}}{\gamma_\theta} + x\phi(x)\Delta_\theta, \tag{2.9}$$

where $\Delta_\theta = \theta - \ell_\theta$. Comparing (2.9) with (2.4) we see that the substitution of ℓ_θ for θ eliminates the $\dot{\theta}$ term, at the cost of adding a perturbation term $x\phi(x)\Delta_\theta$ due to the deviation of θ from ℓ_θ. Note that due to Lemma A.3 and the fact that $\phi(0) = 0$, we can rewrite $\phi(x)$ as $\phi(x) = \bar{\phi}(x)x$, where $\bar{\phi} : \mathbb{R} \to \mathbb{R}$ is a smooth mapping. Consider now a new control law

$$u = -\left(k + \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}\right)x - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\bar{\phi}^2(x)x - \phi(x)\hat{\theta}, \tag{2.10}$$

where $\epsilon_{\Delta_\theta} > 0$ is a constant to balance the linear and the nonlinear terms. This control law along with the parameter update law (2.5) yields

$$\begin{aligned}\dot{V}_\ell &= -\left(k + \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}\right)x^2 - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\phi^2(x) + x\phi(x)\Delta_\theta \\ &\le -kx^2 \le 0. \end{aligned}\tag{2.11}$$

Note that the inequality is established by invoking Lemma A.1 (Young's inequality) and by noting that $x\phi(x)\Delta_\theta \leq \delta_{\Delta_\theta}|x||\phi(x)| \leq \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}x^2 + \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\phi^2(x)$.

Proposition 2.2. *Consider the system* (2.1)*, and the adaptive controller described by* (2.5) *and* (2.10)*, with the parameter θ satisfying Assumption 1.1. Then, all trajectories of the closed-loop system are bounded and $\lim\limits_{t\to+\infty} x(t) = 0$, regardless of the rate of variation of $\theta(t)$.*

Proof. Note that the dissipation inequality (2.11) is in the same form as that of the constant parameter case and we can therefore conclude boundedness of all trajectories of the closed-loop system as well as convergence of x to 0 using the same argument as in the proof of Proposition 2.1, without requiring any restriction on the rate of parameter variation. □

The method of substituting the constant ℓ_θ for the time-varying θ to avoid unnecessary time derivatives is called *congelation of variables* [16][1]. Since restrictions on the rate of parameter variations are not needed in the analysis, controllers designed via the *congelation of variables* method are naturally applicable to systems with fast-varying parameters.

Remark 2.1. *The control law* (2.10) *and the parameter update law* (2.5) *do not depend on ℓ_θ, in the same way as classical adaptive controllers do not depend on θ, thus showing the "adaptive" property of the proposed mechanism. One can interpret the proposed controller as a combination of an adaptive controller, to cope with the unknown parameter ℓ_θ, and a robust controller, to cope with the time-varying perturbation $\Delta_\theta(t)$. This fact can also be revealed by noting that, when θ is a constant, one could select $\ell_\theta = \theta$, hence $\delta_{\Delta_\theta} = 0$, and the control law* (2.10) *is reduced to the classical control law* (2.6)*.*

Remark 2.2. *Assumption 1.4 is only needed when the parameter θ is time-varying. When θ is time-varying, if $\phi(0) \neq 0$, one has $\dot{x} = \phi^\top(x)\theta + u = x\bar{\phi}(x)\theta + u + \phi(0)\theta$, where $\phi(0)\theta$*

[1] Some works predating [16] exploit similar ideas to avoid involving $\dot{\theta}$ in the analysis. For example, in [3] the unknown time-varying controller parameter in the Lyapunov function is replaced with a constant (0, as a matter of fact). In other works one first derives a constant parameter controller via dominance design (instead of directly using a time-varying parameter controller that cancels the time-varying parameter) and then estimates the constant parameter of the dominance controller, see *e.g.* [80], [131]. The decomposition of time-varying parameter into a constant parameter and a time-varying perturbation term is also seen in the online lecture notes [128].

is an unknown time-varying disturbance. This leads to a disturbance rejection problem that is not to be discussed in the thesis. The treatment for the case in which $\phi(0) \neq 0$ is discussed in [27].

Comparison with a Robust Control Scheme

One could agree that the *congelation of variables* scheme resembles a pure robust control scheme in the sense that they both use a nominal parameter, which is ℓ_θ in this case. It should be highlighted that in a robust control scheme ℓ_θ should be known in some sense and it is required for implementation, whereas in the *congelation of variables* scheme ℓ_θ is only needed for analysis. To illustrate this consider a practical scenario in which we have a circuit that has to work with one of three resistors with values 50 Ω, 100 Ω, and 150 Ω, yet which one is used is unknown. In addition, due to temperature variations, the resistances have a fluctuation of ± 10 Ω. In the spirit of the proposed method, ℓ_θ equals either 50 Ω, 100 Ω, or 150 Ω, which is unknown and not used in the controller design, as it is replaced by the dynamically updated $\hat{\theta}$, and $\delta_{\Delta_\theta} = 10$ Ω, which is known and used in the controller design. In the spirit of robust control, one has to determine the nominal resistance of the resistor before designing the controller, and according to the known information, the best guess is $\ell_\theta = 100$ Ω. In this case the maximum deviation from this nominal value is $\delta_{\Delta_\theta} = 60$ Ω, which is caused not only by the parameter variation but also by the imperfect knowledge of the true resistance of the resistor used. This leads to a more conservative design that uses an unnecessarily high gain and may cause severe noise amplification issues. In an extreme case in which the nominal resistance of the used resistor is completely unknown, one cannot design a robust controller while one can still design an adaptive controller using the proposed method.

On the Knowledge of δ_{Δ_θ}

It is natural to question whether it is practical to assume that δ_{Δ_θ} is known (Assumption 1.1), since $\Delta_\theta = \theta - \ell_\theta$ depends on the selection of ℓ_θ and so is δ_{Δ_θ} This is justified by the observation that in many practical scenarios (*e.g.* in the aforementioned resistor example), the amplitude of fluctuation (sometimes is also referred to as an error bound), instead of the mean value of the parameter, is known *a priori*, and therefore one only needs to estimate the nominal value ℓ_θ and apply the *congelation of variables* method to design the controller. In the case in which δ_{Δ_θ} is unknown, one can add an additional dynamical estimate for δ_{Δ_θ}, exploiting the fact that the nonlinear damping term in the control law (2.10) is linearly parametrized. In fact, consider the dynamic feedback controller

$$\begin{aligned}\dot{\hat{\theta}} &= \gamma_\theta x\phi(x),\\ \dot{\hat{k}}_2 &= \gamma_{k_2}\phi^2(x),\\ u &= -k_1x - \hat{k}_2\bar{\phi}^2(x)x - \phi(x)\hat{\theta},\end{aligned} \tag{2.12}$$

with $k_1 > 0$, $\gamma_{(\cdot)} > 0$. Consider a new construction of storage function $V = \frac{1}{2}x^2 + \frac{1}{2\gamma_\theta}(\ell_\theta - \hat{\theta})^2 + \frac{1}{2\gamma_{k_2}}(k_2 - \hat{k}_2)^2$, with $k_2 = \frac{1}{2k_1}\delta^2_{\Delta_\theta}$, $\epsilon_{\Delta_\theta} = \frac{\delta_{\Delta_\theta}}{k_1}$. Its time derivative along the closed-loop system trajectories satisfies

$$\begin{aligned}\dot{V} =& -\left(k_1 - \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}\right)x^2 - \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}x^2 - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\phi^2(x) + x\phi(x)\Delta_\theta\\ &+ (k_2 - \hat{k}_2)\phi^2(x) - \frac{1}{\gamma_{k_2}}(k_2 - \hat{k}_2)\dot{\hat{k}}_2\\ \leq& -\frac{1}{2}k_1x^2 \leq -kx^2 \leq 0,\end{aligned} \tag{2.13}$$

where $k \triangleq \frac{1}{2}k_1$. It is not difficult to see that the results in Proposition 2.2 still hold with this new controller since we end up with the same dissipation inequality. This shows that the knowledge of δ_{Δ_θ} is not a restriction and can be circumvented with simple modifications.

Remark 2.3. *It is worth introducing a convention to clarify the spirit in which we treat unknown quantities. If an unknown indefinite term in the time derivative of the Lyapunov function vanishes as the system parameters become constant, then this term is to be dominated by a static damping design, like the Δ_θ-term in this case, and we do not aim at estimating δ_{Δ_θ}, the bound of $\Delta_\theta(t)$. If an unknown indefinite term is not vanishing even when all system parameters are constant, like the ℓ_θ-term in this case, then this term is to be compensated by a dynamically updated "estimate", which is $\hat{\theta}$ in this case. The reasons for this convention of design are, first, that we do not want to over-extend the dimension* *of the closed-loop system by adding too many dynamic estimates, and second, that we need the static damping terms to counteract fast parameter variations for better transient performance (for the same reason one can use nonlinear damping techniques even for system with constant parameters).*

A Passivity Interpretation

Consider the classical adaptive control problem in which θ is constant. The closed-loop dynamics can be described via a negative feedback loop consisting of two passive systems, namely

$$\Sigma_1 : \begin{cases} \dot{x}_1 &= -kx_1 + \phi(x_1)u_1, \\ y_1 &= x_1\phi(x_1), \end{cases} \tag{2.14}$$

$$\Sigma_2 : \begin{cases} \dot{x}_2 &= \gamma_\theta u_2, \\ y_2 &= x_2, \end{cases} \tag{2.15}$$

where $x_1 \triangleq x$, $x_2 \triangleq \hat{\theta} - \theta$, $u_1 \triangleq -y_2$, $u_2 \triangleq y_1$. The storage functions are $S_1 = \frac{1}{2}x_1^2$ and $S_2 = \frac{1}{2\gamma_\theta}x_2^2$, respectively. It is well-known that the parameter update law (2.5) is not designed to guarantee convergence of $\hat{\theta}$ to θ, though $\hat{\theta}$ is called the parameter estimate by convention, but to make $\hat{\theta} - \theta$ an input/output signal to form a passive interconnection. When θ is time-varying, the dynamics of Σ_2 are described by

$$\Sigma_2 : \begin{cases} \dot{x}_2 &= \gamma_\theta u_2 - \dot{\theta}, \\ y_2 &= x_2, \end{cases} \tag{2.16}$$

which causes the loss of passivity from u_2 to y_2. The *congelation of variables* method can therefore be interpreted as selecting a new signal $\hat{\theta} - \ell_\theta$ that can yield a passive interconnection, while maintaining the passivity of Σ_1 by strengthened damping. Within this framework, the two passive systems are described by

$$\Sigma_1 : \begin{cases} \dot{x}_1 &= -a(x_1,t)x_1 + \phi(x_1)u_1, \\ y_1 &= x_1\phi(x_1), \end{cases} \tag{2.17}$$

$$\Sigma_2 : \begin{cases} \dot{x}_2 &= \gamma_\theta u_2, \\ y_2 &= x_2, \end{cases} \tag{2.18}$$

where $x_1 \triangleq x$, $x_2 \triangleq \hat{\theta} - \ell_\theta$, $u_1 \triangleq -y_2$, $u_2 \triangleq y_2$ and $a(x_1,t) \triangleq \left(k + \frac{1}{2\epsilon_{\Delta_\theta}}\delta_{\Delta_\theta}\right) + \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\bar{\phi}^2(x_1) - \Delta_\theta\bar{\phi}(x_1) \geq k > 0$.

2.1.2 Parameter in the Input Path

In what follows we show how to extend the idea of *congelation of variables* to systems in which a time-varying parameter is coupled with the input by considering the nonlinear system

$$\dot{x} = \phi(x)\theta(t) + b(t)u, \tag{2.19}$$

where the variables are defined as in (2.1) and in addition $b(t) \in \mathbb{R}$ satisfies both Assumption 1.1 and Assumption 1.3. Similarly as before, equation (2.19) can be re-written as

$$\begin{aligned} \dot{x} = {} & \phi(x)\hat{\theta} + \bar{u} + \phi(x)\Delta_\theta + \Delta_b\hat{\varrho}\bar{u} \\ & + (\ell_\theta - \hat{\theta})\phi(x) - \ell_b\left(\frac{1}{\ell_b} - \hat{\varrho}\right)\bar{u}, \end{aligned} \tag{2.20}$$

where $\Delta_b(t) = b(t) - \ell_b$, $\hat{\varrho}$ is an "estimate" of $\frac{1}{\ell_b}$, and $u = \hat{\varrho}\bar{u}$. From classical adaptive control theory (see *e.g.* [75]) we know that the effect of the second line of (2.20) can be cancelled by selecting the parameter update laws (2.5) and

$$\dot{\hat{\varrho}} = -\gamma_\varrho \mathrm{sgn}(\ell_b)\bar{u}x, \tag{2.21}$$

and considering the Lyapunov function candidate $V(x,\hat{\theta},\hat{\varrho}) = \frac{1}{2}x^2 + \frac{1}{2\gamma_\theta}(\theta-\hat{\theta})^2 + \frac{|\ell_b|}{2\gamma_\varrho}(\frac{1}{\ell_b} - \hat{\varrho})^2$, the time derivative of which along the trajectories of (2.20) satisfies

$$\dot{V} = x\phi(x)\hat{\theta} + x\bar{u} + x\phi(x)\Delta_\theta + \Delta_b\hat{\varrho}\bar{u}x. \tag{2.22}$$

Note that the perturbation term $\Delta_b\hat{\varrho}\bar{u}x$ depends on $\bar{u}$ explicitly, which means that we cannot dominate this term by simply adding damping terms to $\bar{u}$, as doing this also alters the perturbation term itself. Instead, we need to make $\Delta_b\hat{\varrho}\bar{u}x$ non-positive by designing $\bar{u}$ and selecting ℓ_b. Consider $\bar{u}$ as a feedback control law with a non-positive nonlinear gain, that is

$$\begin{aligned}\bar{u} &= -\left(k + \frac{1}{2}\left(\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \frac{1}{\epsilon_{\hat{\theta}}}\right) + \frac{1}{2}(\epsilon_{\Delta_\theta}\delta_{\Delta_\theta} + \epsilon_{\hat{\theta}}\hat{\theta}^2)\bar{\phi}^2(x)\right)x \\ &\triangleq -\kappa(x,\hat{\theta})x, \end{aligned} \tag{2.23}$$

where $\epsilon_{\hat{\theta}} > 0$. Note that $\kappa(x,\hat{\theta}) > 0$ by construction.

Proposition 2.3. *Consider the closed-loop system described by* (2.2)*,* (2.5)*,* (2.10)*, and* (2.21)*, with the initial condition of* $\hat{\varrho}$ *satisfying* $\hat{\varrho}(0)\operatorname{sgn}(\ell_b) > 0$*, and with* θ *satisfying Assumption 1.1 and* b *satisfying Assumption 1.1 and 1.3. Then, all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} x(t) = 0$*, regardless of the rates of parameters* $\theta(t)$ *and* $b(t)$*.*

Proof. First, substituting (2.23) into (2.21) yields $\dot{\hat{\varrho}} = \gamma_\varrho \operatorname{sgn}(\ell_b)\kappa x^2$. There are two cases to be discussed.

- When $b(t) > 0$, for all $t \geq 0$, due to Assumption 1.3, there exists a constant ℓ_b such that $0 < \ell_b \leq b(t)$, $\Delta_b > 0$, $\dot{\hat{\varrho}} \geq 0$, which means that any initialization with $\hat{\varrho}(0) > 0$ guarantees that $\hat{\varrho}(t) > 0$, for all $t \geq 0$, and therefore $\Delta_b\hat{\varrho}\bar{u}x = -\Delta_b\hat{\varrho}\kappa x^2 \leq 0$, for all $t \geq 0$.

- Similarly, when $b(t) < 0$, for all $t \geq 0$, there exists ℓ_b such that $b(t) \leq \ell_b < 0$, $\Delta_b < 0$, $\dot{\hat{\varrho}} \leq 0$. Then selecting $\hat{\varrho}(0) < 0$ guarantees $\hat{\varrho}(t) < 0$, for all $t \geq 0$, and $\Delta_b\hat{\varrho}\bar{u}x \leq 0$.

Recalling (2.22), (2.23), and noting that $\Delta_b \hat{\varrho} \bar{u} x \leq 0$ yields

$$\begin{aligned}\dot{V} \leq & -kx^2 - \left(\frac{\epsilon_{\hat{\theta}}}{2} \left(\phi(x)\hat{\theta}\right)^2 + \frac{1}{2\epsilon_{\hat{\theta}}} x^2 - x\phi(x)\hat{\theta} \right) \\ & - \left(\frac{\epsilon_{\Delta_\theta} \delta_{\Delta_\theta}}{2} \phi^2(x) + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}} x^2 + x\phi(x)\Delta_\theta \right) \leq -kx^2 \leq 0. \end{aligned} \tag{2.24}$$

This establishes boundedness of the system trajectories. Finally, noting boundedness of x, $\dot{x}$ and exploiting Lemma A.5 as before, convergence of x to 0 follows. □

Remark 2.4. *This example highlights the flexibility of the congelation of variables method: the congealed parameter $\ell_{(\cdot)}$ can be selected according to the specific usage. It can be a nominal value for robust design, or an "extreme" value to create sign-definiteness, as long as the resulting perturbation $\Delta_{(\cdot)}$ is considered consistently. One can make $\ell_{(\cdot)}$ a time-varying parameter subject to some of the assumptions used in the literature (e.g. $\dot{\ell}_{(\cdot)} \in \mathcal{L}_\infty$, $\dot{\ell}_{(\cdot)} \in \mathcal{L}_1$, see e.g. [84, 88]), and use the congelation of variables method to relax these assumptions. One can also exploit an oscillator-like dynamical system to generate $\ell_{(\cdot)}$ [27], which allows avoiding a conservative design with large $\delta_{\Delta_{(\cdot)}}$ when θ is a large-amplitude periodic signal. This is the reason why the proposed method is named "congelation"[2] not "freeze".*

A Passivity Interpretation

Write (2.20), (2.5), and (2.21) into the system described by

$$\dot{x} = f(x,t) + v, \tag{2.25}$$

$$\dot{\overbrace{\hat{\theta} - \ell_\theta}} = \gamma_\theta x \phi(x) \tag{2.26}$$

$$\dot{\overbrace{\hat{\varrho} - \ell_b^{-1}}} = \gamma_\varrho \operatorname{sgn}(\ell_b) \kappa x^2 \tag{2.27}$$

$$y = x. \tag{2.28}$$

where $f(x,t) \triangleq \phi(x)\hat{\theta} - \kappa x + \phi(x)\Delta_\theta + (\ell_\theta - \hat{\theta})\phi(x) + \ell_b(\ell_b^{-1} - \hat{\varrho})\kappa x$ and the input v is set to $-\Delta_b \hat{\varrho} \kappa y$. The interconnection among the subsystems (2.25)–(2.27) can be understood

[2] The word "congelation" is polysemous: it means both "coagulation" and "freeze/solidification" [87].

using the passivity interpretation in Section 2.1.2. Consider now the system consisting of (2.25)–(2.28) as a whole. One can observe that the selection of κ as in (2.23) renders the system passive from the input v to the output y with dissipation rate kx^2 (see Fig. 2.1 for a schematic illustration). At the same time, the selection of $\hat{\varrho}(0)$ and ℓ_b guarantees that $K(t) \triangleq \Delta_b(t)\hat{\varrho}(t)\kappa(t) > 0$, for all $t \geq 0$. Then, from a passivity perspective, the results in Proposition 2.3 are established by exploiting the fact that K can be arbitrarily selected from $(0, +\infty)$, to achieve robustness against the time-varying parameter $b(t)$.

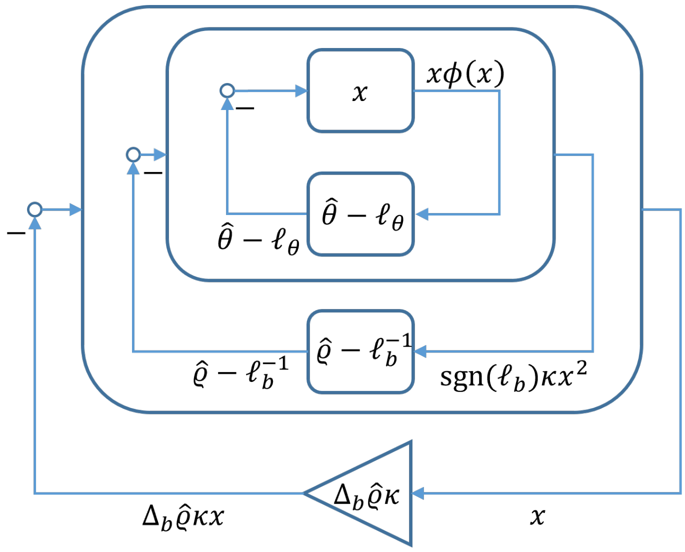

Figure 2.1: Schematic representation of system (2.20), (2.5) **and** (2.21) **as the interconnection of passive subsystems.**

2.2 Immersion-and-Invariance Scheme

In this section we discuss how to combine the *congelation of variables* method with the I&I scheme when the system parameter is time-varying. The core idea of the I&I scheme introduced in [5, 6], is to use a parameter estimate $\hat{\theta} + \beta(x)$ that consists of both a dynamically updated part $\hat{\theta}$ and a state-dependent static part $\beta(x)$, which is in contrast to the parameter estimate used in the passivity-based scheme, where the whole estimate term is dynamically updated. Such a scheme resembles the mechanism of reduced-order observers, which also uses the mixture of dynamically updated filter states and static measurable signals as estimate. Consequently, stability properties of the systems in the I&I scheme should be established via the study of a cascaded $\mathcal{L}_2$ interconnection, instead of a cyclic passive interconnection, similarly to the reduced-order state estimation problems. Due to this difference, our results for the passivity-based scheme do not hold directly without justifications, and further discussions are needed here.

Consider again the nonlinear system (2.1), namely the system $\dot{x} = \phi(x)\theta(t) + u$, with the same assumptions. Consider the dynamic feedback controller

$$\begin{aligned} u =& -k_1 x - k_2\bar{\phi}^2(x)x - \phi(x)\big(\hat{\theta} + \beta(x)\big), \qquad (2.29)\\ \dot{\hat{\theta}} =& -\gamma_\theta\phi(x)\Big(\phi(x)\big((\hat{\theta} + \beta(x)\big) + u\Big), \qquad (2.30)\end{aligned}$$

where $k_1 > 0$, $k_2 \geq 0$, and a possible selection of β is

$$\beta(x) = \gamma_\theta \int_0^x \phi(s)\mathrm{d}s, \qquad (2.31)$$

which is such that

$$\frac{\partial\beta}{\partial x} = \gamma_\theta\phi(x). \qquad (2.32)$$

To understand why the classical design fails to work in the presence of the time-varying parameter, let $k_2 = 0$. Then, substituting (2.29) and (2.30) into (2.1) and defining the so-called *off-the-manifold* variable $z_\theta \triangleq \hat{\theta} - \theta + \beta(x)$ yields the closed-loop system dynamics

$$\begin{aligned} \dot{x} =& -k_1 x + \phi(x)\theta - \phi(x)\big(\hat{\theta} + \beta(x)\big) = -k_1 x - \phi(x) z_\theta, \qquad (2.33)\\ \dot{z}_\theta =& \dot{\hat{\theta}} + \frac{\partial\beta}{\partial x}\big(\phi(x)\theta + u\big) - \dot{\theta} = -\gamma_\theta\phi^2(x)z_\theta - \dot{\theta}. \qquad (2.34)\end{aligned}$$

Consider two storage functions $V_x(x) = \frac{1}{2}x^2$ and $V_{z_\theta}(z_\theta) = \frac{1}{2}z_\theta^2$. It can be shown that their time derivatives along the system trajectories are

$$\begin{aligned} \dot{V}_x =& -k_1x^2 - x\phi(x)z_\theta \leq -\left(k_1 - \frac{1}{2\epsilon_{z_\theta}}\right)x^2 + \frac{1}{2}\epsilon_{z_\theta}\big(\phi(x)z_\theta\big)^2, \qquad (2.35)\\ \dot{V}_{z_\theta} =& -\gamma_\theta\big(\phi(x)z_\theta\big)^2 - \dot{\theta}z_\theta, \qquad (2.36)\end{aligned}$$

respectively. Selecting $V_{xz_\theta}(x, z_\theta) = V_x(x) + \frac{\epsilon_{z_\theta}+2}{2\gamma_\theta} V_{z_\theta}(z_\theta)$ yields

$$\begin{aligned}\dot{V}_{xz_\theta} \leq & - \left(k_1 - \frac{1}{2\epsilon_z}\right)x^2 + \frac{1}{2}\epsilon_{z_\theta}\big(\phi(x)z_\theta\big)^2 - \left(\frac{1}{2}\epsilon_{z_\theta} + 1\right)\big(\phi(x)z_\theta\big)^2 - \frac{\epsilon_{z_\theta}+2}{2\gamma_\theta}\dot{\theta}z_\theta \\ = & - kx^2 - \big(\phi(x)z_\theta\big)^2 - \frac{\epsilon_{z_\theta}+2}{2\gamma_\theta}\dot{\theta}z_\theta, \end{aligned} \tag{2.37}$$

where $k_1 = k + \frac{1}{\epsilon_{z_\theta}}$, with $k > 0$. When $\dot{\theta} = 0$, we have $\dot{V}_{xz_\theta} \leq -kx^2 - \big(\phi(x)z_\theta\big)^2$, with $k > 0$. Proceeding with standard Lyapunov and invariance-like analysis, we can conclude boundedness of closed-loop system trajectories, convergence of x to 0, and in addition, another property that a passivity-based scheme does not guarantee: $\phi(x)z_\theta \in \mathcal{L}_2$. One can also interpret the I&I scheme as finding an error signal z_θ such that the $\phi(x)z_\theta \in \mathcal{L}_2$ and feeding $\phi(x)z_\theta$ to the x-subsystem, which has a finite $\mathcal{L}_2$ gain from $\phi(x)z_\theta$ to the state x. This is, however, not true when θ is time-varying. Under the effect of $\dot{\theta}$, $\phi(x)z_\theta$ is no longer $\mathcal{L}_2$: we lose $\mathcal{L}_2$ property for a reason similar to that for the loss of passivity in the passivity-based scheme.

To restore the $\mathcal{L}_2$ property of the z_θ-subsystem, we follow the spirit of the *congelation of variables* method by substituting ℓ_θ for θ in the definition of the *off-the-manifold* variable, namely, re-defining $z_\theta \triangleq \hat{\theta} - \ell_\theta + \beta(x)$. Consider the closed-loop system with $k_2 > 0$, which yields

$$\dot{x} = - k_1 x - k_2\bar{\phi}^2(x)x + \phi(x)\Delta_\theta - \phi(x)z_\theta, \tag{2.38}$$

$$\dot{z}_\theta = \dot{\hat{\theta}} + \frac{\partial\beta}{\partial x}\big(\phi(x)\theta + u\big) = -\gamma_\theta\phi^2(x)z_\theta + \gamma_\theta\phi^2(x)\Delta_\theta. \tag{2.39}$$

Taking the time derivatives of the functions $V_x(x)$ and $V_{z_\theta}(z_\theta)$ along the trajectories of the closed-loop system, respectively, yields

$$\begin{aligned}\dot{V}_x = & - k_1x^2 - k_2\phi^2(x) + x\phi(x)\Delta_\theta - x\phi(x)z_\theta \\ \leq & - \left(k_1 - \frac{1}{2\epsilon_{z_\theta}} - \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}\right)x^2 - \left(k_2 - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\right)\phi^2(x) + \frac{1}{2}\epsilon_{z_\theta}\big(\phi(x)z_\theta\big)^2, \end{aligned} \tag{2.40}$$

$$\begin{aligned}\dot{V}_{z_\theta} = & - \gamma_\theta\big(\phi(x)z_\theta\big)^2 - \gamma_\theta\phi^2(x)z_\theta\Delta_\theta \\ \leq & - \gamma_\theta\left(1 - \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}\right)\big(\phi(x)z_\theta\big)^2 + \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\phi^2(x), \end{aligned} \tag{2.41}$$

Selecting $V_{xz_\theta}(x, z_\theta) = V_x(x) + cV_{z_\theta}(z_\theta)$, with the scaling coefficient $c \triangleq \frac{\epsilon_{\Delta_\theta}(\epsilon_{z_\theta}+2)}{\gamma_\theta(2\epsilon_{\Delta_\theta}-\delta_{\Delta_\theta})}$, which is positive provided $\epsilon_{\Delta_\theta} > \frac{1}{2}\delta_{\Delta_\theta}$, yields

$$
\begin{aligned}
\dot{V}_{xz_\theta} \leq & -\left(k_1 - \frac{1}{2\epsilon_{z_\theta}} - \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}\right)x^2 - \left(k_2 - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\right)\phi^2(x) + \frac{1}{2}\epsilon_{z_\theta}\big(\phi(x)z_\theta\big)^2 \\
& -\left(\frac{1}{2}\epsilon_{z_\theta} + 1\right)\big(\phi(x)z_\theta\big)^2 + \frac{1}{2}c\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\phi^2(x) \\
= & -\left(k_1 - \frac{1}{2\epsilon_{z_\theta}} - \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}\right)x^2 - \left(k_2 - \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta} - \frac{1}{2}c\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}\right)\phi^2(x) \\
& -\big(\phi(x)z_\theta\big)^2.
\end{aligned} \tag{2.42}
$$

Then, setting

$$
k_1 = k + \frac{1}{2\epsilon_{z_\theta}} + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}} \tag{2.43}
$$

$$
k_2 = \frac{1}{2}\epsilon_{\Delta_\theta}\delta_{\Delta_\theta} + \frac{1}{2}c\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}, \tag{2.44}
$$

with $k > 0$, gives

$$
\dot{V}_{xz_\theta} \leq -kx^2 - \big(\phi(x)z_\theta\big)^2 \leq 0, \tag{2.45}
$$

which is the same as the dissipation inequality (2.37) in the constant parameter case.

Proposition 2.4. *Consider the closed-loop system described by equations* (2.1), (2.29), *and* (2.30). *Then, all trajectories are bounded,* $\lim_{t\to+\infty} x(t) = 0$, *and* $\phi(x)z_\theta \in \mathcal{L}_2$.

Proof. We break the proof into three parts, for boundedness, convergence, and $\mathcal{L}_2$ property, respectively.

Boundedness. First note that V_{xz_θ} is positive definite in x and z_θ, and its time derivative along the system trajectories satisfies $V_{xz_\theta} \leq 0$. Thus x and z_θ are bounded, and due to the definition of ℓ_θ and $\beta(\cdot)$, $\hat{\theta}$ is also bounded.

Convergence of x. Since both x and $\dot{x}$ are bounded, invoking Lemma A.5 yields $\lim_{t\to+\infty} x(t) = 0$.

$\mathcal{L}_2$ *property.* Note that, since $V_{xz_\theta} \geq 0$ and $\dot{V}_{xz_\theta} \leq 0$, one has

$$0 \leq \int_0^{+\infty} \Big(\phi\big(x(s)\big)z_\theta(s)\Big)^2 \mathrm{d}s \leq \int_0^{+\infty} kx^2(s) + \Big(\phi\big(x(s)\big)z_\theta(s)\Big)^2 \mathrm{d}s$$
$$\leq V_{xz_\theta}(0) - V_{xz_\theta}(+\infty) \leq +\infty, \tag{2.46}$$

which proves that $\phi(x)z_\theta \in \mathcal{L}_2$. □

A Small-Gain-Like Interpretation

As we have briefly discussed, the classical scenario of the I&I scheme, namely, with constant system parameters, guarantees stability properties of the closed-loop system by exploiting the cascaded interconnection of the z_θ-subsystem and the x-subsystem. The same analysis does not hold in the time-varying parameter case as the $\dot{\theta}$ term corrupts the $\mathcal{L}_2$ property of $\phi(x)z_\theta$ (see Fig. 2.2 (a)). By redefining the error z_θ in the spirit of the *congelation of variables* method, we remove the $\dot{\theta}$ term from the analysis, at the cost of interconnecting the x-subsystem and the z_θ-subsystem into a cyclic structure in which the x-subsystem "inputs" $\phi(x)$ to the z_θ-subsystem and the z_θ-subsystem "inputs" $\phi(x)z_\theta$ to the x-subsystem (see Fig. 2.2 (b)). Therefore the introduction of the nonlinear damping term $-k_2\bar{\phi}^2(x)x$ is to guarantee a small-gain-like condition for the interconnection. One can see that if we set $\delta_{\Delta_\theta} = 0$, which means the parameter is constant, the control law (2.29), with the gains defined by (2.43) and (2.44), reduces to the classical control law.

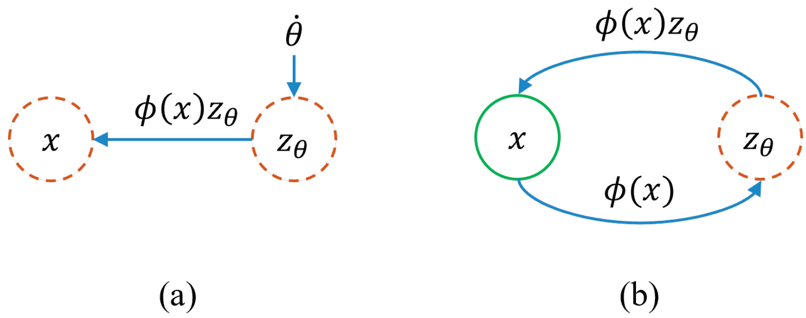

Figure 2.2: A schematic interpretation of the interconnected x and z_θ subsystems: (a) classical adaptive I&I scheme when θ is time-varying and (b) modified interconnection via the congelation of variables method. The colour (and line style) convention is not yet relevant to what we are discussing and will be clarified in Chapter 3 of this thesis.

One can also intuitively interpret the idea of the *congelation of variables* method as removing the effect of $\dot{\theta}$ at the cost of introducing cycles in the closed-loop system, and dominating the cyclic interconnection via a small-gain-like[3] design. This is not seen in the passivity-based scheme because the dominance takes place within the x-subsystem to guarantee passivity of the x-subsystem, instead of guaranteeing a small-gain-like condition between the two subsystems. The fact that the *congelation of variables* method introduces cycles is, however, in general evident, especially when auxiliary filters are used, as we will see in the next section.

2.3 Identification-Based Scheme

The identification-based scheme, as the name suggests, uses a parameter estimator (or equivalently, an identifier) that can identify the values of the system parameters under some conditions, though such a feature is not required in a control problem. To achieve the identification feature, auxiliary filters are used in the design, which increases the dimension of the closed-loop system significantly, even if the plant is a scalar system. Consequently, the application of the *congelation of variables* method also requires more discussions in this context.

To begin with, we revisit the identification-based scheme using the formulation in [75, Chapter 6]. Consider again the scalar nonlinear system (2.1), namely, $\dot{x} = \phi(x)\theta + u$. Since a nonlinear system driven by an exponentially converging parameter estimation error may have finite escape-time (see *e.g.* the example in [75, Section 1.2.1]), we need to design a baseline controller that guarantees boundedness of the solutions of (2.1) before applying standard identification algorithms. To this end, let

$$u = -k(x)x - \phi(x)\hat{\theta} + \mu, \tag{2.47}$$

[3]We call the property that the storage functions and the associated dissipation inequalities of the subsystems, after some scaling, yield an overall dissipation inequality with a negative semi-definite supply rate, a small-gain-like property. This is because in adaptive control problems, the property we need is essentially an algebraic property for dissipation inequalities to allow boundedness and convergence analysis based on Lemma A.5. This is somewhat similar, yet different, from the small-gain properties well-known in the literature, which explicitly exploit stability properties. Therefore we use the term "small-gain-like property" to avoid inaccuracy of expression and confusion.

where $k(x) = k_L + k_\phi \phi^2(x) + \zeta(x)$, with the linear damping gain $k_L > 0$, the nonlinear damping gain $k_\phi > 0$, and the strictly positive term $\zeta(x) > 0$ to be defined. μ is an additional signal added to the baseline control law, and $\hat{\theta}$ is the parameter estimate (the update law of which has to be designed). μ is typically set to 0 as we want to regulate x to 0, yet it can also be used as an excitation signal for identification purposes. Substituting (2.47) into (2.1) yields

$$\dot{x} = -k(x)x + \phi(x)(\theta - \hat{\theta}) + \mu, \tag{2.48}$$

which reveals the following property.

Lemma 2.1. *The system* (2.48) *is ISS with respect to the inputs* $\theta - \hat{\theta}$ *and* μ*.*

Proof. Consider the function $V_x(x) = \frac{1}{2}x^2$. Its time derivative along the trajectories of (2.48) satisfies

$$\begin{aligned} \dot{V}_x = & -k_L x^2 - k_\phi \phi^2 x^2 - \zeta x^2 + x\phi(\theta - \hat{\theta}) + x\mu \\ \leq & -\frac{1}{2}k_L x^2 + \frac{1}{4k_\phi}(\theta - \hat{\theta})^2 + \frac{1}{2k_L}\mu^2. \end{aligned} \tag{2.49}$$

Hence V is an ISS Lyapunov function for the system (2.48), which completes the proof. □

To design a parameter update law for $\hat{\theta}$ based on linear regression we first transform the parametric model (2.48) into a linear regression formulation using the filters

$$\begin{aligned} \dot{\omega}_0 = & -k(x)\omega_0 - \phi(x)\hat{\theta} + \mu, \\ \dot{\omega} = & -k(x)\omega + \phi(x). \end{aligned} \tag{2.50}$$

One could then write

$$x = \omega_0 + \omega\theta + \varepsilon, \tag{2.51}$$

in which the parameterization error ε is governed by the equation

$$\dot{\varepsilon} = \dot{x} - \dot{\omega}_0 - \dot{\omega}\theta - \omega\dot{\theta} = -k(x)\varepsilon - \omega\dot{\theta}. \tag{2.52}$$

Note that $k(x) > k_L > 0$ by definition, thus when θ is constant, that is $\dot{\theta} = 0$, ε converges to 0 exponentially by (2.52). Treating $\hat{x} = \omega_0 + \omega\hat{\theta}$ as a prediction for the state x yields the prediction error

$$\tilde{x} = x - \hat{x} = \omega(\theta - \hat{\theta}) + \varepsilon. \tag{2.53}$$

The parametric model (2.51) or (2.53), compared with the parametric model (2.1) or (2.48), respectively, is described by an algebraic equation instead of a differential equation, which allows using linear regression algorithms, such as the gradient descent method, the least square method, and their variants (see *e.g.* Chapter 4 of [51]), which are widely applied in the area of system identification.

If θ is time-varying, then $\dot{\theta} \neq 0$ and the disturbance term $-\omega\dot{\theta}$ "corrupts" the convergence of ε and the parametric model (2.51) is no longer valid. To circumvent this issue we continue to exploit the *congelation of variables* method by rewriting the parametric models (2.48) and (2.53) as

$$\dot{x} = -k(x)x + \phi(x)(\ell_\theta - \hat{\theta}) + \phi(x)\Delta_\theta + \mu \tag{2.54}$$

and

$$x = \omega_0 + \omega\ell_\theta + \varepsilon, \tag{2.55}$$

respectively. Substituting (2.54) and (2.50) into (2.55) yields

$$\dot{\varepsilon} = -k(x)\varepsilon + \phi(x)\Delta_\theta. \tag{2.56}$$

The resulting prediction error is then described as

$$\tilde{x} = x - \hat{x} = \omega(\ell_\theta - \hat{\theta}) + \varepsilon. \tag{2.57}$$

Note that the use of the new parametric model (2.55) allows replacing the disturbance term $-\omega\dot{\theta}$ in the ε-dynamics with the term $\phi(x)\Delta_\theta$. The latter perturbation term, though preventing concluding convergence of ε directly from (2.55), could be dominated via a joint small-gain-like design without imposing restrictions on $\dot{\theta}$.

To see this, we need to compute the dissipation inequalities for the $\hat{\theta}$, ε, and $\hat{x}$-subsystems. First consider the function $V_{\hat{\theta}}(\hat{\theta}) = \frac{1}{2\gamma}(\ell_\theta - \hat{\theta})^2$. One possible selection for the parameter update law is the classical gradient descent law

$$\dot{\hat{\theta}} = -\gamma \frac{\partial}{\partial \hat{\theta}} \left(\frac{\tilde{x}^2}{2} \right) = \gamma \omega \tilde{x}, \tag{2.58}$$

where $\gamma > 0$ is the adaptation gain. Recalling the parametric model (2.55), the time derivative of $V_{\hat{\theta}}$ along the trajectories of (2.58) is

$$\dot{V}_{\hat{\theta}} = -\omega(\ell_\theta - \hat{\theta})\tilde{x} = -\tilde{x}^2 + \varepsilon\tilde{x} \leq -\frac{1}{2}\tilde{x}^2 + \frac{1}{2}\varepsilon^2. \tag{2.59}$$

Consider now the function $V_\varepsilon(\varepsilon) = \frac{1}{2}\varepsilon^2$ and its time derivative along the trajectories of (2.56), which yields

$$\begin{aligned} \dot{V}_\varepsilon =& -k(x)\varepsilon^2 + \varepsilon\phi(x)\Delta_\theta \\ \leq& -k(x)\varepsilon^2 + \frac{\epsilon_{\Delta\theta}\delta_{\Delta\theta}}{2}\bar{\phi}^2(x)\varepsilon^2 + \frac{\delta_{\Delta\theta}}{2\epsilon_{\Delta\theta}}x^2 \\ \leq& -\left(k(x) - \frac{\epsilon_{\Delta\theta}\delta_{\Delta\theta}}{2}\bar{\phi}^2(x)\right)\varepsilon^2 + \frac{\delta_{\Delta\theta}}{\epsilon_{\Delta\theta}}(\hat{x}^2 + \tilde{x}^2), \end{aligned} \tag{2.60}$$

where $\epsilon_{\Delta\theta} > 0$ is a balancing coefficient that can be selected arbitrarily. Note that the dynamics of the prediction error is governed by the equation

$$\dot{\hat{x}} = \dot{\omega}_0 + \dot{\omega}\hat{\theta} + \omega\dot{\hat{\theta}} = -k(x)\hat{x} + \gamma\omega^2\tilde{x} + \mu. \tag{2.61}$$

Consider finally the function $V_{\hat{x}}(\hat{x}) = \frac{1}{2}\hat{x}^2$, the time derivative of which along the trajectories of (2.61) is

$$\begin{aligned}\dot{V}_{\hat{x}} =& -k(x)\hat{x}^2 + \gamma\omega^2\hat{x}\tilde{x} + \hat{x}\mu \\ \leq& -k(x)\hat{x}^2 + \frac{1}{2}k_L\hat{x}^2 + \frac{\gamma^2\omega^4}{k_L}\tilde{x}^2 + \frac{1}{k_L}\mu^2 \\ \leq& -\left(k(x) - \frac{1}{2}k_L\right)\hat{x}^2 + \frac{\gamma^2\omega^4}{k_L}\tilde{x}^2 + \frac{1}{k_L}\mu^2. \end{aligned} \tag{2.62}$$

One notes that the ω-term in the dissipation inequality (2.62) depends on the state of the dynamical system (2.50), and its boundedness has not been established yet, which causes difficulties in the small-gain analysis. To solve this issue, consider the function $V_\omega(\omega) = \frac{1}{2}\omega^2$ and its time derivative along the trajectories of (2.50), that is, $\dot{V}_\omega = -k(x)\omega^2 + \phi(x)\omega \leq -k_L\omega^2 + \frac{1}{4k_\phi} = -2k_LV_\omega + \frac{1}{4k_\phi}$. This means that $|\omega(t)| \leq \max\{\omega(0), \frac{1}{2\sqrt{k_Lk_\phi}}\} \triangleq \delta_\omega > 0$, for all $t \geq 0$. Note that δ_ω can be computed since $\omega(0)$, k_L, and k_ϕ are known. This allows rewriting (2.62) as

$$\dot{V}_{\hat{x}} \leq -\left(k(x) - \frac{1}{2}k_L\right)\hat{x}^2 + \frac{\gamma^2\delta_\omega^4}{k_L}\tilde{x}^2 + \frac{1}{k_L}\mu^2. \tag{2.63}$$

Select now the additional nonlinear damping term $\zeta(x)$ in the control law (2.47) as

$$\zeta(x) = \zeta_L + \frac{\epsilon_{\Delta\theta}\delta_{\Delta\theta}}{2}\bar{\phi}^2(x), \tag{2.64}$$

with ζ_L a linear damping gain that guarantees the small-gain property of the interconnected $\hat{\theta}$, $\hat{x}$, and ε-subsystems. The existence of such a gain ζ_L is guaranteed by the following lemma.

Lemma 2.2. *Consider the dissipation inequalities* (2.59), (2.60) *and* (2.63). *Let* ζ *be defined as in* (2.64). *There exist constant scaling coefficients* $c_{\hat{x}}$, $c_{\hat{\theta}}$, c_ε, *and a linear damping gain* $\zeta_L > 0$, *such that for the function* $V(\hat{x}, \hat{\theta}, \varepsilon) = c_{\hat{x}}V_{\hat{x}}(\hat{x}) + c_{\hat{\theta}}V_{\hat{\theta}}(\hat{\theta}) + c_\varepsilon V_\varepsilon(\varepsilon)$, *along the trajectories of the closed-loop system described by* (2.61), (2.58), *and* (2.52), *the dissipation inequality*

$$\dot{V} \leq -\hat{x}^2 - \tilde{x}^2 - \varepsilon^2 + b_\mu\mu^2 \tag{2.65}$$

holds, for some $b_\mu > 0$.

Proof. First note that with the additional nonlinear damping term defined in (2.64) the dissipation inequalities (2.59), (2.60) and (2.63) can be re-written as

$$\begin{bmatrix} \dot V_{\hat x} \\ \dot V_{\hat\theta} \\ \dot V_{\varepsilon} \end{bmatrix} \leq - \begin{bmatrix} \frac{1}{2}k_L & -\frac{\gamma^2\delta_\omega^4}{k_L} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ -\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} & -\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} & \zeta_L \end{bmatrix} \begin{bmatrix} \hat x^2 \\ \tilde x^2 \\ \varepsilon^2 \end{bmatrix} + \begin{bmatrix} \frac{1}{2k_L}\mu^2 \\ 0 \\ 0 \end{bmatrix}, \tag{2.66}$$

where the "$\leq$" sign is defined element-wise.

We proceed to prove the claim by construction. Consider the selection for ζ_L:

$$\zeta_L = \frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}\left(\frac{4\gamma^2\delta_\omega^4}{k_L^2} + 3\right), \tag{2.67}$$

then one can verify by some straightforward computations that, with the scaling constants

$$c_{\hat x} = \frac{4}{k_L}, \quad c_{\hat\theta} = 2\left(\frac{4\gamma^2\delta_\omega^4}{k_L^2} + 2\right), \quad c_\varepsilon = \frac{\epsilon_{\Delta_\theta}}{\delta_{\Delta_\theta}}, \tag{2.68}$$

the dissipation inequality (2.65) holds with $b_\mu = \frac{2}{k_L^2}$. □

With the small-gain-like property revealed by Lemma 2.2, one could conclude boundedness and convergence of the closed-loop signals as follows.

Proposition 2.5. *Consider the closed-loop system consisting of the plant* (2.1) *and the adaptive controller* (2.47)*,* (2.50)*,* (2.58)*, with the damping term* (2.64) *and* $\mu = 0$*. Then all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} x(t) = 0$.

Proof. We break the proof into two parts.

Boundedness. Recall that boundedness of ω has been established. With $\mu = 0$, it can be concluded from the dissipation inequality (2.65) in Lemma 2.2 that $\hat\theta$, $\hat x$ and ε are bounded. By (2.57), boundedness of ω, $\hat\theta$, and ε implies that $\tilde x$ is also bounded. By Lemma 2.1, Assumption 1.1 and the fact that $\hat\theta$ is bounded, we can conclude that x is bounded. Boundedness of x, $\hat\theta$ and smoothness of ϕ together with (2.50) imply that ω_0 is bounded.

Convergence of x. Note that $\dot{\hat{x}}$ is bounded due to boundedness of x, $\hat{x}$, $\tilde{x}$, and ω, and that $\dot{\tilde{x}} = -\big(k(x) + \gamma\omega^2\big)\tilde{x} + \phi(x)(\theta - \hat{\theta})$ is bounded due to boundedness of x, ω, $\hat{\theta}$ and Assumption 1.1. Due to boundedness of $\hat{x}$, $\dot{\hat{x}}$, $\tilde{x}$, and $\dot{\tilde{x}}$, invoking Lemma A.5, we can conclude that $\lim_{t\to+\infty} \hat{x}(t) = 0$ and $\lim_{t\to+\infty} \tilde{x}(t) = 0$. Finally recall that $x = \hat{x} + \tilde{x}$, and therefore $\lim_{t\to+\infty} x(t) = 0$, which completes the proof. □

Remark 2.5. *Observing the two parametric model* (2.54) *and* (2.57) *one could see that the parameter ℓ_θ and the parameter estimate $\hat{\theta}$ coupled with ϕ in* (2.54) *are "swapped" from the input side of the filter* (2.50) *to the output side (regarding the filter state ω as an output) in* (2.57)*, coupled with the filtered regressor ω. The swapping lemma [90], as well as its equivalents and variants (see e.g. Appendix A of [51] and Appendix F of [75]), are widely used in the adaptive control literature to justify the use of the filters* (2.50)*. The classical use of the swapping lemma in the stability proof is to establish boundedness and square-integrability for the cascaded subsystems depicted in Fig. 2.3(a), which is not directly applicable* [4] *to Fig. 2.3(b). To this end, we use Lemma 2.2 instead of the swapping lemma to establish boundedness and convergence. Though stated differently, the idea behind the analysis based on Lemma 2.2 and the classical analysis based on the swapping lemma are similar, and the resulting identifier (filters and the parameter update law) design are the same except for the additional nonlinear damping term ζ.*

Remark 2.6. *In the passivity-based congelation of variables scheme discussed in Section 2.1, the meaning of congealed parameter ℓ_θ is not intuitive since it is treated as a non-implementable unknown parameter to be used only in the analysis for establishing boundedness of $\hat{\theta}$, whereas there is no guarantee that $\hat{\theta}$ converges to ℓ_θ. In the identification-based scheme there is an interpretation of ℓ_θ from an identification perspective. Note that $\dot{V}_\varepsilon \leq -(k_L - k_\phi\phi^2)\varepsilon^2 + \varepsilon\phi^\top\Delta_\theta \leq -k_L\varepsilon^2 + \frac{1}{4k_\phi}\Delta_\theta^2$ hence if θ varies in such a way that $\Delta_\theta \in \mathcal{L}_2$ (e.g. θ converges to ℓ_θ exponentially), then $\varepsilon \in \mathcal{L}_2$. It is well known that if ω is persistently exciting and $\varepsilon \in \mathcal{L}_2$, then $\hat{\theta}$ converges to ℓ_θ. This means that the congelation of variables method preserves the identification capability of the adaptive controller under proper excitation conditions (enforced by μ). This feature once again highlights that by us-*

[4] It should be noted that as a generic mathematical result on dynamical operators, the *swapping lemma* does not impose any restrictions on the system parameters. The design based on this lemma, however, requires that the signal $\frac{\mathrm{d}}{\mathrm{d}t}(\theta - \hat{\theta})$ be implementable in the filters. In the classical time-invariant case $\frac{\mathrm{d}}{\mathrm{d}t}(\theta - \hat{\theta})$ and $-\dot{\hat{\theta}}$ can be used interchangeably, and the latter can be directly obtained from the parameter update law, whereas when θ is time-varying, $\frac{\mathrm{d}}{\mathrm{d}t}(\theta - \hat{\theta})$ is no longer implementable in the filters.

ing the congealed parameter ℓ_θ, we do not lose the "adaptive" nature, which distinguishes the proposed scheme from other methods that dominate the system without making the system literally "adaptive".

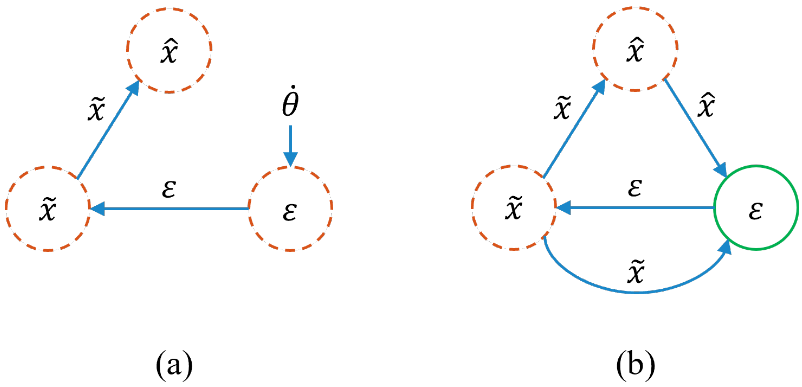

Figure 2.3: A schematic interpretation of the interconnected $\hat{x}$, $\tilde{x}$, and ε subsystems. (a) shows the interconnection of the classical identification-based scheme when θ is time-varying and (b) shows the interconnection after modified by the congelation of variables method.

A Small-Gain-Like Interpretation

One can see from Fig. 2.3 that similarly to the case of the I&I scheme, the congelation of θ removes the undesired unknown $\dot{\theta}$ from the scope of the analysis at the cost of interconnecting the three subsystems into a cyclic structure. The dominance design using the nonlinear damping term $\zeta(x)$ enforces a small-gain-like condition over the interconnected subsystems, which allows establishing boundedness and convergence of the closed-loop signals.

However, as there are more than one intersecting cycles in the interconnection, the small-gain-like analysis in general requires constructive proofs, like the one we did for Lemma 2.2, if we only consider the standard analysis tools. Such a constructive method does work for establishing the boundedness and convergence results that we need, but is not going to be pursued in this thesis, mainly for two reasons. First, the complexity of such constructions grows together with the dimension of the closed-loop system: the more complex auxiliary filters we use, the more complicated the analysis is. Sticking to the constructive method prevents the further development of the proposed schemes once we move forward to more complex systems. Second, the constructive method does not reveal the dependency among the design parameters and in general one has to select the design in a certain form to make sure that the construction works. Third, and possibly most importantly, the constructive method creates an illusion that the boundedness and convergence analysis can only be established based on "smart" constructions of Lyapunov functions, which may distract the reader (as well as the author) away from the "big picture" of the proposed schemes. In the light of this, we should develop a suitable analysis and synthesis tool to reveal the nature of the proposed schemes in an abstract and intuitive manner.

Chapter 3

Dominance Design over Interconnections: Active Nodes

From the aforementioned discussion, one can identify the two ingredients of the *congelation of variables* method.

- **The parameter congelation** consists in substituting ℓ_θ for θ to remove the undesired $\dot{\theta}$ term from the analysis and design, at the cost of creating additional interconnections.

- **The dominance design over the interconnected systems** consists in enforcing a small-gain-like condition over the interconnected systems via feedback control design.

The methodology associated to the first step has been introduced with simple examples in Chapter 2. In this chapter, we show how to accomplish the second step, especially for interconnected systems with underlying directed graph possessing multiple cycles.

It is worth mentioning that the results to be discussed in this chapter, chronologically speaking, were not only developed after but also inspired by the development of the *congelation of variables* method for the solution of adaptive output-feedback problems in the presence of time-varying parameters, in which the interconnection of the subsystems in the closed-loop system is more complex than that arising in the classical constant parameter case, and for which constructive analysis/synthesis methods do not work well. These results are presented before the adaptive control results that originally motivated them, because the author believes that understanding these results can significantly facilitate the understanding of the application of the *congelation of variables* method to complex scenarios and it is logically natural.

To start with, some definitions and theorem are recalled since they are useful in the remainder of the chapter.

Definition 3.1 (*Z-matrix*)**.** *A matrix A is called a Z-matrix (or negated Metzler matrix) if all its off-diagonal elements are non-positive, that is, $(A)_{ij} \leq 0$, $i \neq j$.*

Definition 3.2 (*M-matrix*)**.** *A matrix $A = B + sI$, where B is a square Z-matrix and s is a real number not smaller than the spectral radius of B, is called an M-matrix.*

Theorem 3.1. *Let A be a Z-matrix. Then the following conditions[1] are equivalent.*

1. *A is a non-singular M-matrix.*
2. *$A^\top$ is a non-singular M-matrix.*
3. *All principal minors of A are positive.*
4. *All leading principal minors of A are positive.*
5. *A^{-1} exists and $A^{-1} \geq 0$.*
6. *There exists a vector $v > 0$ such that $Av > 0$.*

3.1 Systems with Quadratic Supply Rates

Consider a network of n interconnected dynamical systems in which each node, a dynamical system denoted as Σ_i, $i = 1, \ldots, n$, has $n-1$ inputs $u_{ij}(t) \in \mathbb{R}^{n^u_{ij}}$, $j = 1, \ldots, i-1, i+1, \ldots, n$, one output $y_i(t) \in \mathbb{R}^{n^y_i}$, and satisfies the dissipation inequality

$$\dot{V}_i \leq -a_i|y_i|^2 + \sum_{j=1, j\neq i}^{i-1} b_{ij}|u_{ij}|^2, \tag{3.1}$$

[1]These are extracted from 40 equivalent conditions listed in [101], in which the proof of this result is available.

with respect to a storage function $V_i : \mathbb{R}^{n_i} \to \mathbb{R}_+$ of class $\mathcal{C}^1$, where n_i is the dimension of the state vector of Σ_i, $a_i > 0$, and $b_{ij} \geq 0$. The nodes are interconnected via the equations $u_{ij} = y_j$, for all $b_{ij} \neq 0$, whereas $b_{ij} = 0$ means that Σ_j is not connected to Σ_i, that is $u_{ij} = 0$. Then the node dissipation inequalities can be written into the compact form

$$\dot{\bar{V}} \leq -E\phi(y), \tag{3.2}$$

where $\bar{V} \triangleq [V_1, \ldots, V_n]^\top$, $\phi(y) \triangleq \left[|y_1|^2, \ldots, |y_n|^2\right]^\top$, and[2]

$$E \triangleq \begin{bmatrix} a_1 & -b_{12} & \cdots & -b_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ -b_{n1} & \cdots & -b_{n(n-1)} & a_n \end{bmatrix}. \tag{3.3}$$

The structure of the network can be described by a directed graph, the underlying weighted adjacency matrix B of which is defined by

$$(B)_{ij} \triangleq \begin{cases} 0, & \text{if } i = j, \\ b_{ij}, & \text{otherwise,} \end{cases} \tag{3.4}$$

that is, the negated off-diagonal part of E. For example, consider a network system with the node dissipation inequalities described by the matrix

$$E = \begin{bmatrix} a_1 & 0 & -b_{13} & -b_{14} & -b_{15} & -b_{16} \\ -b_{21} & a_2 & 0 & 0 & 0 & 0 \\ -b_{31} & -b_{32} & a_3 & 0 & 0 & 0 \\ -b_{41} & 0 & 0 & a_4 & 0 & 0 \\ -b_{51} & 0 & 0 & -b_{54} & a_5 & -b_{56} \\ -b_{61} & 0 & 0 & 0 & 0 & a_6 \end{bmatrix}. \tag{3.5}$$

The underlying directed graph is the one shown in Fig. 3.1.

In practice one is interested in the dissipation inequality for the overall network system. More specifically, one may wonder whether there exist positive scaling coefficients

[2]The definition of E is different from the definition of its counterpart in [21], as the definition used here can be related to an adjacency matrix convention commonly used in the control community.

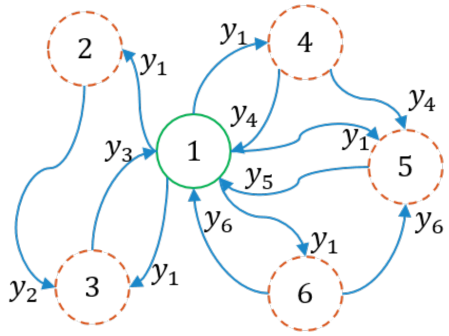

Figure 3.1: The underlying directed graph of a network with node dissipation inequalities specified by (3.2) **and** (3.5)**.**

$c_1, \ldots, c_n$ such that the storage function of the overall network system, defined by

$$V = \sum_{i=1}^{n} c_i V_i = c^\top \bar{V}, \tag{3.6}$$

where $c \triangleq [c_1, \ldots, c_n]^\top$ satisfies the dissipation inequality $\dot{V} \leq 0$. In addition, in the case in which the V_i's are positive definite and radially unbounded one may require that

$$\dot{V} \leq -W(y) \leq 0, \tag{3.7}$$

where $W(\cdot)$ is a positive definite function of $y \triangleq [y_1, \ldots, y_n]^\top$. From the dissipation inequality (3.7) one can conclude Lyapunov stability of the equilibrium of the overall system and convergence of all y_i's to 0 by Lemma A.5, or similar invariance-like analysis. Such a property is sufficient for the adaptive control results in the thesis, which focus on establishing boundedness of closed-loop signals and convergence of system state or output. A specific stability property, however, is not the main concern of this chapter, although it reveals a possible area of applicability of the forthcoming results.

The aim of the chapter is to study how to use the structure of the network to make the dissipation inequality (3.7) hold, in other words, we do not impose any condition on V and simply focus on $\dot{V}$. To this end, the next result provides a condition for the existence of the scaling coefficients c_i, $i = 1, \ldots, n$, mentioned above.

Theorem 3.2. *Consider the function V defined by* (3.6)*. If the matrix E in* (3.3) *is a non-singular M-matrix then, for all $\sigma \triangleq [\sigma_1, \ldots, \sigma_n]^\top > 0$, there exists a vector of scaling* *coefficients* $c > 0$ *depending on* σ *such that the dissipation inequality*

$$\dot{V} \leq -\sigma^\top \phi(y) = -\sum_{i=1}^{n} \sigma_i |y_i|^2 \leq 0 \tag{3.8}$$

holds.

Proof. To begin with, note that

$$\dot{V} = \sum_{i=1}^{n} c_i \dot{V}_i \leq \; -c^\top E \phi(y) \leq 0, \tag{3.9}$$

where $\phi(y) = \left[|y_1|^2, \ldots, |y_n|^2\right]^\top$. Invoking condition 5) of Theorem 3.1 we can construct the scaling vector as $c = (E^\top)^{-1}\sigma$, for all $\sigma > 0$ and $c > 0$ due to the fact that $(E^\top)^{-1}$ is non-singular and $(E^\top)^{-1} \geq 0$. The conclusion is therefore obtained invoking condition 6) of Theorem 3.1. □

Remark 3.1. *There are several variants of Theorem 3.2 available in the literature, for example, the criteria for $\mathcal{L}_p$-stability based on the so-called test matrix, that is essentially the matrix* (3.3) *written in terms of $\mathcal{L}_p$ gains, see [130, Section 6.2]. Theorem 3.2, however, provides purely an algebraic result which does not require any assumption on the stability properties of each node.*

It is worth noting that the condition expressed by Theorem 3.2 is not generic: it is straightforward to build networks for which it is not satisfied. One of such networks is a single-cycle interconnection containing two scalar nodes that violates the small-gain condition, that is either the condition $\frac{b_{21}}{a_1}\frac{b_{12}}{a_2} < 1$ or, equivalently, the condition $a_1 a_2 - b_{12} b_{21} > 0$. (The counterpart of this small-gain condition for more complex networks is precisely condition 4) of Theorem 3.1.) The small-gain analysis for the considered example reveals the fact that if one is allowed to adjust the coefficients a_i's arbitrarily one can always enforce the dissipation inequality (3.7), provided there is a distributed controller on each of the nodes of the network to make the a_i's tunable design parameters. In practice, however, this is not always feasible, for example, because of dynamics that cannot be controlled, or economical concerns do not allow using as many distributed controllers as nodes. This highlights the fact that, even if Theorem 3.2 provides a tool for network analysis, we need to answer a question from a design perspective: how many controllers are needed to enforce the dissipation inequality (3.7) and where these should be placed considering the structure of the network?

To answer this question we define a special class of nodes.

Definition 3.3. *A node Σ_i is called an active node if it satisfies the dissipation inequality* (3.1) *with an adjustable $a_i \in [\underline{a}_i, +\infty)$, with $\underline{a}_i > 0$. The property "Σ_i is active" is denoted by $i \in \mathcal{I}_A$, where $\mathcal{I}_A$ is the index set of the active nodes.*

We now make a convention for graphic representation: an *active node* is represented by a solid green circle (*e.g.* node Σ_1 in Fig. 3.1) and a non-active node is represented by a red dashed circle. Exploiting the concept of *active nodes* we now present a feasibility condition for the considered design problem.

Theorem 3.3. *The matrix* (3.3) *can be made a non-singular M-matrix by adjusting the parameters a_{i_A}, $i_A \in \mathcal{I}_A$, if every directed cycle of the underlying directed graph describing the network contains at least one vertex associated with an active node.*

To prove Theorem 3.3 we need to first identify a directed cycle in the graph. It turns out that it is much easier to identify a directed graph without any directed cycle, that is a directed acyclic graph (DAG). Although most of the criteria for determining DAGs are algorithm-based, we give an algebraic condition on the relation between a DAG and the matrix (3.3).

Lemma 3.1. *The matrix* (3.3) *of a network which is described by a DAG satisfies the condition*

$$\det(E) = \prod_{i=1}^{n} a_i. \tag{3.10}$$

Proof. We use Laplace expansion (also called the cofactor expansion) to compute the determinant. Note that one of the properties of a DAG is that there is at least one vertex that has no incoming edges and at least one vertex that has no outgoing edge. Without loss of generality assume that vertex 1 (related to node Σ_1) is the vertex with no incoming edge. This means that $(E)_{11} = a_1$ is the only non-zero element in the first row of E, namely $(E^\top)_1$, which yields

$$\det(E) = a_1(-1)^{(1+1)}\det(E_1) = a_1\det(E_1), \tag{3.11}$$

where E_1 is the matrix obtained from E after deleting the row and column containing $(E)_{11}$. Note that this is equivalent to deleting vertex 1 and the edges attached to it from the graph. Recall now that a DAG is such that the removal of any of its vertices yields a new (reduced) DAG. From the graph thus obtained we can select a vertex without incoming edges, say, vertex 2 without loss of generality. Noting that $(E)_{22} = (E_1)_{11} = a_2$ yields

$$\begin{aligned}\det(E) &= a_1\det(E_1) = a_1a_2(-1)^{(1+1)}\det(E_2)\\ &= a_1a_2\det(E_2),\end{aligned} \tag{3.12}$$

where E_2 is the matrix obtained from E_1 after deleting the row and column containing $(E_1)_{11}$. Repeating this reduction operation until all the vertices of the original graph have been deleted and noting that every vertex that has been selected for the reduction process has no incoming edge, yields a Laplace expansion in which only one factor a_i is present at each step. As a result, condition (3.10) holds and the proof is complete. □

Proof of Theorem 3.3: We use condition 3) of Theorem 3.1 to prove the claim. Lemma 3.1 implies that if any subgraph is a DAG the corresponding principal minor cannot contain b_{ij} terms. Thus the b_{ij} terms in the principal minors are all contributed by the directed cycles. More specifically, for directed graphs with more than four vertices, the b_{ij} terms can also be associated with the union of disjoint directed cycles [77, Property 1]. In the light of this fact and for convenience of the discussion, one can assign an index $l = 1, \ldots, l_{\max}$ to each directed cycle and union of disjoint directed cycles, and let $\mathcal{C}_l$ be the vertex index set of the directed cycle (*e.g.*, the vertex index set $\{1, 2\}$ for the directed cycle "$1 \to 2 \to 1$") or the union of disjoint directed cycle (*e.g.*, the vertex index set $\{1, 2, 3, 4\}$ for the union of the directed cycles "$1 \to 2 \to 1$" and "$3 \to 4 \to 3$"). Consider now the largest principal minor $\det(E)$ and note that

$$\det(E) = \prod_{i \in \mathcal{I}} a_i + \sum_{l=1}^{l_{\max}} \Big(\prod_{i \in \mathcal{I} \backslash \mathcal{C}_l} a_i \Pi_b \Big), \tag{3.13}$$

where $\mathcal{I}$ is the set of all node/vertex indices of the graph, the structure of which is specified by E, $\Pi_b \triangleq s_l \prod_{j,k \in \mathcal{C}_l, k \in \mathcal{S}_j} b_{kj}$, and s_l is the sign of the cofactor related to each $\mathcal{C}_l$. Therefore $\det(E)$ can be re-written as

$$\det(E) = \sum_{l=1}^{l_{\max}} \Big(\prod_{i \in \mathcal{I} \backslash \mathcal{C}_l} a_i \Big(\frac{1}{l_{\max}} \prod_{j \in \mathcal{C}_l} a_j + \Pi_b \Big) \Big). \tag{3.14}$$

It is easy to see that since there is at least one *active node* in each directed cycle or, equivalently, there exists $i_A \in \mathcal{C}_l$ such that we can always satisfy the condition

$$\frac{1}{l_{\max}} \prod_{j \in \mathcal{C}_l} a_j + \Pi_b > 0 \tag{3.15}$$

by selecting one of the a_{i_A}'s sufficiently large, one can guarantee the condition $\det(E) > 0$. Since this analysis can also be applied to the submatrix related to each of the principal minors of E, there has to be a selection of a_{i_A} for each *active node* such that all principal minors of E are positive. This proves that E is a non-singular *M-matrix* and completes the proof. □

Remark 3.2. *Theorem 3.3 provides a sufficient condition of the existence of α_{i_A}, $i_A \in \mathcal{I}_A$, which makes E a non-singular M-matrix. This condition is in general not necessary, since if E is already a non-singular M-matrix, then there is no need to use active nodes to enforce such a property. However, a robust version of Theorem 3.3, which holds for a family of E's (instead of a certain E), may make the condition in Theorem 3.3 also necessary. The robust counterpart of Theorem 3.3 is to be studied in future work.*

To illustrate the result expressed by Theorem 3.3 consider the example illustrated in Fig. 3.1. The matrix E related to this graph is a non-singular *M-matrix* with a_1 sufficiently large, as node Σ_1 is an *active node*, and its associated vertex is contained in every directed cycle of the graph. Theorem 3.2 and Theorem 3.3, therefore, reveal that as long as there is at least one *active node* in every directed cycle, there exist positive scaling coefficients

$c_1, \ldots, c_n$ and design parameters a_{i_A}, $i_A \in \mathcal{I}_A$ such that the dissipation inequality (3.7) holds. Note that this condition simply guarantees the feasibility of the underlying design problem, yet does not provide an approach to carry out the design computationally.

3.2 Systems with Nonlinear Supply Rates

The systems discussed in Section 3.1 have dissipation inequalities with quadratic supply rates. The advantage of this formulation is that each term in the dissipation inequality is one-to-one related to a node subsystem, which allows carrying out analysis and design via the underlying directed graph. This, however, puts restrictions that do not allow the use of many common nonlinear control design techniques, *e.g.* the use of nonlinear damping. Although the quadratic supply rate case does take nonlinear dynamic systems into account, it is natural to ask whether the results developed in Section 3.1 can be applied to generic nonlinear supply rates. In the light of this, the rest of this subsection generalizes the matrix-based conditions discussed in Section 3.1 for network systems with *sum-type* nonlinear supply rates[3]. We consider an n-node network, with node dissipation inequalities given by

$$\dot{V}_i = -\alpha_i(V_i) + \sum_{j=1, j\neq i}^{n} \beta_{ij}(V_j), \tag{3.16}$$

where $\alpha_i \in \mathcal{K}_\infty$ and $\beta_{ij} \in \mathcal{K}_\infty \cup \{0\}$. This can be understood as a more general version of (3.9) and the gain functions β_{ij}'s provides a more general form of the b_{ij} in (3.4) which also describe an underlying weighted directed graph. This allows discussing the dissipativity properties of the network systems by discussing the weight properties of the underlying directed graph. Unlike the case of quadratic supply rates, in which finding a vector of constant scaling coefficients is sufficient to render the desired property of the overall dissipation inequality, in the case of nonlinear supply rates the constant scaling technique is not applicable since one cannot separate β_{ij} and V_j from the $\beta_{ij} \circ V_j$ term which is state-dependent. In this case, the *state-dependent scaling* method [52] is applicable. More

[3]In this section we express the supply rates in terms of the node storage functions V_i's instead of the output y_i's, for convenience. These can be equivalently expressed using y_i as explained in [53].

specifically, the aim of this subsection is to find the condition on α_i and β_{ij} that guarantees the existence of continuous functions $\lambda_i : \mathbb{R}_+ \to \mathbb{R}_+$, $i = 1, \ldots, n$ satisfying

$$\lambda_i(s) > 0, \ \forall s > 0, \tag{3.17}$$

$$\int_0^{+\infty} \lambda_i(s) \mathrm{d}s = +\infty, \tag{3.18}$$

such that the time derivative of

$$V = \sum_i^n \int_0^{V_i} \lambda_i(s) \mathrm{d}s \tag{3.19}$$

is negative definite in $V_1, \ldots, V_n$, namely

$$\dot{V} = -\sum_{i=1}^n \left(\lambda_i(V_i)\alpha_i(V_i) - \sum_{j \in \mathcal{P}_i} \lambda_i(V_i)\beta_{ij}(V_j) \right) < 0, \tag{3.20}$$

where $\mathcal{P}_i$ is the set comprised of the indices of the direct predecessors of the vertex i. One can see that this method is called *state-dependent scaling* because the construction of the overall dissipation inequality (3.19) allows $\lambda_i(V_i)$ to scale the supply rate to the node Σ_i in a similar way as c_i works in the quadratic case.

We proceed by removing the coupling between $\lambda_i(V_i)$ and $\beta_{ij}(V_j)$ since they are depending on different variables, namely V_i and V_j. By Lemma A.2, one can write

$$-\lambda_i(V_i)\beta_{ij}(V_j) \geq -\int_0^{\lambda_i(V_i)} f_i(s)\mathrm{d}s - \int_0^{\beta_{ij}(V_j)} f_i^{-1}(s)\mathrm{d}s, \tag{3.21}$$

which removes the coupling between the V_i-dependent λ_i and the V_j dependent β_{ij}. This allows deriving a sufficient condition for (3.20) to hold, namely

$$\tilde{r}_i(V_i) \triangleq \lambda_i(V_i)\alpha_i(V_i) - p_i \int_0^{\lambda_i(V_i)} f_i(s)\mathrm{d}s - \sum_{j \in \mathcal{S}_i} \int_0^{\beta_{ji}(V_i)} f_j^{-1}(s)\mathrm{d}s > 0, \tag{3.22}$$

for $V_i > 0$, $i = 1, \ldots, n$, where $p_i \triangleq |\mathcal{P}_i|$ and $\mathcal{S}_i$ is the set comprised of the indices of the direct successors of the vertex i. $\tilde{r}_i(V_i)$ can be understood as the redundant damping in

V_i. In fact, if

$$r_i\big(\lambda_i(V_i), \alpha_i(V_i)\big) \triangleq \lambda_i(V_i)\alpha_i(V_i) - p_i \int_0^{\lambda_i(V_i)} f_i(s)\mathrm{d}s \tag{3.23}$$

is treated as the provided damping then

$$\hat{r}_i(V_i) \triangleq \sum_{j \in \mathcal{S}_i} \int_0^{\beta_{ji}(V_i)} f_j^{-1}(s)\mathrm{d}s \tag{3.24}$$

can be treated as the demanded damping. Therefore the implication from (3.22) to (3.20) can be interpreted as follows: if each node system dissipates the energy that it injects to its successor node systems, then the overall network system is dissipative. Consider now two types of vertices (using the index i for illustration).

1. $|\mathcal{P}_i| = 0$, that is, the vertex i has no incoming edges. This means λ_i can always be selected as a sufficiently "large" function so that (3.22) is satisfied, since $r_i(\lambda_i, \alpha_i)$ is not upperbounded in λ_i if $p_i = 0$. Then the vertex i can be deleted from the graph as the information of the vertex i is not needed in the analysis of the rest of the graph.

2. $|\mathcal{P}_i| \geq 1$, that is, the vertex i has at least one incoming edge. In this case $r_i(\lambda_i, \alpha_i)$ is upperbounded in λ_i due to the negative term $-p_i \int_0^{\lambda_i(V_i)} f_i(s)\mathrm{d}s$ and therefore (3.22) can only hold under a small-gain condition on α_i and β_{ij}.

In the light of these observations, we can recursively delete the vertices with $p_i = 0$ and the outgoing edges attached to these vertices until obtaining a graph in which each vertex has $p_i \geq 1$, and such deletions do not alter the small-gain condition we intend to derive. Thus, without loss of generality, we can consider a graph with $p_i \geq 1$, for $i = 1, \dots, n$, from the beginning.

To proceed, multiplying $\frac{1}{p_i}$ on both sides of (3.23) and invoking Lemma A.2 yields

$$\frac{1}{p_i} r_i = \lambda_i \frac{1}{p_i}\alpha_i - \int_0^{\lambda_i} f_i(s)\mathrm{d}s \leq \int_0^{\frac{1}{p_i}\alpha_i} f_i^{-1}(s)\mathrm{d}s \tag{3.25}$$

and the equality holds if and only if $f(\lambda_i) = \frac{1}{p_i}\alpha_i$, which indicates that

$$\max_{\lambda_i} \ r_i(\lambda_i, \alpha_i) = p_i \int_0^{\frac{1}{p_i}\alpha_i} f_i^{-1}(s)\mathrm{d}s, \tag{3.26}$$

with the maximizer

$$\lambda_i(V_i) = f_i^{-1} \circ \frac{1}{p_i}\alpha_i \circ V_i. \tag{3.27}$$

Hence, a sufficient condition for (3.22) to hold is

$$p_i T_i \circ \frac{1}{p_i}\alpha_i - \sum_{j\in\mathcal{S}_i} T_j \circ \beta_{ji} \geq \delta_i, \tag{3.28}$$

for $i = 1, \ldots, n$, where $\delta_i \in \mathcal{K}_\infty$, and the transform[4] T_i is defined by

$$T_i(s) \triangleq \int_0^s f_i^{-1}(\sigma)\mathrm{d}\sigma. \tag{3.29}$$

We now provide two lemmas which are useful in the subsequent analysis.

Lemma 3.2. *Consider a convex function $g : \mathbb{R} \to \mathbb{R}$ such that $g(0) \leq 0$ and two positive real numbers $x_2 > x_1 > 0$. The following inequality holds*

$$g(x_2) - g(x_1) \geq g(x_2 - x_1). \tag{3.30}$$

Proof. By the super-additivity of the convex function g for any positive real numbers $\bar{x}_1$ and $\bar{x}_2$, one has $g(\bar{x}_1 + \bar{x}_2) \geq g(\bar{x}_1) + g(\bar{x}_2)$. Hence substituting $\bar{x}_1 = x_1$ and $\bar{x}_2 = x_2 - x_1$ yields (3.30). □

Lemma 3.3. *Consider the function T_i defined in* (3.29)*, the functions $\alpha_i \in \mathcal{K}_\infty$, $\beta_{ji} \in \mathcal{K}_\infty$, and $\delta_i \in \mathcal{K}_\infty$ as in* (3.28)*, and a function $\underline{\alpha}_i \in \mathcal{K}_\infty$ such that $\underline{\alpha}_i(s) \leq \alpha_i(s)$, $\forall s \geq 0$. Then the condition*

$$T_i \circ \frac{1}{p_i}\underline{\alpha}_i - \sum_{j\in\mathcal{S}_i} T_j \circ \beta_{ji} = \delta_i \tag{3.31}$$

implies (3.28)*.*

[4] The transform is known as the Legendre-Fenchel transform and it is used in [69] for simplifying ISS/iISS Lyapunov functions.

Proof. By the strict monotonicity of T_i, equation (3.31) implies

$$T_i \circ \frac{1}{p_i}\alpha_i - \sum_{j\in\mathcal{S}_i} T_j \circ \beta_{ji} = \bar{\delta}_i, \tag{3.32}$$

and noting that $p_i \geq 1$ further implies

$$p_i T_i \circ \frac{1}{p_i}\alpha_i - \sum_{j\in\mathcal{S}_i} T_j \circ \beta_{ji} = \bar{\bar{\delta}}_i, \tag{3.33}$$

where $\bar{\delta}_i$ and $\bar{\bar{\delta}}_i$ are some positive definite functions and $\bar{\bar{\delta}}_i(s) \geq \bar{\delta}_i(s) \geq \delta_i(s)$, $\forall s \geq 0$. This completes the proof. □

Before proceeding note that T_i is a convex function, since $T_i' = f_i^{-1}$ is strictly increasing as required by the condition in Lemma A.2. We treat convexity as a constraint on T_i to be enforced later. To introduce a recursive algorithm to "solve" for T_i, we re-write (3.32) using a matrix-like representation and take a fully connected 3-node system as an example, that is

$$\left[\begin{array}{ccc} T_1 \circ \frac{1}{p_1}\alpha_1 & -T_1 \circ \beta_{12} & -T_1 \circ \beta_{13} \\ -T_2 \circ \beta_{21} & T_2 \circ \frac{1}{p_2}\alpha_2 & -T_2 \circ \beta_{23} \\ -T_3 \circ \beta_{31} & -T_3 \circ \beta_{32} & T_3 \circ \frac{1}{p_3}\alpha_3 \\ \hline \bar{\delta}_1 & \bar{\delta}_2 & \bar{\delta}_3 \end{array}\right] \triangleq \left[\begin{array}{c} E_T^{\mathrm{I}} \\ \hline \sigma^{\mathrm{I}} \end{array}\right]. \tag{3.34}$$

Applying Gaussian elimination to the first element of the second and the third column using right composition (instead of multiplication) yields

$$\left[\begin{array}{ccc} T_1 \circ \frac{1}{p_1}\alpha_1 & 0 & 0 \\ -T_2 \circ \beta_{21} & (E_T^{\mathrm{II}^0})_{22} & (E_T^{\mathrm{II}})_{23} \\ -T_3 \circ \beta_{31} & (E_T^{\mathrm{II}})_{32} & (E_T^{\mathrm{II}})_{33} \\ \hline \bar{\delta}_1 & (\sigma^{\mathrm{II}})_2 & (\sigma^{\mathrm{II}})_3 \end{array}\right]. \tag{3.35}$$

where

$$
\begin{aligned}
(E_T^{\mathrm{II}^0})_{22} &\triangleq T_2 \circ \frac{1}{p_2}\alpha_2 - T_2 \circ \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12},\\
(E_T^{\mathrm{II}})_{23} &\triangleq \ -T_2 \circ \beta_{23} - T_2 \circ \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{13},\\
(E_T^{\mathrm{II}})_{32} &\triangleq \ -T_3 \circ \beta_{32} - T_3 \circ \beta_{31} \circ \alpha_1^{-1} \circ p_1\beta_{12},\\
(E_T^{\mathrm{II}})_{33} &\triangleq T_3 \circ \frac{1}{p_3}\alpha_3 - T_3 \circ \beta_{31} \circ \alpha_1^{-1} \circ p_1\beta_{13},\\
(\sigma^{\mathrm{II}^0})_2 &\triangleq \bar{\delta}_2 + \bar{\delta}_1 \circ \alpha_1^{-1} \circ p_1\beta_{12},\\
(\sigma^{\mathrm{II}})_3 &\triangleq \bar{\delta}_3 + \bar{\delta}_1 \circ \alpha_1^{-1} \circ p_1\beta_{13}.
\end{aligned}
\tag{3.36}
$$

If all operators are linear, the second pivot equals $T_2 \circ \big(\frac{1}{p_2}\alpha_2 - \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}\big)$ due to the left distributivity of multiplication, but since this does not hold for general composition operation (that is, in general $T_2 \circ \frac{1}{p_2}\alpha_2 - T_2 \circ \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12} \neq T_2 \circ (\frac{1}{p_2}\alpha_2 - \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12})$), additional care is needed. Define

$$
\hat{\alpha}_2 \triangleq \alpha_2 - p_2\beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}. \tag{3.37}
$$

Suppose that there exists $\underline{\hat{\alpha}}_2$ such that $\hat{\alpha}_2 \geq \underline{\hat{\alpha}}_2 \in \mathcal{K}_\infty$. Then $\frac{1}{p_2}\alpha_2 > \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}$ and, by Lemma 3.2,

$$
\begin{aligned}
&T_2 \circ \frac{1}{p_2}\alpha_2 - T_2 \circ \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}\\
&\geq T_2 \circ \left(\frac{1}{p_2}\alpha_2 - \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}\right)\\
&= T_2 \circ \frac{1}{p_2}\hat{\alpha}_2 \geq T_2 \circ \frac{1}{p_2}\underline{\hat{\alpha}}_2.
\end{aligned}
\tag{3.38}
$$

Due to the monotonicity of T_2, replacing $T_2 \circ \frac{1}{p_2}\alpha_2 - T_2 \circ \beta_{21} \circ \alpha_1^{-1} \circ p_1\beta_{12}$ in (3.35) with $T_2 \circ \frac{1}{p_2}\underline{\hat{\alpha}}_2$ effectively replaces α_2 in (3.34) with some $\underline{\alpha}_2 \in \mathcal{K}_\infty$ such that $\underline{\alpha}_2 \leq \alpha_2$ and, by Lemma 3.3, we can rewrite (3.35) as

$$
\left[\begin{array}{ccc}
T_1 \circ \frac{1}{p_1}\underline{\hat{\alpha}}_1 & 0 & 0\\
-T_2 \circ \beta_{21} & (E_T^{\mathrm{II}})_{22} & (E_T^{\mathrm{II}})_{23}\\
-T_3 \circ \beta_{31} & (E_T^{\mathrm{II}})_{23} & (E_T^{\mathrm{II}})_{33}\\
\hline
\delta_1 & (\sigma^{\mathrm{II}})_2 & (\sigma^{\mathrm{II}})_3
\end{array}\right]
\triangleq
\left[\begin{array}{c}
E_T^{\mathrm{II}}\\
\hline
\sigma^{\mathrm{II}}
\end{array}\right],
\tag{3.39}
$$

where

$$
\begin{aligned}
(E_T^{\mathrm{II}})_{22} &\triangleq T_2 \circ \frac{1}{p_2}\underline{\hat{\alpha}}_2, \\
(\sigma^{\mathrm{II}})_2 &\triangleq \delta_2 + \delta_1 \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{12},
\end{aligned} \tag{3.40}
$$

$\delta_1 = \bar{\delta}_1$, $0 < \delta_2 \leq \bar{\delta}_2$, $\underline{\hat{\alpha}}_1 \triangleq \alpha_1$, and $\underline{\hat{\alpha}}_2 \in \mathcal{K}_\infty$ is constructed such that $\underline{\hat{\alpha}}_2 \leq \hat{\alpha}_2$. Exploiting the condition that $\underline{\hat{\alpha}}_2 \in \mathcal{K}_\infty$, we can continue to apply Gaussian elimination to the second term on the third column. The resulting matrix is

$$
\left[\begin{array}{ccc}
T_1 \circ \frac{1}{p_1}\underline{\hat{\alpha}}_1 & 0 & 0 \\
-T_2 \circ \beta_{21} & (E_T^{\mathrm{II}})_{22} & 0 \\
-T_3 \circ \beta_{31} & (E_T^{\mathrm{II}})_{32} & (E_T^{\mathrm{III}})_{33} \\
\hline
\delta_1 & (\sigma^{\mathrm{II}})_2 & (\sigma^{\mathrm{III}})_3
\end{array}\right] \triangleq \left[\begin{array}{c} E_T^{\mathrm{III}} \\ \hline \sigma^{\mathrm{III}} \end{array}\right], \tag{3.41}
$$

where

$$
\begin{aligned}
(E_T^{\mathrm{III}})_{33} &\triangleq T_3 \circ \frac{1}{p_3}\underline{\hat{\alpha}}_3, \\
(\sigma^{\mathrm{III}})_{13} &\triangleq \delta_3 + \delta_1 \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13} + \delta_2 \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{23} \\
&\quad + \delta_1 \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{12} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{23} \\
&\quad + \delta_2 \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{21} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13} \\
&\quad + \delta_1 \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{12} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{21} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13},
\end{aligned} \tag{3.42}
$$

and $0 < \delta_3 \leq \bar{\delta}_3$. $\underline{\hat{\alpha}}_3 \in \mathcal{K}_\infty$ is constructed such that $\underline{\hat{\alpha}}_3 \leq \hat{\alpha}_3$ (assume that this is possible for the time being), where

$$
\begin{aligned}
\hat{\alpha}_3 \triangleq \alpha_3 &- p_3\beta_{31} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13} - p_3\beta_{32} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{23} \\
&- p_3\beta_{31} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{12} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{23} \\
&- p_3\beta_{32} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{21} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13} \\
&- p_3\beta_{31} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{12} \circ \underline{\hat{\alpha}}_2^{-1} \circ p_2\beta_{21} \circ \underline{\hat{\alpha}}_1^{-1} \circ p_1\beta_{13}.
\end{aligned} \tag{3.43}
$$

It is claimed that as long as there exist $\mathcal{K}_\infty$-class functions $\underline{\hat{\alpha}}_1$, $\underline{\hat{\alpha}}_2$, $\underline{\hat{\alpha}}_3$ that bound $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$, respectively, from below, one can construct convex T_1, T_2, T_3 by properly selecting

$\delta_1, \delta_2, \delta_3 \in \mathcal{K}_\infty$. A simple way to see this is to solve (3.41) for T_1, T_2, T_3 in terms of δ_1, δ_2, δ_3. These allow writing T_1, T_2, T_3 as positive sums of composition terms led by δ_1, δ_2, δ_3. By properly selecting δ_1, δ_2, δ_3, these positive terms can be individually made convex, and the convexity is preserved under positive summation, which further justifies the use of Lemma A.2. Finally, using (3.27) and (3.29) yields

$$\lambda_i(V_i) = \frac{\mathrm{d}T_i}{\mathrm{d}s} \circ \frac{1}{p_i}\alpha_i \circ V_i, \tag{3.44}$$

which is the state-dependent scaling function needed.

For large-scale systems it is not practically useful to derive a closed-form expression for $\hat{\alpha}_i$ as it can be shown that the total number of the composition terms in $\hat{\alpha}_i$ is described by the *Sylvester's sequence* [123] subtracted by one[5], which grows doubly exponentially with n, the number of node systems. Instead, we seek a recursive algorithm that equivalently performs iterative Gaussian elimination (as demonstrated above) and at the same time reveals the underlying topological meaning of the terms in $\hat{\alpha}_i$. This recursive algorithm is described in Algorithm 3.1, the main idea of which is that each entry, indexed by (i, j), of E_T (the lower triangular matrix defined by (3.41)), contains the composition terms formed by a term from an upper entry (k, j) in the same column, the inverse of $\underline{\hat{\alpha}}_k$ (which is the lower-bound of $\hat{\alpha}_k$), and a term from a left entry (i, k) in the same row, sequentially composed from right to left, and such a construction holds recursively. From the aforementioned iterative Gaussian elimination procedures, it is not difficult to see the reasoning behind: each term in the (i, j) entry was brought from the (i, k) entry as a side effect of eliminating the (k, j) entry using $\underline{\hat{\alpha}}_k$. The computation of $(\sigma^n)_k$ is straightforward once $\hat{\alpha}_k$ has been computed. Since $(\sigma^n)_k$ is derived by the same Gaussian elimination procedure as the one that derives $\hat{\alpha}_k$, $(\sigma^n)_k$ is simply the positive sum of each composition term in $\hat{\alpha}_k$, after removing the sign and replacing the first function of each composition term with the δ_l, where l is the column index of the first function.

With the help of the recursive Algorithm 3.1, we can conclude the following small-gain-like result.

[5]To see this, first note that the expression of $\hat{\alpha}_i$ is a sum-of-composition formula and let N_i denote the number of composition terms. Then one can find that $N_1 = 1$ and $N_n = 1 + \sum_{i=1}^{n-1} N_i^2$, or equivalently $M_n = M_{n-1}^2 - M_{n-1} + 1$, with $M_n \triangleq N_n + 1$. It can be proved that $M_n = \lfloor k^{2^n} + \frac{1}{2} \rfloor$, where $k \approx 1.264085$, and $\lfloor \cdot \rfloor$ denotes the floor function. See [1] for detail.

Algorithm 3.1 Recursive algorithm to compute the lower triangular matrix M_T^n and the pivot gain functions $\hat{\alpha}$ and $\underline{\hat{\alpha}}$ (see Algorithm 3.2 for the helper functions).

Input: E_T^{I}
Output: E_T^n, $\hat{\alpha}$, and $\underline{\hat{\alpha}}$

```
Initialize E_T^n to an empty matrix of the same size as E_T^I
for each (E_T^n)_ij in E_T^n do
    if i < j then
        (E_T^n)_ij ← 0                        ▷ The upper triangular entries are eliminated
    else if i = j then
        α̂_i ← COMPUTEENTRY(i, i, E_T^n)
        α̂_i ← p_i · REMOVET(i, α̂_i)
        Find α̲̂_i ∈ K_∞ such that α̲̂_i ≤ α̂_i
        (E_T^n)_ii ← T_i ∘ (1/p_i) α̲̂_i       ▷ The diagonal entries are expressed as T_i ∘ (1/p_i) α̲̂_i due to Lemma 3.2 and 3.3
    else
        (E_T^n)_ij ← COMPUTEENTRY(i, j, E_T^n)  ▷ Compute the lower triangular entries without merging the T_i terms
    end if
end for
```

Theorem 3.4. *Consider a network system with node dissipation inequalities given by* (3.16). *If there exists* $\underline{\hat{\alpha}}_i \in \mathcal{K}_\infty$ *such that* $\underline{\hat{\alpha}}_i \leq \hat{\alpha}_i$, $i = 1, \ldots, n$, *with* $\hat{\alpha}_i$ *computed by Algorithm 3.1, then there exist continuous functions* $\lambda_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $i = 1, \ldots, n$ *such that* (3.17) *and* (3.18) *hold, and the overall dissipation inequality defined by* (3.19) *satisfies*

$$\dot{V} \leq -\sum_{i=1}^{n} \delta_i(V_i) < 0, \tag{3.45}$$

for $V_i > 0$, $i = 1, \ldots, n$.

Proof. Given the existence of the $\mathcal{K}_\infty$-class functions $\underline{\hat{\alpha}}_i$, $i = 1, \ldots, n$, Algorithm 3.1 is feasible and its output is a lower-triangular matrix E_T^n and the functions $\hat{\alpha}_k$, which further yields the functions $(\sigma^n)_k$. For ease of expression, in what follows we drop the superscript n without causing ambiguity, since all the variables used have the same superscript. Let σ denote $[(\sigma^n)_1, \ldots, (\sigma^n)_n]^\top$. From the previous analysis we know that

$$\mathit{1}^\top E_T = \sigma^\top. \tag{3.46}$$

Algorithm 3.2 Helper functions for Algorithm 3.1.

function COMPUTEENTRY(i, j, E_T)
  if $i = 1$ or $j = 1$ **then** ▷ *Base Case*
    return $(E_T)_{ij}$
  else ▷ *Recursive Case*
    $\Xi \leftarrow (E_T)_{ij}$ ▷ *ξ stores the expression to be returned*

    for $k = 1$ to $\min(i-1, j-1)$ **do**
      $\Xi_L \leftarrow$ COMPUTEENTRY(i, k, E_T)
      ▷ *Ξ_L: expression of the (i,k)-entry (left expression)*
      $\Xi_U \leftarrow$ COMPUTEENTRY(k, j, E_T)
      $\Xi_U \leftarrow -$REMOVET(k, Ξ_U)
      ▷ *Ξ_U: expression of the (k,j)-entry (upper expression), with T_k and the minus sign of each term removed*
      $\hat{\alpha}_k \leftarrow$ COMPUTEENTRY(k, k, E_T)
      $\hat{\alpha}_k \leftarrow p_k \cdot$REMOVET$(k, \hat{\alpha}_k)$
      Find $\underline{\hat{\alpha}}_k \in \mathcal{K}_\infty$ such that $\underline{\hat{\alpha}}_k \leq \hat{\alpha}_k$
      for each ξ_U in Ξ_U **do**
        for each ξ_L in Ξ_L **do**
          $\Xi \leftarrow \Xi + \xi_L \circ \underline{\hat{\alpha}}_k^{-1} \circ (p_k \cdot \xi_U)$ ▷ *This accounts for the effect of eliminating the (k,j)-entry on the (i,j)-entry*
        end for
      end for
    end for
    return Ξ
  end if
end function

function REMOVET(i, Ξ)
  for each ξ in Ξ **do**
    remove T_i from ξ
  end for
  return Ξ
end function

Each element in σ is a positive sum of composition terms, with each composition term led by one of the δ_i, $i = 1, \ldots, n$. In addition, the off-diagonal elements of E_T are negative sums of composition terms led by one of the T_i's, thus $T_1, \ldots, T_n$ resulting from (3.46) can be written as positive sums of composition terms led by one of the δ_i's. Note that to meet the conditions of Lemma A.2, $T_1, \ldots, T_n$ are required to be strictly convex: this is to guarantee that their derivatives $f_1^{-1}, \ldots, f_n^{-1}$ are strictly increasing. This is enforced by properly selecting $\delta_i \in \mathcal{K}_\infty$ such that each individual composition term is strictly convex. Since convexity is preserved under positive summation, the resulting

$T_1, \ldots, T_n$ are also strictly convex. Compute $\lambda_i(V_i)$ using (3.44), and consider the overall storage function (3.19). By Lemma 3.3,

$$\begin{aligned}\dot{V} \leq & -\sum_{i=1}^{n}\left(\lambda_i(V_i)\alpha_i(V_i) - \sum_{j\in\mathcal{P}_i}\lambda_i(V_i)\beta_{ij}(V_j)\right)\\ \leq & -\sum_{i=1}^{n}\left(p_i T_i \circ \frac{1}{p_i}\alpha_i \circ V_i - \sum_{j\in\mathcal{S}_i} T_j \circ \beta_{ji} \circ V_i\right)\\ \leq & -\sum_{i=1}^{n}\delta_i(V_i) \end{aligned} \tag{3.47}$$

and $\dot{V} < 0$, for $V_i > 0$, $i = 1, \ldots, n$, which completes the proof. □

Remark 3.3. *One can regard $\underline{\hat{\alpha}}_i \in \mathcal{K}_\infty$ as the counterpart of the small-gain condition for network systems, since such a condition guarantees the existence of λ_i such that $\dot{V} < 0$. For a two-node system, where $\hat{\alpha}_1 \triangleq \alpha_1$, $\hat{\alpha}_2 \triangleq \alpha_2 - \beta_{21} \circ \alpha_1^{-1} \circ \beta_{12}$, $\alpha_i \in \mathcal{K}_\infty$, $\beta_{ij} \in \mathcal{K}_\infty$, $i = 1, 2$, the condition $\hat{\alpha}_1 \geq \underline{\hat{\alpha}}_1 \in \mathcal{K}_\infty$ is satisfied by defining $\underline{\hat{\alpha}}_1 \triangleq \alpha_1$ and $\hat{\alpha}_2 \geq \alpha_2 \in \mathcal{K}_\infty$ is implied by the classical small-gain condition*

$$\alpha_2^{-1} \circ \tau_2\beta_{21} \circ \alpha_1^{-1} \circ \tau_1\beta_{12} < \mathrm{Id}, \tag{3.48}$$

where $\tau_1 > 1$, $\tau_2 > 1$, and Id *is the identity mapping. To see this, doing some manipulations on* (3.48) *yields $\hat{\alpha}_2 = \alpha_2 - \beta_{21}\circ\alpha_1^{-1}\circ\beta_{12} > \alpha_2 - \beta_{21}\circ\alpha_1^{-1}\circ\tau_1\beta_{12} > (\tau_2 - 1)\beta_{21}\circ\alpha_1^{-1}\circ\beta_{12}$. Then defining $\underline{\hat{\alpha}}_2 \triangleq (\tau_2 - 1)\beta_{21} \circ \alpha_1^{-1} \circ \beta_{12} \in \mathcal{K}_\infty$ reveals that* (3.48) *implies the condition $\hat{\alpha}_2 \geq \underline{\hat{\alpha}}_2$.*

Similarly to Definition 3.3, one can define the notion of *active nodes*, provided there is an adjustable α function for a node.

Definition 3.4. *A node Σ_i is called an active node, denoted $i \in \mathcal{I}_A$, if it satisfies the dissipation inequality* (3.16) *with an adjustable $\alpha_i \in \mathcal{K}_\infty$, and $\alpha_i \geq \underline{\alpha}_i$, for some positive definite function $\underline{\alpha}_i$.*

It turns out that one can select the α_{i_A}, $i_A \in \mathcal{I}_A$, to let the algebraic condition of Theorem 3.4 hold, and the feasibility of this is closely related to the location of the *active nodes* in the network system, even though the condition and the proof of Theorem 3.4 are completely different from Theorem 3.2. In other words, there is a nonlinear counterpart of Theorem 3.3, as stated below.

Theorem 3.5. *The condition of Theorem 3.4, namely, the existence of* $\underline{\hat{\alpha}}_i \in \mathcal{K}_\infty$*, holds for some admissible class* $\mathcal{K}_\infty$ *functions* α_{i_A}*,* $i_A \in \mathcal{I}_A$*, if every directed cycle of the underlying directed graph contains at least one vertex associated with an active node.*

Proof. To prove the claim note the graph-theoretic interpretation of the COMPUTEENTRY function of Algorithm 3.2: for each vertex i, the expression of $\hat{\alpha}_i$ is obtained from α_i subtracting several composition terms, with each composition term associated to one of the closed paths starting from vertex i and returning to vertex i, regarding the off-diagonal elements of E_T^{I} as an analogy to the adjacency matrix. The upper expression ξ_U contains the outgoing path, and the left expression ξ_L contains the incoming path, and such interconnection is made recursively. In this sense, $\hat{\alpha}_i$ can be written as

$$\hat{\alpha}_i = \alpha_i - \sum_{l=1}^{l_{\max}} \left(p_i \beta_{ik_m^l} \circ \underline{\hat{\alpha}}_{k_m^l}^{-1} \circ p_{k_m^l} \beta_{k_m^l k_{m-1}^l} \circ \cdots \circ \underline{\hat{\alpha}}_{k_1^l}^{-1} \circ p_{k_1^l} \beta_{k_1^l i} \right), \tag{3.49}$$

where $k_j^l \in \mathcal{C}_l^i$, $\mathcal{C}_l^i$ is the index set of the lth circuit path[6] from vertex i to itself, in the subgraph consisting of vertices 1 to i and the associated edges; $l_{\max}$ is the total number of such sets; $m \triangleq |\mathcal{C}_l^i|$; and $k_j^l \in \mathcal{S}_{k_{j-1}^l}$. We proceed by invoking Lemma A.6, which indicates that if α_i is adjustable in the sense stated in Definition 3.4, then $\underline{\hat{\alpha}}_i$ is also adjustable in the same sense. Since the negative terms in (3.49) are associated with circuit paths, having at least an *active node* in every cycle implies that there is at least an *active node* in every circuit, and therefore the negative terms in (3.49) can be made arbitrarily small by adjusting $\underline{\hat{\alpha}}_{i_A}$ functions. Thus, any $\alpha_i \in \mathcal{K}_\infty$ can dominate $\hat{\alpha}_i$ and guarantee the existence of $\underline{\hat{\alpha}}_i \in \mathcal{K}_\infty$ such that $\underline{\hat{\alpha}}_i \leq \hat{\alpha}_i$, for $i = 1, \ldots, n$. Hence the conditions of Theorem 3.4 hold, which completes the proof. □

3.3 Systems with Linearly Parametrized Supply Rates

Section 3.2 has shown that the notion of *active nodes* is applicable to a generic class of nonlinear systems with sum-type dissipation inequalities. While on one hand, this

[6]The term "circuit path" is used here since the path may not be a simple circuit, or equivalently a cycle: it can contain nested cycles. See (3.43) for an example: the last term indicates the path $3 \to 1 \to 2 \to 1 \to 3$, which contains two cycles.

extends the results in Section 3.1 to more general scenarios, on the other hand, it makes the formulation of the damping functions α_i abstract and therefore difficult to be used in practical designs. Thus in this subsection, we consider a "middle ground" between the schemes discussed in Section 3.1 and Section 3.2 which allows parametrizing the supply rates and allows the implementation of standard nonlinear control design techniques that are not well catered by the quadratic supply rate case.

To start with, consider a two-node system with storage functions $V_1 = y_1^2$, $V_2 = y_2^2$, and dissipation inequalities

$$\begin{aligned}\dot{V}_1 &\leq -a_1 y_1^2 + b_{12} y_2^2 - a_3 y_1^4, \\ \dot{V}_2 &\leq b_{21} y_1^2 - a_2 y_2^2 + b_{23} y_1^4,\end{aligned} \tag{3.50}$$

where $y_1 \in \mathbb{R}$, $y_2 \in \mathbb{R}$. Obviously, the selection $\phi(y) = [y_1^2, y_2^2]^\top$ in the spirit of Section 3.1 does not work for this system since it does not take the y_1^4-terms into account. In the spirit of Section 3.2 one can write $\alpha_1(s) = a_1 s + a_3 s^2$, $\alpha_2(s) = a_2 s$, $\beta_{12}(s) = b_{12} s$, $\beta_{21} = b_{21} s$, which properly describes the system. Nevertheless, the parametrization of α_1, even though a_1 and a_2 can be selected arbitrarily, does not fulfil the condition for an *active node* specified by Definition 3.4. This suggests finding an alternative formulation to allow exploiting parametrization of node dissipation inequalities. Alternatively, based on the formulation in Section 3.1, we could augment ϕ with an extra positive definite term, in this case, y_1^4, and define the augmented vector $\hat{\phi}(y) = [y_1^2, y_2^2, y_1^4]^\top$. Compared to the case in Section 3.1, in which each term of the supply rate is one-to-one related to a node, now y_1^2 and y_1^4 are both related to the same node, node Σ_1. In other words, node Σ_1 has two supply rate basis functions (basis functions for short), y_1^2 and y_1^4, while node Σ_2 has only one basis function y_2^2.

To generalize this idea, suppose that there are q_i basis functions for the dissipation inequality of node Σ_i and $\hat{q} \triangleq \sum_{i=1}^n q_i$ basis functions in total. Define $\hat{\phi}(y) = [\phi^\top(y), \varphi_1^2(y_1), \ldots, \varphi_1^{q_1}(y_1), \ldots, \varphi_n^2(y_n), \ldots, \varphi_n^{q_n}(y_n)]^\top \in \mathbb{R}^{\hat{q}}$, where $\varphi_i^j(y_i)$ denotes the j-th basis function of node Σ_i, and $\phi(y) = [\varphi_1^1(y_1), \ldots, \varphi_n^1(y_n)]^\top \in \mathbb{R}^n$ is the vector containing the primary basis functions[7] of each node, with the other elements in $\hat{\phi}(y)$ referred to as

[7] One can select any one of the basis functions associated with a node as the "primary" basis function of the node. The rest of them are treated as the "secondary" basis functions.

secondary basis functions. Under this definition of $\hat{\phi}(y)$, the node dissipation inequalities are

$$\dot{V}_i \leq -\sum_{k=1}^{q_i} a_i^k \varphi_i^k(y_i) + \sum_{j=1, j\neq i}^{n} \left(\sum_{k=1}^{q_j} b_{ij}^k \varphi_j^k(y_j) \right), \tag{3.51}$$

with $a_i^k > 0$, $b_{ij}^k > 0$, and these can be written into a vector form similarly to (3.2), namely

$$\dot{\boldsymbol{V}} \leq \; -\bar{E}\hat{\phi}, \tag{3.52}$$

where $\bar{E} \in \mathbb{R}^{n\times\hat{q}}$ contains the coefficients of the node dissipation inequalities with opposite signs. With the same construction of the overall storage function as (3.6), the overall dissipation inequality can be derived as

$$\dot{V} \leq -c^\top \bar{E}\hat{\phi}. \tag{3.53}$$

Define the left $n \times n$ submatrix of $\bar{E}$ as E, and the right $n \times (\hat{q} - n)$ submatrix as $\tilde{E}$. For example, the matrix $\bar{E}$ associated with the two-node system with dissipation inequalities (3.50) is

$$\bar{E} = \left[\begin{array}{cc|c} a_1 & -b_{12} & a_3 \\ -b_{21} & a_2 & -b_{23} \end{array} \right] \triangleq \left[\begin{array}{c|c} E & \tilde{E} \end{array} \right]. \tag{3.54}$$

To make the notation natural in the matrix $\bar{E}$ we remove the superscript of the coefficients in the matrix expression (3.51) and replace the subscript of $a_{(\cdot)}$ and the second subscript of $b_{(\cdot)}$ (indicating the predecessor node) with the index of the corresponding basis function in $\hat{\phi}(y)$ (or equivalently, the column index of the coefficient in $\bar{E}$). This defines an alternative set of indices $\hat{\imath}(i, k)$ for coefficients associated with the kth basis function of node Σ_i. For example in (3.54) we have $\hat{\imath}(1, 2) = 3$ for the second basis function of node Σ_1 and therefore we replace a_1^2 with a_3 and replace b_{21}^2 with b_{23}. Since $\bar{E}$ is not a *Z-matrix*, if we want to use the results established in Section 3.1 it is better to restore the one-to-one relation between each basis function and each matrix dimension by augmenting $\bar{E}$ so that we have a *Z-matrix* to analyze. From a graph theoretic perspective, this requires adding some *augmented* vertices to the underlying directed graph so that the total number of vertices is the same as the basis functions rather than the number of nodes. To this end, rewrite the overall dissipation inequality as

$$\dot{V} \le -\hat{c}^\top \hat{E}\hat{\phi}, \tag{3.55}$$

where $\hat{E}$ is an $\hat{q} \times \hat{q}$ *Z-matrix* augmented from $\bar{E}$, and $\hat{c} \in \mathbb{R}^{\hat{q}}$ is augmented from c. More specifically, $\hat{E}$ can be written as

$$\hat{E} = \left[\begin{array}{c|c} E & \tilde{E} - U\mathrm{diag}(a_{n+1}, \ldots, a_{\hat{q}}) \\ \hline 0_{(\hat{q}-n)\times n} & \mathrm{diag}(a_{n+1}, \ldots, a_{\hat{q}}) \end{array}\right] \tag{3.56}$$

where $U \in \mathbb{R}^{n\times(\hat{q}-n)}$ is defined as

$$(U)_{ij} \triangleq \begin{cases} 1, & \text{if } (\tilde{E})_{ij} = a_{(\cdot)} > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.57}$$

In other words, the underlying graph described by $\hat{E}$ is obtained by adding vertices (indexed by $\hat{\imath}$) associated with the secondary basis functions, and then by connecting these augmented vertices to the graph described by E according to the dissipation inequalities, which is associated with the primary basis functions and the original network. Since the scaling operation is implemented on the n node systems, not on the $\hat{q} - n$ augmented vertices, as they do not originate from new node systems, the last $\hat{q} - n$ elements of the augmented scaling vector $\hat{c}$ are generated by the first n elements, that is, the original scaling vector c. It is not difficult to find that this fact can be described by the constraint

$$L\hat{c} = 0, \tag{3.58}$$

where $L \in \mathbb{R}^{(\hat{q}-n)\times n}$ is defined by

$$L \triangleq \left[\begin{array}{c|c} U^\top & -I_{\hat{q}-n} \end{array}\right]. \tag{3.59}$$

Consider again the two-node system (3.50). Then $U = [1, 0]^\top$ and the augmented version of (3.54) is

$$\hat{E} = \begin{bmatrix} a_1 & -b_{12} & 0 \\ -b_{21} & a_2 & -b_{23} \\ 0 & 0 & a_3 \end{bmatrix}. \tag{3.60}$$

Since the third column of $\hat{M}$ comes from the first dissipation inequality, which should be multiplied by c_1, we also need to define $\hat{c} = [c_1, c_2, c_3]^\top$, with the constraint $c_1 = c_3$ or equivalently, $L\hat{c} = [1, 0, -1]\hat{c} = c_1 - c_3 = 0$.

To derive the counterpart of Theorem 3.2 we have to first answer two questions: 1) how to determine whether $\hat{E}$ is a non-singular *M-matrix* by checking the original matrix E; and 2) how to use condition 6) of Theorem 3.1 considering the additional constraint (3.58). The answer to the first question is given by the result below.

Lemma 3.4. *$\hat{E}$ is a non-singular M-matrix if and only if E is a non-singular M-matrix.*

Proof. Consider the block triangular structure of the matrix $\hat{E}$ defined by (3.56). The leading principal minors with order higher than n have the same sign as $\det(E)$. Therefore using condition 4 of Theorem 3.1 completes the proof. □

To answer the second question we need first to clarify the problem we are trying to deal with. Condition 6) of Theorem 3.1 guarantees the existence of $\hat{c} > 0$ such that $\hat{E}^\top \hat{c} > 0$ if $\hat{E}$ is a non-singular *M-matrix*, which, however, is not sufficient in this case as there is the additional constraint (3.58) on $\hat{c}$. Thus, we need to add an additional condition to $\hat{E}$ such that $\hat{c}$ also satisfies the constraint (3.58). This leads to the following result.

Lemma 3.5. *Consider the non-singular M-matrix $\hat{E}$. There exists $\hat{c} > 0$ such that $\hat{E}^\top \hat{c} > 0$ and $L\hat{c} = 0$ if and only if there exists $\hat{\sigma} > 0$ such that $\hat{\sigma}$ is in the kernel of $L(\hat{E}^\top)^{-1}$.*

Proof. To prove the "if" part assume that there exists $\hat{\sigma} > 0$ such that $L(\hat{E}^\top)^{-1}\hat{\sigma} = 0$. Then selecting $\hat{c} = (\hat{E}^\top)^{-1}\hat{\sigma}$ satisfies the constraint $L\hat{c} = 0$. Due to condition 5) of Theorem 3.1, $(\hat{E}^\top)^{-1}$ is non-singular and $(\hat{E}^\top)^{-1} \geq 0$, and therefore $\hat{c} > 0$. Left-multiplying $\hat{c}$ by $\hat{E}^\top$ yields $\hat{E}^\top \hat{c} = \hat{E}^\top (\hat{E}^\top)^{-1}\hat{\sigma} = \hat{\sigma} > 0$, which completes proof.

To prove the "only if" part assume that there exists $\hat{c} > 0$ such that $\hat{E}^\top \hat{c} > 0$ and

$L\hat{c} = 0$, yet for all $\hat{\sigma} > 0$, $L(\hat{E}^\top)^{-1}\hat{\sigma} \neq 0$. Then we can select $\hat{\sigma} = \hat{E}^\top \hat{c} > 0$, and therefore $L(\hat{E}^\top)^{-1}\hat{\sigma} = L\hat{c} \neq 0$, which causes a contradiction. Hence the "only if" part is proved. □

Having proved Lemma 3.4 and Lemma 3.5 we can proceed to give a criterion for the existence of scaling coefficients such that the dissipation inequality (3.7) holds.

Theorem 3.6. *Consider the node dissipation inequalities given by* (3.52)*. There exists a vector of scaling coefficients $c > 0$ such that V constructed by* (3.6) *satisfies the dissipation inequality* (3.7) *if both the conditions below are satisfied.*

1. *The $n \times n$ leading principal submatrix of $\hat{E}$ is a non-singular M-matrix.*
2. *There exists $\hat{\sigma} > 0$ such that $\hat{\sigma}$ is in the kernel of $L(\hat{E}^\top)^{-1}$.*

Proof. By Lemma 3.4, condition 1) of the proposition implies that $\hat{E}$ is also a non-singular *M-matrix.* Then by invoking Lemma 3.5, condition 2) of the proposition guarantees that there exists $\hat{c} > 0$ such that $\hat{E}^\top \hat{c} = \hat{\sigma} > 0$ and $L\hat{c} = 0$. Recall that

$$\dot{V} \leq -\hat{c}^\top \hat{E}\hat{\phi}(y) = -\hat{\sigma}^\top \hat{\phi}(y) \leq 0, \tag{3.61}$$

in which case $\hat{\phi}^\top(y)\hat{\sigma}$ is the positive definite function $W(y)$ in (3.7). Finally construct $c > 0$ using the first n elements of $\hat{c}$, which completes proof. □

Theorem 3.6 reveals the fact that the $n \times n$ leading principal submatrix of $\hat{E}$ plays an important role. It is easier to understand this from a graph perspective. Since all the augmented columns in $\hat{E}$ have only one non-zero element on the diagonal entries, the associated vertices in the underlying directed graph have only outgoing edges but no incoming edges (see Fig. 3.2), which guarantees that no directed cycle contains these *augmented* vertices. In other words, the augmentation of the graph does not create new directed cycles and all directed cycles in the graph are specified by the $n \times n$ leading principal submatrix of $\hat{E}$.

Note now that each node can have more than one basis function, and therefore we need to slightly extend Definition 3.3.

Definition 3.5. *A node Σ_i is called an active node, denoted by $i \in \mathcal{I}_A$, if it satisfies the dissipation inequality* (3.51) *and for $k = 1, \dots, q_i$, the damping coefficients a_i^k are*

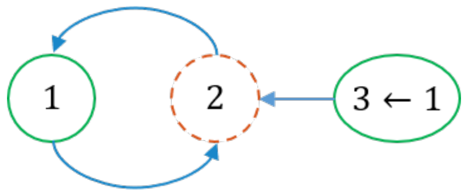

Figure 3.2: The underlying directed graph specified by (3.60)**. The notation "**$3 \leftarrow 1$**" means that vertex 3 is an augmented vertex which originates from node** Σ_1**.**

adjustable in $[\underline{a}_i^k, +\infty)$*, with* $\underline{a}_i^k > 0$*. The indices of all vertices (including the augmented vertices) of the underlying graph that originate from active nodes make up the set* $\hat{\mathcal{I}}_A$*.*

This definition allows the damping coefficient $a_{(\cdot)}$ of all vertices (including the augmented vertices) in the underlying directed graph that originate from *active nodes*, to be adjustable. For instance, in the underlying directed graph of the two-node example, both vertex 1 and vertex 3 originate from node Σ_1, and thus both a_1 and a_3 are adjustable if node Σ_1 is an *active node*. Having clarified this, we are ready to see how to enforce the dissipation inequality (3.7) with *active nodes*.

Theorem 3.7. *For all* $\hat{\sigma} > 0$ *there exists a selection of* $a_{i_A}^k$*,* $k = 1, \ldots, q_{i_A}$*, and a vector of scaling coefficients* $c > 0$*, depending on* $\hat{\sigma}$*, such that* V *constructed by* (3.6) *satisfies the dissipation inequality*

$$\dot{V} \leq -\hat{\sigma}^\top \hat{\phi}(y) \tag{3.62}$$

if both the conditions below are satisfied.

1. *Every directed cycle of the underlying directed graph describing the network contains at least one vertex that originates from an active node.*

2. *Every augmented vertex originates from an active node.*

Proof. We first consider the original n-vertex graph without the augmented vertices. Condition 1) of this proposition and Theorem 3.3 indicate that for all $\sigma > 0$ there exists a selection of $a_{i_A}^1$, $i_A \in \mathcal{I}_A$, that is, the damping coefficients associated with the primary basis function of the *active nodes*, and $c > 0$, depending on the choice of σ, such that the $n \times n$ leading principal submatrix E satisfies $E^\top c = \sigma$. Define $\hat{c} \triangleq [c^\top, \tilde{c}^\top]^\top$, with

$\tilde{c} \in \mathbb{R}^{\hat{q}-n}$ to be determined. Note that (3.58) and (3.59) yield $U^\top c - \tilde{c} = 0$, or equivalently, $\tilde{c} = U^\top c$. Since $U > 0$ and $c > 0$, we have $\tilde{c} > 0$, which provides a valid candidate for $\hat{c} > 0$.

We proceed to prove that for such a $\hat{c}$ we can select $a^k_{i_A}$, $k = 2, \ldots, q_{i_A}$, $i_A \in \mathcal{I}_A$, that is, the damping coefficient of the secondary basis functions of the *active nodes*, such that the claim holds. To see this, note that

$$\underbrace{\left[\begin{array}{c|c} E^\top & 0_{n\times(\hat{q}-n)} \\ \hline -\hat{E}_{12}^\top & \mathrm{diag}(a_{n+1}, \ldots, a_{\hat{q}}) \end{array}\right]}_{\hat{E}^\top} \underbrace{\left[\begin{array}{c} c \\ \hline \tilde{c} \end{array}\right]}_{\hat{c}} = \underbrace{\left[\begin{array}{c} \sigma \\ \hline \tilde{\sigma} \end{array}\right]}_{\hat{\sigma}}, \tag{3.63}$$

where $\hat{E}_{12} \triangleq U\mathrm{diag}(a_{n+1}, \ldots, a_{\hat{q}}) - \tilde{E}$ is independent of the $a_{(\cdot)}$ coefficients (cancelled by the subtraction) and only depends on the $b_{(\cdot)}$ coefficients and therefore $\hat{E}_{12} \geq 0$. Thus we have $-\hat{E}_{12}^\top c + [a_{n+1}, \ldots, a_{\hat{q}}]\tilde{c} = \tilde{\sigma}$, or equivalently, $[a_{n+1}, \ldots, a_{\hat{q}}]^\top = \big(\mathrm{diag}(\tilde{c})\big)^{-1}(\tilde{\sigma} + \hat{E}_{12}^\top c)$, for all $\tilde{\sigma} > 0$, which gives a vector of positive candidates for $a_{n+1}, \ldots, a_{\hat{q}}$. It remains to check whether these $a_{(\cdot)}$ are in the interval specified by Definition 3.5. If the resulting $a_{(\cdot)}$ is below the lowerbound, we replace the value with the lowerbound. This guarantees that $\hat{E}^\top \hat{c} \geq \hat{\sigma}$, and therefore $\dot{V} \leq -\hat{c}^\top \hat{E}\hat{\phi}(y) \leq -\hat{\sigma}^\top \hat{\phi}(y)$. Combining the arbitrariness of $\sigma > 0$ and $\tilde{\sigma} > 0$ we conclude that $\hat{\sigma} = [\sigma^\top, \tilde{\sigma}^\top]^\top > 0$ is also arbitrary. Hence the proof is complete. □

Remark 3.4. *It is not difficult to see that condition 1) of Theorem 3.7 allows enforcing condition 1) of Theorem 3.6 and condition 2) of Theorem 3.7 allows enforcing condition 2) of Theorem 3.6. In this sense, Theorem 3.6 and Theorem 3.7 view the same fact from an analysis perspective and a synthesis perspective, respectively. A similar remark also applies to Theorem 3.2 and Theorem 3.3.*

Remark 3.5. *Different selection of primary basis functions may result in different augmented graphs described by $\hat{E}$. It is worth noting that if the conditions in Theorem 3.7 hold for one of these graphs, they also hold for the other graphs. The reason for this is that, in this case, a node system associated with multiple augmented vertices is active due to Condition 2 of Theorem 3.7. Therefore when an augmented vertex is turned into a primary vertex due to a different selection of primary basis function, this vertex still originates from an active node and therefore verifies Condition 1 of Theorem 3.7 even if new* *cycles containing this vertex are created by that selection. Similarly, the previous primary vertex is turned into an augmented vertex, but since the node system discussed is active, this augmented vertex verifies Condition 2 of Theorem 3.7.*

3.4 Control Synthesis via Active Nodes

In Sections 3.1–3.3 we have focused on the existence of a scaling for the node storage functions to construct an overall network storage function that satisfies a dissipation inequality with a given dissipation margin, as well as the feasibility conditions based on the location of the *active nodes* to enforce such a dissipation inequality. In this subsection, we move forward to answer the "how to" part of the problem from a synthesis perspective.

3.4.1 Placement of Active Nodes

From Propositions 3.3, 3.5, and 3.7 we know that the *active nodes* should be placed in the network such that every directed cycle in the underlying graph contains at least one of the vertices associated with the *active nodes*. Hence placing the *active nodes*, from a graph-theoretic perspective, boils down to finding a set of vertices such that after the removal of these vertices (and the edges attached to them), the remaining graph is acyclic. This set is commonly known as the *feedback vertex (node) set* (FVS) in the graph theory and the computing theory literature. In general one may want to use a minimum number of active nodes to achieve the desired network dissipation inequality, which leads to the minimum FVS problem. This problem is proved to be *NP-complete* for directed graphs [66]. Many contributions have studied the exact solution of the minimum FVS problem, see *e.g.* [77, 105, 117]. We present here a method exploiting the *permanent* of a matrix to search for the minimum FVS. The *permanent* of a matrix $A \in \mathbb{R}^n$ is defined as

$$\operatorname{per}(A) = \sum_{\pi} \prod_{i=1}^{n} (A)_{\pi(i),i}, \tag{3.64}$$

where π is one of the permutations of the set $\{1, \dots, n\}$. A more intuitive interpretation is that $\operatorname{per}(A)$ can be computed by using Laplace expansion without switching the sign of the product terms as when computing $\det(A)$. As discussed in Section 3.1, each product term in $\det(E)$ is associated with directed cycles or unions of disjoint directed cycles of the underlying directed graph of E. Note that by defining the $n \times n$ matrix Ω_E as

$$(\Omega_E)_{ij} = \begin{cases} 1, & \text{if } (E)_{ij} \neq 0, \\ 0, & \text{otherwise,} \end{cases} \tag{3.65}$$

the number of Laplace expansion terms of $\det(E)$ equals $\mathrm{per}(\Omega_E)$. It is then natural to understand that vertex i is contained in $\big(\mathrm{per}(\Omega_E) - \mathrm{per}(\Omega_{Ei})\big)$ directed cycles and unions of disjoint directed cycles, where Ω_{Ei} is the matrix obtained by deleting the ith row and the ith column of Ω_E. This leads to the following lemma.

Lemma 3.6. *Consider E and its underlying graph. Let vertex i and vertex j be in the same directed cycle. Vertex i is contained in at least as many directed cycles as vertex j is, if and only if*

$$\mathrm{per}(\Omega_{\mathrm{Ei}}) \leq \mathrm{per}(\Omega_{\mathrm{Ej}}). \tag{3.66}$$

Proof. Let m_i and m_j be the total number of directed cycles and unions of disjoint directed cycles containing the vertex i and the vertex j, respectively. We have $m_i = \big(\mathrm{per}(\Omega_E) - \mathrm{per}(\Omega_{Ei})\big)$, and $m_j = \big(\mathrm{per}(\Omega_E) - \mathrm{per}(\Omega_{Ej})\big)$. Therefore, $m_i \geq m_j$ if and only if (3.66) holds. Note that m_i and m_j also count the number of unions of disjoint directed cycles, but since vertex i and vertex j are in the same directed cycle, the contribution of the considered cycle to the values m_i and m_j (via the considered cycle and the unions of the considered cycle with other disjoint directed cycles) are the same. Hence $m_i \geq m_j$ if and only if the vertex i is contained in at least as many directed cycles as the vertex j is, which completes the proof. □

This lemma indicates that one can find the most "important" vertex in a given directed cycle by determining the vertex which has the largest decrease in $\mathrm{per}(\Omega_E)$ after the deletion of itself and its attached edges, as this vertex is contained in more directed cycles than any other vertex in the considered direct cycle, and therefore should be included in the solution of the minimum FVS. Algorithm 3.3 is developed under this spirit to place the *active nodes* efficiently. The for-loop part of the algorithm is straightforward as this fulfils the requirement of Theorem 3.7. The while-loop part is essentially an exact algorithm to solve the underlying minimum FVS problem.

Algorithm 3.3 Algorithm to place the minimum number of *active nodes* in a given network system described by (3.52).

Input: The dissipation coefficient matrix $\bar{E}$
Output: Index set of the placed active nodes $\mathcal{I}_A$
Build a directed graph $\mathcal{G}$ based on E, the left $n \times n$ submatrix of $\bar{E}$, in which each vertex stores the original associated node index
Compute U from $\bar{E}$ as per (3.57)
▷ *The for-loop below allocates active nodes to satisfy condition 2) of Theorem 3.7*
for each row i of U **do**
    if row i of U is not all-zero **then**
        Attach node index i to $\mathcal{I}_A$
        Delete vertex i and its attached edges from $\mathcal{G}$
    end if
end for
▷ *The while-loop below solves for the minimum FVS of the remaining graph*
while $\mathcal{G}$ is not acyclic **do**
    for each vertex i in $\mathcal{G}$ **do**
        $m_i \leftarrow \mathrm{per}(\Omega_{\mathcal{G}}) - \mathrm{per}(\Omega_{\mathcal{G}i})$
    end for
    $i_A \leftarrow \underset{i}{\operatorname{argmin}}\, m_i$
    Attach the original node index stored in vertex i_A to $\mathcal{I}_A$
    Delete vertex i_A and its attached edges from $\mathcal{G}$
end while

Proposition 3.1. *The while-loop in Algorithm 3.3 finds one*[8] *of the minimum FVSs of* $\mathcal{G}$.

Proof. First note two facts: 1) $\mathcal{G}$ contains a finite number of directed cycles and 2) at least one vertex in each directed cycles has to be included in the minimum FVS. If $\mathcal{G}$ is acyclic, the minimum FVS is the empty set, which is consistent with the algorithm as the while-loop breaks without attaching any indices. Otherwise $m_{i_A} \geq 1$, and vertex i_A is contained in at least one directed cycle. Due to fact 2), one has to include at least one vertex to the minimum FVS for each such cycle, and due to Lemma 3.6, vertex i_A is the vertex such that the deletion of itself and its attached edges removes the most directed cycles in $\mathcal{G}$. At each step of the while-loop, the selected vertex i_A is among the vertices, one of which has to be selected, and is the one that removes the most directed cycles from the finite number mentioned in fact 1). Hence the algorithm removes all directed cycles in $\mathcal{G}$ with minimum steps, which completes the proof. □

[8] The minimum FVS of a graph is, in general, not unique.

Algorithm 3.3 is at most of complexity $O\big(p(n)2^n\big)$ or equivalently $O^*(2^n)$, where $p(\cdot)$ is a polynomial. To see this note that the complexity is dominated by the while-loop solving for the minimum FVS and, in each single loop, the complexity of the algorithm is dominated by the computation of the permanent. Computing the permanent of an $n \times n$ matrix is of $O(n2^n)$ using *Ryser's formula* in the *Gray code* order [96, 110], and in the worst case the number of permanents that need to be computed is a polynomial of n, which means that the overall complexity is $O^*(2^n)$. The efficiency can be further improved if one computes the permanent using the method proposed by [114], which reduces the overall complexity to $O^*\big((2-\epsilon)^n\big)$, where $\epsilon > 0$ is a constant depending on the sparsity of the graph. This is comparable to the complexity of existing exact methods (*e.g.* [105]) for the solution of the minimum FVS problem of directed graphs.

In general the exact methods solve for the minimum FVS in directed graphs in exponential time, which is not acceptable when the considered network is large. In practice, the number of available *active nodes* can also be limited (this is also why one needs to find a placement method using the minimum number of *active nodes*), which restricts the associated FVS problem to a *parametrized* FVS problem, in which the parameter is the upperbound of the cardinality of the FVS. The algorithms proposed in [14, 83] can solve the *parametrized* FVS problem in linear time or report that no FVS smaller than the given bound exists. Furthermore, if we allow trading off the computation time against the number of *active nodes* used, the approximate method proposed in [55] can be used for computing permanent and this reduces the computation time to polynomial in a probabilistic sense. It should be noted that such a reduction in computation time is achieved at the cost of precision and the resulting set may not be a minimum FVS as Proposition 3.1 may not hold, but the result is guaranteed to be an FVS by definition since the while-loop in Algorithm 3.3 checks whether the graph is acyclic before its termination, and therefore it still computes a valid placement of *active nodes*.

3.4.2 Computation of Damping and Scaling Coefficients

Once the location of the *active nodes* has been determined, and the associated vertices comprise a FVS, one can solve for the damping coefficients of the *active nodes* and the corresponding scaling coefficients. In what follows we show that the solution for these coefficients can be formulated as the solution of a set of linear inequalities, despite the nonlinear coupling between the coefficients. We consider the linearly parametrized nonlinear supply rates discussed in Section 3.3. Recall that for the network system with node dissipation inequalities (3.52) and the network storage function (3.6), the network dissipation inequality is (3.55). As a result, the dissipation inequality (3.7) holds if $\hat{E}^\top \hat{c} > 0$, for some $\hat{c} > 0$ subject to (3.59). Note that $\hat{E}$ can be decomposed as $\hat{E} = M + \tilde{M}$, where

$$(M)_{ij} = \begin{cases} \underline{a}_i, & \text{if } i \in \hat{\mathcal{I}}_A,\ j = i, \\ (\hat{E})_{ij}, & \text{otherwise,} \end{cases} \tag{3.67}$$

and

$$(\tilde{M})_{ij} = \begin{cases} \tilde{a}_i, & \text{if } i \in \hat{\mathcal{I}}_A,\ j = i, \\ 0, & \text{otherwise,} \end{cases} \tag{3.68}$$

with $\tilde{a}_i \triangleq a_i - \underline{a}_i$. Note that in $\tilde{M}^\top \hat{c}$, for $i_A \in \hat{\mathcal{I}}_A$, the decision variables $\tilde{a}_{i_A}$ are multiplied by $\hat{c}_{i_A}$, which are also decision variables, creating nonlinear couplings. To linearize these nonlinear terms, define $d_j \triangleq \hat{c}_{i_{Aj}} \tilde{a}_{i_{Aj}}$, where $i_{Aj} \in \hat{\mathcal{I}}_A$ is the vertex index of the j-th (augmented) vertex that originates from the *active nodes*, for $j = 1, \ldots, |\hat{\mathcal{I}}_A|$, and $d \triangleq [d_1, \ldots, d_{|\hat{\mathcal{I}}_A|}]^\top$. Note that the damping coefficients $a_{i_{Aj}}$ can be obtained from d by using the relation

$$a_{i_{Aj}} = \frac{d_j}{c_{i_{Aj}}} + \underline{a}_{i_{Aj}}. \tag{3.69}$$

This allows defining a constant matrix

$$(N)_{ij} = \begin{cases} 1, & \text{if } i \in \hat{\mathcal{I}}_A,\ j = i, \\ 0, & \text{otherwise,} \end{cases} \tag{3.70}$$

such that $\tilde{M}^\top \hat{c} = N^\top d$ and therefore $\hat{E}^\top \hat{c} = M^\top \hat{c} + N^\top d$. This yields a system of linear inequalities and equations given by

$$\begin{aligned} M^\top \hat{c} + N^\top d &\geq \hat{\sigma}, \\ L\hat{c} &= 0, \\ \hat{c} &> 0, \\ d &> 0. \end{aligned} \tag{3.71}$$

This describes an admissible region that is non-empty due to Theorem 3.7 yet not a direct solution for $\hat{c}$ and d. In the light of this, one can re-write (3.71) into an optimization problem with some objective function f to be specified, namely

$$\begin{aligned} &\min_v f(v) \\ \text{subject to} \quad &Rv \geq \hat{\sigma}, \\ &Qv = 0, \\ &\underline{v} \leq v \leq \bar{v}, \end{aligned} \tag{3.72}$$

where $v \triangleq [\hat{c}^\top, d^\top]^\top$; $R \triangleq [M^\top, N^\top]$; $Q \triangleq [L, 0]$; $\hat{\sigma} > 0$ is the same as the one used in Theorem 3.7 that specifies a minimum guaranteed dissipation rate $-\hat{\sigma}^\top \hat{\phi}(y)$; $\underline{v} \triangleq [\varepsilon^\top, 0_{1\times|\hat{\mathcal{I}}_A|}]^\top$, with $\varepsilon > 0$ to keep $\hat{c}$ away from 0; and $\bar{v} > 0$ is an "upperbound" to make the admissible region compact and suitable for the use of off-the-shelf optimization solvers. Typically one may want to keep the damping coefficients of the *active nodes* as small as possible while fulfilling all the design specifications. In this case one may consider the objective function $f(v) = \sum_{j=1}^{|\hat{\mathcal{I}}_A|} \frac{d_j}{c_{i_{Aj}}} = \sum_{j=1}^{|\hat{\mathcal{I}}_A|} \tilde{a}_{i_{Aj}}$ or a weighted version of this. Since this cost is nonlinear and may complicate the computation, one can also consider the linear cost $f(v) = w^\top v$, where w is a vector of constant weights. One only needs to tune w in such a way that $w_{i_{Aj}} \ll w_{\hat{q}+j}$, which leads to a large $\hat{c}_{i_{Aj}}$ and a small d_j, yielding a small $a_{i_{Aj}}$. This reduces (3.72) to a linear programming problem and allows searching for v efficiently, which is especially favourable for large-scale network systems. It should be noted that the selection of the objective function only affects the solution for $\hat{c}$ and a_{i_A}, $i_A \in \hat{\mathcal{I}}_A$, whereas all solutions yield the same network dissipation inequality $\dot{V} \leq -\hat{\sigma}^\top \hat{\phi}(y)$.

We complete our discussion by showing that Theorem 3.7 also guarantees that the admissible region of the reformulated optimisation problem (3.72) is non-empty.

Proposition 3.2. *Consider the problem* (3.72) *and assume that the conditions of Proposition 3.7 are satisfied. Then, for all $\varepsilon > 0$, there exists a $\bar{v}$, depending on ε, such that the admissible region of* (3.72) *is non-empty.*

Proof. By Theorem 3.7, setting a constant $\hat{\sigma} > 0$, there exist $\hat{c} > 0$ and $a_{i_A} > \underline{a}_{i_A}$, $i_A \in \hat{\mathcal{I}}_A$ such that $\hat{E}^\top \hat{c} = \hat{\sigma}$ and $L\hat{c} = 0$. Consider now a scalar λ and a new scaling vector $\hat{c}' \triangleq \lambda \hat{c}$. Keeping a_{i_A} unchanged, $i_A \in \hat{\mathcal{I}}_A$, we have $d' = \mathrm{diag}(\tilde{a}_{i_A})\hat{c}' = \lambda \mathrm{diag}(\tilde{a}_{i_A})\hat{c} = \lambda d$, which yields $v' = \lambda v$. Recall that $\underline{v} = [\varepsilon^\top, \mathit{0}]^\top$. For all $\varepsilon > 0$, there exists $\lambda \geq 1$, depending on ε, such that $v' \geq \underline{v}$. Then it is straightforward to construct $\bar{v} \geq v'$ such that $\underline{v} \leq v' \leq \bar{v}$. Finally note that v' also satisfies the first two conditions of (3.72) since $\lambda \geq 1$, and hence the admissible region contains at least v'. This completes the proof. $\square$

3.4.3 Adaptation of Damping Coefficients

The result presented in Section 3.4.2 provides a solution for both the scaling coefficients for all node systems and the damping coefficients for the *active nodes*. From a control synthesis perspective, the computation of the scaling coefficients is unnecessary as they do not appear in the implementation, and the existence of such scaling coefficients are sufficient to establish desired stability properties. Lifting the requirement of computing the scaling coefficients allows obtaining damping coefficients for the *active nodes* via adaptation instead of explicit computation. In what follows we present how to implement decentralized identifiers to achieve this.

For an *active node* Σ_{i_A}, $i_A \in \mathcal{I}_A$, one can substitute the parameter estimates $\hat{a}^k_{i_A}$, $k = 1, \ldots, q_{i_A}$ for the damping coefficients $a^k_{i_A}$. Define $\theta_{i_A} = [a^1_{i_A}, \ldots, a^{q_{i_A}}_{i_A}]^\top$, $\hat{\theta}_{i_A} = [\hat{a}^1_{i_A}, \ldots, \hat{a}^{q_{i_A}}_{i_A}]^\top$, $\tilde{\theta}_{i_A} = \hat{\theta}_{i_A} - \theta_{i_A}$, $\bar{\varphi}_{i_A}(y_{i_A}) = [\bar{\varphi}^1_{i_A}(y_{i_A}), \ldots, \bar{\varphi}^{q_{i_A}}_{i_A}(y_{i_A})]^\top$ and consider the parameter update law

$$\dot{\hat{\theta}}_{i_A} = \mathrm{proj}\left(\theta_{i_A}, \Gamma_{i_A}\bar{\varphi}_{i_A}(y_{i_A})\right), \tag{3.73}$$

where $\Gamma_{i_A} = \Gamma^\top_{i_A} \succ 0$ is the adaptation gain, and $\mathrm{proj}(\theta, v)$ is the *projection operator* proposed in [102], which projects the vector v onto the tangent space (at $\hat{\theta}$) of a convex region Θ (with smooth boundary) to which θ belongs, if $\hat{\theta}$ reaches the boundary of Θ and

v is pointing out of Θ. Therefore the condition

$$\tilde{\theta}_{i_A} \operatorname{proj}\left(\theta_{i_A}, \Gamma_{i_A}\bar{\varphi}_{i_A}(y_{i_A})\right) \leq \tilde{\theta}_{i_A}\Gamma_{i_A}\bar{\varphi}_{i_A}(y_{i_A}) \tag{3.74}$$

holds. By Theorem 3.7, the existence of θ is guaranteed and one can always select Θ such that it is contained in the cone defined by Definition 3.5 and sufficiently large to contain the true parameter θ. It turns out that one only needs to augment the node storage functions with a positive definite function of the parameter estimation error $\tilde{\theta}$ to obtain a result similar to Theorem 3.7 without the exact knowledge of θ nor the computation of it.

Proposition 3.3. *Consider the augmented node storage functions* $V_{i_A}^{\theta} = V_{i_A} + \frac{1}{2}\tilde{\theta}_{i_A}^{\top}\Gamma_{i_A}^{-1}\tilde{\theta}_{i_A}$ *for all active nodes* Σ_{i_A}*,* $i_A \in \mathcal{I}_A$*, with* $a_{i_A}^k$*,* $k = 1, \ldots, q_{i_A}$*, replaced by the parameter estimates* $\hat{\theta}_{i_A}$ *(updated by* (3.73)*), and the original node storage functions with dissipation inequalities* (3.51) *for the other nodes. If the two conditions on the active nodes in Theorem 3.7 hold then, for all* $\hat{\sigma} > 0$*, there exists* $c > 0$ *and convex regions of projection* Θ_{i_A}*, depending on* $\hat{\sigma}$*, such that the network dissipation inequality*

$$\dot{V}^{\theta} \leq -\hat{\sigma}^{\top}\hat{\phi}(y) \tag{3.75}$$

holds, where the augmented network storage function is given by

$$V^{\theta} = \sum_{i \notin \mathcal{I}_A} c_{i_A} V_i + \sum_{i_A \in \mathcal{I}_A} c_{i_A} V_{i_A}^{\theta}. \tag{3.76}$$

Proof. First consider the time derivative of $V_{i_A}^{\theta}$, namely

$$\begin{aligned}
\dot{V}_{i_A}^{\theta} &= \dot{V}_{i_A} + \overbrace{\frac{1}{2}\tilde{\theta}_{i_A}^{\top}\Gamma_{i_A}^{-1}\theta}^{\cdot} \\
&\leq -\sum_{k=1}^{q_{i_A}} \hat{a}_{i_A}^k \varphi_{i_A}^k(y_{i_A}) + \sum_{j=1, j\neq i_A}^{n} \left(\sum_{k=1}^{q_j} b_{i_A j}^k \varphi_j^k(y_j)\right) \\
&\quad + \tilde{\theta}_{i_A}^{\top}\Gamma_{i_A}^{-1} \operatorname{proj}\left(\theta_{i_A}, \Gamma_{i_A}\bar{\varphi}_{i_A}(y_{i_A})\right).
\end{aligned} \tag{3.77}$$

By inequality (3.74) we obtain

$$
\begin{aligned}
\dot{V}_{i_A}^{\theta} \leq & -\sum_{k=1}^{q_{i_A}} a_{i_A}^k \varphi_{i_A}^k(y_{i_A}) + \sum_{j=1, j\neq i_A}^{n} \left(\sum_{k=1}^{q_j} b_{i_A j}^k \varphi_j^k(y_j) \right) \\
& - \sum_{k=1}^{q_{i_A}} (\hat{a}_{i_A}^k - a_{i_A}^k) \varphi_{i_A}^k(y_{i_A}) + \tilde{\theta}_{i_A}^\top \bar{\varphi}_{i_A}(y_{i_A}) \\
= & -\sum_{k=1}^{q_{i_A}} a_{i_A}^k \varphi_{i_A}^k(y_{i_A}) + \sum_{j=1, j\neq i_A}^{n} \left(\sum_{k=1}^{q_j} b_{i_A j}^k \varphi_j^k(y_j) \right). \qquad (3.78)
\end{aligned}
$$

Similarly to what is done in passivity-based adaptive control, the parameter estimation error terms are cancelled by the parameter update law (3.73). This means that the node dissipation inequalities associated with the augmented storage function in the adaptive case have the same supply rates as that of the original storage functions (3.51) in the case in which the $a_{i_A}^k$ are known. This allows using the unknown coefficients $a_{i_A}^k$ as if they were known (the *certainty-equivalence principle*) in the subsequent analysis as long as their nominal value is contained in the convex regions Θ_{i_A}. This reduces the problem to the same problem solved by Theorem 3.7. Then the rest of the proof is straightforward by invoking Theorem 3.7, and by noting that the existence of Θ_{i_A} which satisfies the constraints in Definition 3.5 and contains the nominal value of $a_{i_A}^k$, is guaranteed by the existence of $a_{i_A}^k$. This completes the proof. □

Remark 3.6. *The regions of projection Θ_{i_A} has to be determined when implementing the parameter update law* (3.73)*: this requires some knowledge of $a_{i_A}^k$. This is not as restrictive as the requirement to know the exact value of $a_{i_A}^k$ in two senses. First, one can make Θ_{i_A} sufficiently large to relax the requirement of such knowledge. Second, it is possible to implement the update law without projection, for example,*

$$
\dot{\hat{a}}_{i_A}^k = \gamma_{i_A}^k \varphi_{i_A}^k(y_{i_A}), \qquad (3.79)
$$

with adaptation gain $\gamma_{i_A}^k > 0$ and initial condition $\hat{a}_{i_A}^k(0) \geq \underline{a}_{i_A}^k$. In this case the parameter estimates are decoupled and each of them is non-decreasing over time. Due to the initial conditions we set, these parameter estimates satisfy the constraints of Definition 3.5. This is a special case of (3.73) *in which Γ is diagonal and one can prove the counterpart of Proposition 3.3 in a similar approach. It should, however, be noted that the implementation* *without projection is not robust in the presence of disturbances.*

In addition to removing the need for explicit computation, the adaptive update laws used, either (3.73) or (3.79), only require information on the node Σ_{i_A}, which allows the implementation to be decentralized. This is more favourable in some scenarios in which the information of other nodes is unavailable, compared to the method in Section 3.4.2.

3.5 A Control Synthesis Example

In this section let us see an example that demonstrates the advantage of *active nodes* in stabilizing an interconnected system without explicit computation. Consider the three-node interconnected nonlinear system described by

$$\begin{aligned}
\Sigma_1 &: \ \dot{y}_1 = y_3 + u, \\
\Sigma_2 &: \ \dot{y}_2 = y_1 - y_2 + y_1^2, \\
\Sigma_3 &: \ \dot{y}_3 = y_2 - y_3,
\end{aligned} \tag{3.80}$$

and a controller given by

$$u = -k_1 y_1 - k_2 y_1^3, \tag{3.81}$$

where $k_1 > 0$ and $k_2 > 0$ are adjustable parameters. The design objective is to regulate y_1, y_2 and y_3 to 0. In the spirit of this chapter, the first step is to derive the dissipation inequality for each node system. Define the standard quadratic storage functions $V_i = \frac{1}{2}y_i^2$, $i = 1, 2, 3$, for each node system. Taking the time derivatives of these storage functions along the trajectories of the closed-loop system and invoking Lemma A.1 (Young's inequality) yields the dissipation inequalities

$$\begin{aligned}
\dot{V}_1 &\leq -\left(k_1 - \frac{1}{2}\right) y_1^2 + \frac{1}{2} y_3^2 - k_2 y_1^4, \\
\dot{V}_2 &\leq y_1^2 - \frac{1}{2} y_2^2 + y_1^4, \\
\dot{V}_3 &\leq \frac{1}{2} y_2^2 - \frac{1}{2} y_3^2,
\end{aligned} \tag{3.82}$$

which can be written in to the compact form $\dot{V} \leq \bar{E}\hat{\phi}(y)$, where $\hat{\phi}(y) \triangleq [y_1^2, y_2^2, y_3^2, y_1^4]^\top$ and

$$\bar{E} \triangleq \begin{bmatrix} k_1 - \frac{1}{2} & 0 & -\frac{1}{2} & -k_2 \\ -1 & \frac{1}{2} & 0 & -1 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}. \tag{3.83}$$

Using (3.56), one can augment (3.83) into the matrix

$$\hat{E} \triangleq \begin{bmatrix} k_1 - \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} & 0 & -1 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -k_2 \end{bmatrix}. \tag{3.84}$$

The underlying *augmented* directed graph of (3.84) is depicted in Fig. 3.3. Since k_1 and

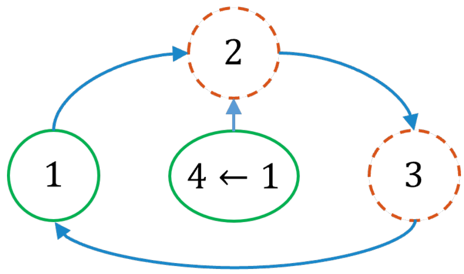

Figure 3.3: The underlying directed graph specified by (3.84)**.**

k_2 are adjustable design parameters, Σ_1 is an *active node* according to Definition 3.5. Also observe that vertex 1 is contained in the only directed cycle of the graph, and therefore the conditions in Theorem 3.7 are satisfied. By Theorem 3.7 one can conclude that for all positive constants σ_1, σ_2, σ_3, σ_4, there exist some $k_1 > 0$, $k_2 > 0$, and positive constants c_1, c_2, c_3, depending on $\sigma_{(\cdot)}$, such that the overall storage function $V \triangleq c_1V_1 + c_2V_2 + c_3V_3$ satisfies the dissipation inequality

$$\dot{V} \leq -\sigma_1 y_1^2 - \sigma_2 y_2^2 - \sigma_3 y_3^2 - \sigma_4 y_1^4 \leq 0. \tag{3.85}$$

Instead of computing k_1 and k_2 explicitly, a shortcut is to use the adaptive controller given by

$$\begin{aligned}\dot{\hat{k}}_1 &= \gamma_1 y_1^2, \quad \hat{k}_1(0) \geq 0, \\ \dot{\hat{k}}_2 &= \gamma_2 y_1^4, \quad \hat{k}_2(0) \geq 0, \\ u &= -\hat{k}_1 y_1 - \hat{k}_2 y_1^3, \end{aligned} \tag{3.86}$$

where $\gamma_1 > 0$ and $\gamma_2 > 0$ are the adaptation gains. Note that by Proposition 3.3, replacing controller (3.81) with controller (3.86) and re-defining $V_1 = \frac{1}{2}y_1^2 + \frac{1}{2\gamma_1}(k_1 - \hat{k}_1)^2 + \frac{1}{2\gamma_2}(k_2 - \hat{k}_2)^2$ does not change the overall dissipation inequality (3.85). Using standard boundedness analysis and Lemma A.5 we can conclude that all closed-loop signals are bounded and y_1, y_2, y_3 converge to 0 asymptotically. This is verified by the simulation results shown in Fig. 3.4.

Note that by using the adaptive controller (3.86) we augment the state of Σ_1 with $\hat{k}_1$ and $\hat{k}_2$, and this subsystem is not ISS. The proposed method, unlike the classical small-gain based on ISS, is still valid for this system as it only exploits the algebraic properties of the dissipation inequalities, which is referred to as the small-gain-like analysis/synthesis in the thesis. From this example we can see that by checking the locations of *active nodes* in the graph one can easily check the feasibility of the controller design problem. Even the synthesis of the controller can be completed without explicit computation if adaptive control techniques are applied.

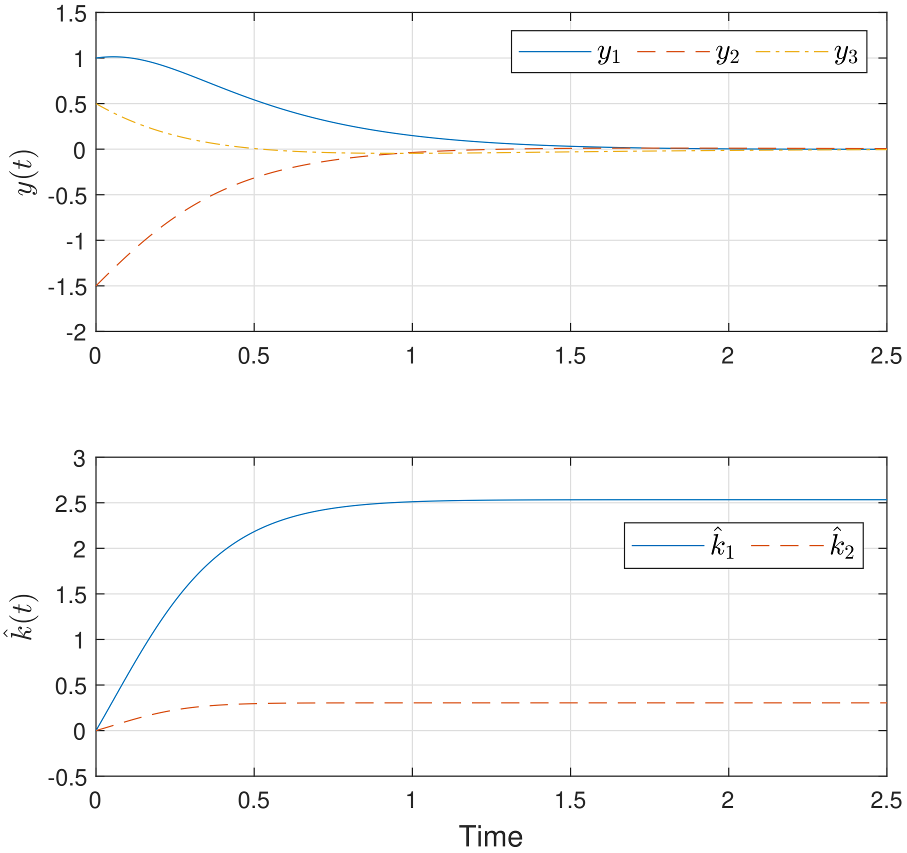

Figure 3.4: Time histories of the closed-loop state variables from the initial conditions $y_1(0) = 1$, $y_2(0) = -1.5$, $y_3(0) = 0.5$, $\hat{k}_1(0) = \hat{k}_2(0) = 0$, and with the gains $\gamma_1 = 6$, $\gamma_2 = 1$.

Chapter 4

State-Feedback Design for Lower-triangular Systems

In this chapter we move back to the discussion of the *congelation of variables* method. As seen in Chapter 2, the *congelation of variables* method removes the undesired disturbance caused by the $\dot{\theta}$ term via introducing the *congealed* parameter ℓ_θ, and either maintains the passive interconnection or achieves a small-gain-like condition, via strengthened damping design. This chapter shows that the same perspective is still applicable to a more general class of nonlinear systems, namely, systems described by the equations

$$\begin{aligned}
\dot{x}_1 &= \phi_1^\top(x_1)\theta(t) + x_2, \\
&\vdots \\
\dot{x}_i &= \phi_i^\top(\underline{x_i})\theta(t) + x_{i+1}, \\
&\vdots \\
\dot{x}_n &= \phi_n^\top(x)\theta(t) + u,
\end{aligned} \tag{4.1}$$

where $i = 2, \ldots, n-1$; $x(t) = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ is the state; $u(t) \in \mathbb{R}$ is the input; $\theta(t) \in \mathbb{R}^q$ is the vector of unknown parameters satisfying Assumption 1.1; and $\phi_i : \mathbb{R}^i \to \mathbb{R}^q$, $i = 1, \ldots, n$ are the regressors satisfying Assumption 1.4. This class of systems are commonly said to be in *lower-triangular form*, due to the fact that the nonlinearities ϕ_i's depend only on $\underline{x_i}$.

Remark 4.1. *The condition $\phi_i(0) = 0$ implies that $\phi_i^\top(0)\theta(t) = 0$, for all $t \geq 0$, which allows zero control effort at $x = 0$. One can easily see that if $\phi_i(0) \neq 0$, $\phi_i^\top(0)\theta(t)$ becomes an unknown time-varying disturbance (similar to the one in Remark 2.2), yielding a disturbance rejection/attenuation problem not discussed here. By Lemma A.3 (Hadamard's lemma), one can express the regressors as $\phi_i(\underline{x_i}) = \bar{\Phi}_i(\underline{x_i})\underline{x_i}$, where the $\bar{\Phi}_i$'s are smooth mappings.*

The challenge brought by the more complex structure of the system is that the parametric uncertainty term $\phi_i^\top(\underline{x_i})\theta$ enters the system via integrators which are different from the integrator via which the control input u enters. This requires modifying the control law as the designs in Chapter 2 are only suitable for nonlinear systems in which the uncertainty enters the system via the same integrator as the control input u does, that is, for systems satisfying the so-called *matching condition*[1]. This is because in such a case one can use the *certainty-equivalence principle*, that is, adding a term $-\phi^\top(x)\hat{\theta}$ directly to the control law to compensate for the effect of the uncertainty $\phi^\top(x)\theta$ since it enters the same integrator, which yields a linearly parametrized error term $\phi^\top(x)(\theta - \hat{\theta})$ that can be dealt with by adaptively updating $\hat{\theta}$, as seen in Chapter 2.

It is well known that to solve the adaptive control problem when the *matching condition* fails, one needs to find a change of coordinates and a feedback control such that the closed-loop system, described in the new coordinates, contains a linearly parametrized error term, similar to the *matched* case, which can be dealt with using the three classes of identifiers discussed in Chapter 2. The spirit of such a change of coordinates has been exploited extensively, explicitly or implicitly, in nonlinear control problems. In this chapter we adopt the formulation and interpretation used in [75] for the aforementioned change of coordinates, that is, the so-called *integrator backstepping* scheme, or *backstepping* for short, due to its popularity in the adaptive control literature. The core spirit of the *backstepping* method is to treat the term x_{i+1} as a virtual control input to the dynamics of x_i, and to design a stabilizing parametrized virtual control law $\alpha_i(\underline{x_i}, \hat{\theta})$ for x_{i+1} to "track". The "tracking" error $z_{i+1} \triangleq x_{i+1} - \alpha_i$ is the resulting coordinate. By controlling z_{i+1} using x_{i+1}, one obtains z_{i+2}. Repeating the same procedure in a recursive manner, with ascending subscript of x, effectively yields the change of coordinates from x to z

[1] In model reference adaptive control literature (see *e.g.* [126]) the same terminology is used, with a different meaning, for model-plant matching.

step by step in the reverse order of the connection of the integrators, and this is why this procedures is called *integrator backstepping.* Since the uncertainty term is directly compensated by the virtual control law at the same integrator, the parameter update law can be designed by similar approaches as in Chapter 2. Meanwhile, the perturbation terms caused by the *congelation of variables* appear in the dynamics of every integrator and has to be dominated step by step consistent with the order of the recursive design of the virtual control laws: this is important to establish boundedness and convergence properties. In the rest of this chapter, we see how to combine the *congelation of variables* method with *backstepping* when using different classes of identifiers.

4.1 Passivity-Based Scheme with Backstepping

Consider system (4.1). The forward step-by-step derivation of the *backstepping* variables for this system is presented in [75, Chapter 4]. For the sake of conciseness and clarity, the *backstepping* method is presented the other way round, that is, the definition of the relevant variables is given before we proceed to explain the reasons for such definitions. Thus, for each step i, $i = 1, \dots, n$, define:

• the error variables

$$z_0 = 0, \tag{4.2}$$

$$z_i = x_i - \alpha_{i-1}, \tag{4.3}$$

• the new regressor vectors

$$w_i(\underline{x_i}, \hat{\theta}) = \phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j, \tag{4.4}$$

• the tuning functions

$$\tau_0 = 0, \tag{4.5}$$

$$\tau_i(\underline{x_i}, \hat{\theta}) = \tau_{i-1} + w_i z_i = \sum_{j=1}^{i} w_i z_i, \tag{4.6}$$

- the virtual control laws

$$
\begin{aligned}
\alpha_0 &= 0, && (4.7)\\
\alpha_i(\underline{x_i}, \hat{\theta}) &= -z_{i-1} - (c_i + \zeta_i)z_i - w_i^\top \hat{\theta} \\
&\quad + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} x_{j+1} + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma_\theta \tau_i + \sum_{j=2}^{i-1} \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}} \Gamma_\theta w_i z_j, && (4.8)
\end{aligned}
$$

where $c_i > 0$ is the constant feedback gain; $\zeta_i(\underline{x_i}, \hat{\theta})$ is the nonlinear feedback gain to be defined; and $\Gamma_\theta = \Gamma_\theta^\top \succ 0$ is the adaptation gain.

Remark 4.2. *By constructing* (4.3) *to* (4.8) *recursively, it is not hard to see that* $z_i(\underline{x_i}, \hat{\theta})$, $w_i(\underline{x_i}, \hat{\theta})$, $\tau_i(\underline{x_i}, \hat{\theta})$, $\alpha_i(\underline{x_i}, \hat{\theta})$ *are smooth and* $z_i(0, \hat{\theta}) = 0$, $w_i(0, \hat{\theta}) = 0$, $\tau_i(0, \hat{\theta}) = 0$, $\alpha_i(0, \hat{\theta}) = 0$. *Note also that the* $\hat{\theta}$*-dependent change of coordinates between* $\underline{z_i}$ *and* $\underline{x_i}$ *is smooth, invertible, and* $\underline{x_i} = 0 \Leftrightarrow \underline{z_i} = 0$, *thus we can directly express* w_i *as* $w_i = \bar{W}_i(\underline{x_i}, \hat{\theta})\underline{z_i}$ *with* W_i *smooth, by Lemma A.3.*

We are now ready to perform a step-by-step computation to explain the roles of these variables in the context of the *congelation of variables.*

Step 1. Computing the z_1-dynamics yields

$$
\dot{z}_1 = \dot{x}_1 = z_2 + \alpha_1 + w_1^\top \theta. \tag{4.9}
$$

Consider the function $V_1 = \frac{1}{2} z_1^2$, which yields

$$
\begin{aligned}
\dot{V}_1 &= z_1(z_2 + \alpha_1 + w_1^\top \theta) \\
&= -(c_1 + \zeta_1)z_1^2 + z_1 z_2 + z_1 w_1^\top (\theta - \hat{\theta}) \\
&= -(c_1 + \zeta_1)z_1^2 + z_1 z_2 + z_1 w_1^\top \Delta_\theta + (\ell_\theta - \hat{\theta})^\top w_1 z_1. && (4.10)
\end{aligned}
$$

The so-called tuning function τ_1 can be treated as a temporary candidate for the parameter update law since considering $\bar{V}_1 = V_1 + \frac{1}{2}|\ell_\theta - \hat{\theta}|^2_{\Gamma_\theta^{-1}}$ and $\dot{\hat{\theta}} = \Gamma_\theta \tau_1$, one obtains

$$
\begin{aligned}
\dot{\bar{V}}_1 &= -(c_1 + \zeta_1)z_1^2 + z_1 z_2 + z_1 w_1^\top \Delta_\theta + (\ell_\theta - \hat{\theta})^\top (w_1 z_1 - \tau_1) \\
&= -(c_1 + \zeta_1)z_1^2 + z_1 z_2 + z_1 w_1^\top \Delta_\theta, && (4.11)
\end{aligned}
$$

in which the $(\ell_\theta - \hat{\theta})$ term is cancelled.

Step 2. The dynamics of z_2 are described by

$$\dot{z}_2 = \dot{x}_2 - \dot{\alpha}_1 = z_3 + \alpha_2 + w_2^\top \theta - \frac{\partial \alpha_1}{\partial x_1} x_2 - \frac{\partial \alpha_1}{\partial \hat{\theta}} \dot{\hat{\theta}}. \tag{4.12}$$

Consider $V_2 = V_1 + \frac{1}{2} z_2^2$. This yields

$$\begin{aligned}
\dot{V}_2 = \; & \dot{V}_1 - z_1 z_2 - (c_2 + \zeta_2) z_2^2 + z_2 z_3 + z_2 w_2^\top \Delta_\theta \\
& + z_2 \frac{\partial \alpha_1}{\partial \hat{\theta}} (\Gamma_\theta \tau_2 - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top w_2 z_2 \\
= \; & -(c_1 + \zeta_1) z_1^2 - (c_2 + \zeta_2) z_2^2 + z_2 z_3 + (z_1 w_1^\top + z_2 w_2^\top) \Delta_\theta \\
& + z_2 \frac{\partial \alpha_1}{\partial \hat{\theta}} (\Gamma_\theta \tau_2 - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top \tau_2.
\end{aligned} \tag{4.13}$$

Step i, $i = 3, \ldots, n-1$. The z_i-dynamics can be written as

$$\dot{z}_i = \dot{x}_i - \dot{\alpha}_{i-1} = z_{i+1} + \alpha_i + w_i^\top \theta - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} x_{j+1} - \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \dot{\hat{\theta}}. \tag{4.14}$$

Continue to consider $V_i = V_{i-1} + \frac{1}{2} z_i^2$. This yields

$$\begin{aligned}
\dot{V}_i = \; & \dot{V}_{i-1} - z_{i-1} z_i - (c_i + \zeta_i) z_i^2 + z_i z_{i+1} + z_i w_i^\top \Delta_\theta \\
& + z_i \sum_{j=2}^{i-1} \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}} \Gamma_\theta w_i z_j + z_i \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} (\Gamma_\theta \tau_i - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top w_i z_i \\
= \; & -\sum_{j=1}^{i} (c_j + \zeta_j) z_j^2 + z_i z_{i+1} + \sum_{j=1}^{i} z_j w_j^\top \Delta_\theta \\
& + \left(\sum_{j=1}^{i-1} z_{j+1} \frac{\partial \alpha_j}{\partial \hat{\theta}} \right) (\Gamma_\theta \tau_i - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top \tau_i.
\end{aligned} \tag{4.15}$$

Finally, select the actual control law and the parameter update law.

$$u = \alpha_n(x, \hat{\theta}), \tag{4.16}$$

$$\dot{\hat{\theta}} = \Gamma_\theta \tau_n, \tag{4.17}$$

which allows the analysis of the last step.

Step n. Differentiating z_n yields

$$\dot{z}_n = \dot{x}_n - \dot{\alpha}_{n-1} = u + w_n^\top \theta - \sum_{j=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial x_j} x_{j+1} - \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}} \dot{\hat{\theta}}. \tag{4.18}$$

Consider $V_n = V_{n-1} + \frac{1}{2} z_n^2 = \frac{1}{2}|z|^2$, where $z \triangleq [z_1, \ldots, z_n]^\top$. This yields

$$\begin{aligned}
\dot{V}_n &= \dot{V}_{n-1} - z_{n-1} z_n - (c_n + \zeta_n) z_n^2 + z_i w_i^\top \Delta_\theta \\
&\quad + z_n \sum_{j=2}^{n-1} \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}} \Gamma_\theta w_n z_j + z_n \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}} (\Gamma_\theta \tau_n - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top w_n z_n \\
&= -\sum_{j=1}^{n} (c_j + \zeta_j) z_j^2 + \sum_{j=1}^{n} z_j w_j^\top \Delta_\theta \\
&\quad + \left(\sum_{j=1}^{n-1} z_{j+1} \frac{\partial \alpha_j}{\partial \hat{\theta}} \right) (\Gamma_\theta \tau_n - \dot{\hat{\theta}}) + (\ell_\theta - \hat{\theta})^\top \tau_n.
\end{aligned} \tag{4.19}$$

Note that the second last term is 0 due to the parameter update law (4.17), and thus

$$\dot{V}_n = -\sum_{j=1}^{n} (c_j + \zeta_j) z_j^2 + \sum_{j=1}^{n} z_j w_j^\top \Delta_\theta + (\ell_\theta - \hat{\theta})^\top \tau_n. \tag{4.20}$$

Then, considering the function $\bar{V}_n = V_n + \frac{1}{2}|\ell_\theta - \hat{\theta}|^2_{\Gamma_\theta^{-1}}$ and its time derivative along the closed-loop system trajectories yields

$$\begin{aligned}
\dot{\bar{V}}_n &= -\sum_{j=1}^{n} (c_j + \zeta_j) z_j^2 + \sum_{j=1}^{n} z_j w_j^\top \Delta_\theta + (\ell_\theta - \hat{\theta})^\top (\tau_n - \Gamma_\theta^{-1} \dot{\hat{\theta}}) \\
&= -\sum_{j=1}^{n} (c_j + \zeta_j) z_j^2 + \sum_{j=1}^{n} z_j w_j^\top \Delta_\theta.
\end{aligned} \tag{4.21}$$

In the classical case, in which $\zeta_i = 0$, $i = 1, \ldots, n$, and $\Delta_\theta(t) = 0$, for $t \geq 0$, we can conclude that $\dot{\bar{V}}_n$ is negative semi-definite and one can invoke Lemma A.5 for the boundedness and convergence analysis. In the presence of time-varying parameters, one additional step is needed to establish the negative semi-definiteness of $\dot{\bar{V}}_n$ via a strengthened damping design of the ζ_i terms, which is the main difference of the proposed (virtual) control laws compared to the classical (virtual) control laws. This is addressed in the following result.

Proposition 4.1. *Consider system* (4.1)*, the parameter update laws* (4.27)*, and the con-*

trol law (4.16) *with the nonlinear damping gains*

$$\zeta_i(\underline{x_i}, \hat{\theta}) = \frac{1}{2}\left((n-i+1)\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_i|_F^2\right), \tag{4.22}$$

with $\epsilon_{\Delta_\theta} > 0$. *Then, all trajectories of the closed-loop system are bounded and* $\lim\limits_{t\to+\infty} x(t) = 0$.

Proof. By Remark 4.2 and invoking Lemma A.1 yields

$$\begin{aligned} z_j w_j^\top \Delta_\theta &= z_j \Delta_\theta^\top \bar{W}_j \underline{z_j} \\ &\le \frac{1}{2}\left(\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_j|_{\mathrm{F}}^2\right) z_j^2 + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}|\underline{z_{j-1}}|^2. \end{aligned} \tag{4.23}$$

Recall that $V_n = \frac{1}{2}|z|^2$, $\bar{V}_n = V_n + \frac{1}{2}|\ell_\theta - \hat{\theta}|^2_{\Gamma_\theta^{-1}}$ and consider (4.21) and (4.22) together with the inequality (4.23). This yields

$$\begin{aligned} \dot{\bar{V}}_n =& -\sum_{j=1}^n (c_j + \zeta_j) z_j^2 + \sum_{j=1}^n z_j w_j^\top \Delta_\theta \\ \le& -\sum_{j=1}^n \left(c_j + \frac{1}{2}\left((n-j+1)\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_j|_{\mathrm{F}}^2\right)\right) z_j^2 \\ & + \sum_{j=1}^n \left(\frac{1}{2}\left(\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_j|_{\mathrm{F}}^2\right) z_j^2 + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}|\underline{z_{j-1}}|^2\right) \\ =& -\sum_{j=1}^n c_j z_j^2 \le 0. \end{aligned} \tag{4.24}$$

Boundedness. Since the function $\bar{V}_n(z, \hat{\theta})$ is positive definite and radially unbounded in $(z, \ell_\theta - \hat{\theta})$, equation (4.24) shows that both z and $\hat{\theta}$ are bounded. Due to Remark 4.2, x is bounded, and so are the variables defined from (4.2) to (4.8).

Convergence. Since $\dot{z}$ can be expressed using smooth functions of $\theta(t)$ and the variables defined by (4.2) to (4.8), which are all bounded, $\dot{z}$ is also bounded. Hence by invoking Lemma A.5 one can conclude that $\lim\limits_{t\to+\infty} z(t) = 0$, which further indicates that $\lim\limits_{t\to+\infty} x(t) = 0$, by Remark 4.2. □

Remark 4.3. *From the proof of Proposition 4.1 one can see that the function of adding nonlinear damping terms* ζ_i *to the control law is to counteract the effect of the perturbation*

$z^\top W^\top \Delta_\theta$, where $W \triangleq [w_1, \dots, w_n]$. This makes the z-subsystem, with the storage function V_n, passive from the input $\ell_\theta - \hat{\theta}$ to the output Wz, in the presence of the perturbation term $z^\top W^\top \Delta_\theta$ resulting from the congelation of variables. This allows the z-subsystem, together with the passive identifier, to compose a negative feedback interconnection of passive systems. Therefore, from a passivity perspective and despite the additional complexity, the design for lower-triangular systems is essentially in the same spirit as the design for scalar systems discussed in Section 2.1.

Time-varying Input Coefficient

When using the passivity-based scheme, it is possible to modify the aforementioned *backstepping* scheme with modifications similar to the ones used in Section 2.1.2. To illustrate this, consider the *lower-triangular* nonlinear system

$$
\begin{aligned}
\dot{x}_1 &= \phi_1^\top(x_1)\theta(t) + x_2, \\
&\vdots \\
\dot{x}_i &= \phi_i^\top(\underline{x_i})\theta(t) + x_{i+1}, \\
&\vdots \\
\dot{x}_n &= \phi_n^\top(x)\theta(t) + b(t)u.
\end{aligned}
\tag{4.25}
$$

The only difference between (4.25) and (4.1) is that there is a time-varying input coefficient $b(t) \in \mathbb{R}$, which is unknown and satisfies Assumption 1.1 and Assumption 1.3. The way to deal with $b(t)$ is to select a control law with a nonlinear feedback gain $\kappa(x, \hat{\theta}) > 0$, like what is done for the scalar systems. The major difference in this case is that the gain $\kappa(x, \hat{\theta})$ has to be multiplied by z_n to yield the $-\kappa z_n^2$ term. Therefore the design of the gain should be done in the last step of the *backstepping* procedures[2]. Thus, select the adaptive

[2] [78] has proposed a final-step control law with similar properties and analysis, which also uses a control law with a nonlinear negative feedback gain, albeit to achieve inverse optimality. This result inspired the proposed design to counteract the time-varying $b(t)$.

controller described by

$$u = -\hat{\varrho}\kappa(x, \hat{\theta})z_n, \tag{4.26}$$

$$\dot{\hat{\theta}} = \Gamma_\theta \tau_n, \tag{4.27}$$

$$\dot{\hat{\varrho}} = \gamma_\varrho \mathrm{sgn}(\ell_b)\kappa(x, \hat{\theta})z_n^2, \tag{4.28}$$

where z_n and τ_n are given by the recursive definitions (4.2); and $\kappa(x, \hat{\theta}) > 0$ is the nonlinear feedback gain to be determined. Since the first $n-1$ steps are identical to the case in which $b(t) = 1$, for all $t \geq 0$, consider the nth step directly. Let $V(z, \hat{\theta}, \hat{\varrho}) = \frac{1}{2}|z|^2 + \frac{1}{2}|\ell_\theta - \hat{\theta}|^2_{\Gamma^{-1}} + \frac{|\ell_b|}{2\gamma_\varrho}|\frac{1}{\ell_b} - \hat{\varrho}|^2$. Taking the time-derivative of V along the trajectories of the closed-loop system yields

$$\begin{aligned} \dot{V} = & -\sum_{i=1}^{n}(c_i + \zeta_i)z_i - \kappa z_n^2 + \Delta + z_n\psi \\ & + (\ell_\theta - \hat{\theta})^\top \left(\sum_{i=1}^{n-1} w_i z_i - \Gamma_\theta^{-1}\dot{\hat{\theta}} \right) + \ell_b \left(\frac{1}{\ell_b} - \hat{\varrho} \right) \left(\bar{\alpha}_n z_n - \frac{\dot{\hat{\varrho}}}{\gamma_\varrho} \right), \end{aligned} \tag{4.29}$$

where

$$\Delta \triangleq \sum_{i=1}^{n-1} z_i w_i^\top \Delta_\theta - \Delta_b \hat{\varrho}\kappa z_n^2, \tag{4.30}$$

$$\psi \triangleq z_{n-1} + w_n^\top \hat{\theta} - \sum_{j=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial x_j} x_{j+1} - \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}} \Gamma_\theta \tau_n - \sum_{j=2}^{n-1} \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}} \Gamma_\theta w_n z_j. \tag{4.31}$$

The second line of (4.29) is eliminated by the parameter update laws (4.27) and (4.28), and the non-positivity of $-\Delta_b \hat{\varrho}\kappa z_n^2$ can be established in the same way as in Section 2.1.2, thanks to the fact that $\kappa(x, \hat{\theta}) > 0$ and that $\Delta_b(t)$ and $\hat{\varrho}(t)$ can be made positive for all $t \geq 0$ by properly selecting ℓ_b and $\hat{\varrho}(0)$, respectively. In addition, note that a similar observation as the one in Remark 4.2 suggests that one can express ψ as $\psi = \bar{\psi}^\top(x, \hat{\theta})z$ with $\bar{\psi}$ smooth. The rest of the design boils down to determining the nonlinear damping gains ζ_i and κ to dominate the Δ_θ-terms, as discussed in what follows.

Proposition 4.2. *Consider the system* (4.25) *and the adaptive controller described by*

(4.89), (4.27), *and* (4.28). *Let the nonlinear damping gains be defined as*

$$\zeta_i(\underline{x_i}, \hat{\theta}) = \frac{1}{2}\left((n-i+1)\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_i|_{\mathrm{F}}^2 + \frac{1}{\epsilon_{\bar{\psi}}}\right), \tag{4.32}$$

$$\kappa(x, \hat{\theta}) = c_n + \zeta_n + \frac{1}{2}\epsilon_{\bar{\psi}}|\bar{\psi}|^2, \tag{4.33}$$

with $c_n > 0$ *and* $\epsilon_{(\cdot)} > 0$. *The parameter estimate* $\hat{\varrho}$ *is initialized such that* $\mathrm{sgn}(\hat{\varrho}(0)) = \mathrm{sgn}(b)$. *Then, all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} x(t) = 0$.

Proof. Recalling Remark 4.2 and invoking Lemma A.1 yields

$$\begin{aligned} z_i w_i^\top \Delta_\theta &= z_i \Delta_\theta^\top \bar{W}_i \underline{z_i} \\ &\le \frac{1}{2}\left(\frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}} + \epsilon_{\Delta_\theta}\delta_{\Delta_\theta}|\bar{W}_i|_{\mathrm{F}}^2\right) z_i^2 + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}|\underline{z_{i-1}}|^2, \end{aligned} \tag{4.34}$$

$$\begin{aligned} z_n \psi &= z_n \bar{\psi} z \\ &\le \frac{1}{2}\left(\frac{1}{\epsilon_{\bar{\psi}}} + \epsilon_{\bar{\psi}}|\bar{\psi}|^2\right) z_n^2 + \frac{1}{2\epsilon_{\bar{\psi}}}|\underline{z_{n-1}}|^2. \end{aligned} \tag{4.35}$$

Applying these inequalities to (4.29) yields

$$\begin{aligned} \dot{V} =& -\sum_{i=1}^{n}(c_i + \zeta_i)z_i - \kappa z_n^2 + \sum_{i=1}^{n-1} z_i w_i^\top \Delta_\theta - \Delta_b \hat{\varrho}\kappa z_n^2 + z_n\psi \\ \le& -\sum_{i=1}^{n}\left(c_i + \frac{1}{2\epsilon_\psi}\right)z_i - \frac{1}{2}\epsilon_{\bar{\psi}}|\bar{\psi}|^2 z_n^2 - \Delta_b\hat{\varrho}\kappa z_n^2 + z_n\psi \\ \le& -\sum_{i=1}^{n} c_i z_i^2 - \Delta_b \hat{\varrho}\kappa z_n^2. \end{aligned} \tag{4.36}$$

Note now the following two facts: 1) there exists ℓ_b such that $0 < |\ell_b| \le |b|$, $\mathrm{sgn}(\ell_b) = \mathrm{sgn}(b)$, and $\mathrm{sgn}(\Delta_b) = \mathrm{sgn}(b)$, due to Assumption 1.3 and; 2) the selection of $\hat{\varrho}(0)$ such that $\mathrm{sgn}(\hat{\varrho}(0)) = \mathrm{sgn}(b)$ guarantees that $\hat{\varrho}$ driven by (4.28) satisfies $\mathrm{sgn}\big(\hat{\varrho}(t)\big) = \mathrm{sgn}(b)$. Therefore, $-\Delta_b\hat{\varrho}\kappa z_n^2 = -|\Delta_b||\hat{\varrho}|\kappa z_n^2 \le 0$. This yields

$$\dot{V} \le -\sum_{i=1}^{n} c_i z_i^2 \le 0, \tag{4.37}$$

which has the same supply rate as that of (4.24). Note that V is positive definite and radially unbounded in $(z, \ell_\theta - \hat{\theta}, \ell_b^{-1} - \hat{\varrho})$. Therefore, using the same argument as in the proof of Proposition 4.1, boundedness of all closed-loop system trajectories and convergence of x to 0 can be concluded. □

Remark 4.4. *The passivity interpretations in Section 2.1.1 and Section 2.1.2 are still valid in the backstepping case, with some necessary changes to the interconnecting signals. The schematic interpretation for the closed-loop system consisting of the system* (4.25) *and the adaptive controller described by* (4.26)*,* (4.27)*, and* (4.28) *is shown in Fig. 4.1.*

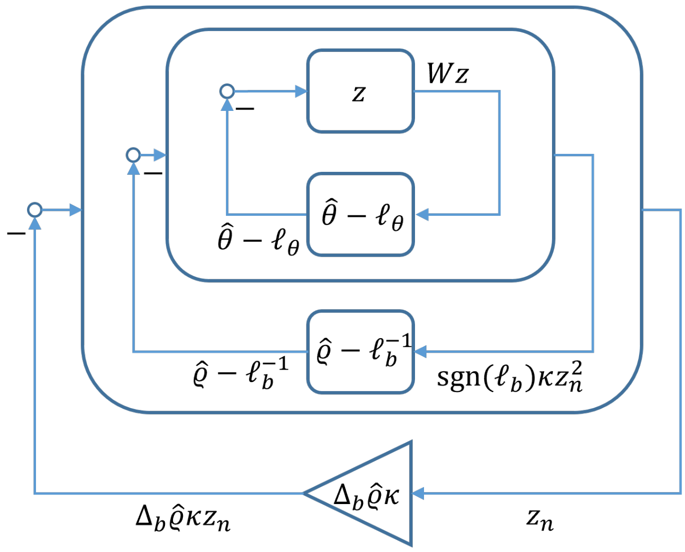

Figure 4.1: Representation of the closed-loop system described by equations (4.25) (4.26)**,** (4.27)**, and** (4.28) **as the interconnection of passive subsystems.**

4.2 Immersion-and-Invariance Scheme with Dynamic Scaling

This section extends the results in Section 2.2 to n-dimensional nonlinear systems. Consider a linearly parametrized nonlinear system (not necessarily in *lower-triangular* form) described by the equations

$$\dot{x} = f_u(x,u) + \Phi^\top(x)\theta, \tag{4.38}$$

where $x(t) \in \mathbb{R}^n$ is the state; $u(t) \in \mathbb{R}^m$ is the input; $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a smooth mapping; $\Phi : \mathbb{R}^n \to \mathbb{R}^{q\times n}$ is a smooth mapping satisfying Assumption 1.4; and $\theta(t) \in \mathbb{R}^q$ is the vector of time-varying parameters satisfying Assumption 1.1. The parameter estimate is given by the sum of a dynamic part $\hat{\theta}$ and a static part $\beta(x)$. In the spirit of the I&I scheme, one would expect the mapping $\beta : \mathbb{R}^n \to \mathbb{R}^q$ to be selected in such a way that

$$\frac{\partial \beta}{\partial x} = \Gamma_\theta \Phi(x), \tag{4.39}$$

which guarantees that the *off-the-manifold* error $z_\theta \triangleq \hat{\theta} - \ell_\theta + \beta(x)$ satisfies $\dot{z}_\theta = -\Gamma_\theta \Phi(x)(z_\theta - \Delta_\theta)$, with $\Gamma_\theta = \Gamma_\theta^\top \succ 0$. However, a challenge in this case is that the derivation of the β function cannot be obtained directly from the integral of Φ over x as in (2.31), since x is not a scalar. The need for solving the PDE (4.39) can be restrictive in practical scenarios. To avoid such a restriction, the rest of the section presents a joint estimator-controller scheme by exploiting the *dynamic scaling* technique.

4.2.1 Dynamic Scaling Estimator

The *dynamic scaling* technique has been originally developed for high-gain observers [103], and extended to the I&I scheme with constant system parameters in [5, 65]. The key of this method is to scale the off-the-manifold variable $\theta - \ell_\theta + \beta$ by a factor $\frac{1}{r(t)}$, namely, letting the new error variable be

$$z_\theta \triangleq \frac{\hat{\theta} - \ell_\theta + \beta(x, \hat{x})}{r}, \tag{4.40}$$

where $r(t) \in \mathbb{R}$, the update law of which is to be determined; $\hat{\theta}(t) \in \mathbb{R}^q$ is the dynamic part of the parameter estimate $\hat{\theta} + \beta(x, \hat{x})$; $\ell_\theta \in \mathbb{R}^q$ is the constant vector of *congealed* parameters; and $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^q$ is the static part of the parameter estimate and is selected as

$$\beta(x, \hat{x}) = \Gamma_\theta \Phi(\hat{x}) x, \tag{4.41}$$

with $\Gamma_\theta = \Gamma_\theta^\top \succ 0$. The auxiliary state[3] $\hat{x}$ is updated using the filter

$$\dot{\hat{x}} = f_u(x, u) + \Phi^\top(x)\big(\hat{\theta} + \beta(x, \hat{x})\big) - L(x, r, \tilde{x})\tilde{x}, \tag{4.42}$$

[3] $\hat{x}$ can also be interpreted as an estimate of x, but it is called auxiliary state since x is known and $\hat{x}$ is introduced solely to help avoid solving the PDE (4.39).

where $\tilde{x} \triangleq \hat{x} - x$, and $L(x, r, \tilde{x})$ is the injection gain to be determined. The dynamic parameter estimate $\hat{\theta}$ is driven by the update law

$$\dot{\hat{\theta}} = -\Gamma_\theta \Phi(\hat{x}) \Big(f_u(x,u) + \Phi^\top(x) \big(\hat{\theta} + \beta(x,\hat{x})\big) \Big) - \frac{\partial \beta}{\partial \hat{x}} \dot{\hat{x}}. \tag{4.43}$$

Then, the resulting $\tilde{x}$-dynamics and z_θ-dynamics are described by the equations

$$\dot{\tilde{x}} = \dot{\hat{x}} - \dot{x} = -L(x, r, \tilde{x})\tilde{x} + r\Phi^\top(x)\left(z_\theta - \frac{\Delta_\theta}{r} \right) \tag{4.44}$$

and

$$\begin{aligned} \dot{z}_\theta &= \frac{1}{r}\left(\dot{\hat{\theta}} + \frac{\partial \beta}{\partial x}\dot{x} + \frac{\partial \beta}{\partial \hat{x}}\dot{\hat{x}} \right) - \frac{\dot{r}}{r^2}\big(\hat{\theta} - \ell_\theta + \beta(x,\hat{x})\big) \\ &= -\Gamma_\theta \Phi(\hat{x})\Phi^\top(x)\left(z_\theta - \frac{\Delta_\theta}{r} \right) - \frac{\dot{r}}{r} z_\theta, \end{aligned} \tag{4.45}$$

respectively. Note that due to the smoothness of $\Phi(\cdot)$ and by Lemma A.3, one has $\Phi(\hat{x}) = \Phi(x) + D(x,\tilde{x})(I_n \otimes \tilde{x})$ for some smooth mapping $D : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{q \times n^2}$. With this fact in mind the dynamics of z_θ can be re-written as

$$\dot{z}_\theta = -\Gamma_\theta \big(\Phi(x) + D(x,\tilde{x})(I_n \otimes \tilde{x})\big)\Phi^\top(x)\left(z_\theta - \frac{\Delta_\theta}{r} \right) - \frac{\dot{r}}{r} z_\theta. \tag{4.46}$$

Lemma 4.1. *Consider the system* (4.38)*, the auxiliary state filter* (4.42)*, and the dynamic parameter estimate update law* (4.43)*, with the filter injection gain defined by*

$$L(x, r, \tilde{x}) = \lambda r^2 I + \bar{L}(x, r, \tilde{x}), \tag{4.47}$$

and the dynamic scaling coefficient updated as

$$\dot{r} = \kappa n |D(x,\tilde{x})|_{\mathrm{F}}^2 |\tilde{x}|^2 r, \tag{4.48}$$

where $\bar{L}(x, r, \tilde{x}) = \epsilon\kappa n r^2 |D(x,\tilde{x})|_{\mathrm{F}}^2 I$*,* $r(0) = 1$*,* $\kappa I - 2\Gamma_\theta \succeq 0$*,* $\lambda > 0$*, and* $\epsilon > 0$*. Then the function* $V_{z_\theta \tilde{x} r}(z_\theta, \tilde{x}, r) = 2|z_\theta|^2_{\Gamma_\theta^{-1}} + \frac{1}{2}\lambda|\tilde{x}|^2 + \frac{1}{2}\lambda\epsilon r^2$ *satisfies the dissipation inequality*

$$\dot{V}_{z_\theta \tilde{x} r} \leq -|\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda^2|\tilde{x}|^2 + 6\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2. \tag{4.49}$$

Moreover, if the parameter vector θ is constant, then $z_\theta \in \mathcal{L}_\infty$, $r \in \mathcal{L}_\infty$, $\tilde{x} \in \mathcal{L}_\infty \cap \mathcal{L}_2$, and $\Phi^\top(x) z_\theta \in \mathcal{L}_2$.

Proof. First note that r is non-decreasing due to (4.48). Thus $r(t) \geq r(0) = 1$ and $|\frac{\Delta_\theta(t)}{r}| \leq \delta_{\Delta_\theta}$, for all $t \geq 0$. Consider now the function $V_{z_\theta}(z_\theta) = \frac{1}{2} z_\theta^\top \Gamma_\theta^{-1} z_\theta$. Taking its time derivative along the trajectories of (4.46) yields

$$\begin{aligned}
\dot{V}_{z_\theta} \leq & -|\Phi^\top(x) z_\theta|^2 + \left(\frac{1}{4}|\Phi^\top(x) z_\theta|^2 + \delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2\right) \\
& + \left(\frac{1}{4}|\Phi^\top(x) z_\theta|^2 + |D(x,\tilde{x})(I_n \otimes \tilde{x})|_{\mathrm{F}}^2 |z_\theta|^2\right) \\
& + \left(\frac{1}{4}\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2 + |D(x,\tilde{x})(I_n \otimes \tilde{x})|_{\mathrm{F}}^2 |z_\theta|^2\right) \\
& - \kappa n |D(x,\tilde{x})|_{\mathrm{F}}^2 |\tilde{x}|^2 z_\theta^\top \Gamma_\theta^{-1} z_\theta \\
\leq & -\frac{1}{2}|\Phi^\top(x) z_\theta|^2 + \frac{5}{4}\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2.
\end{aligned} \tag{4.50}$$

Proceed by considering the function $V_{z_\theta\tilde{x}}(z_\theta,\tilde{x}) = \frac{4}{\lambda} V_{z_\theta}(z_\theta) + \frac{1}{2}|\tilde{x}|^2$, the time derivative of which along (4.44) and (4.46) satisfies

$$\begin{aligned}
\dot{V}_{z_\theta\tilde{x}} \leq & -\frac{2}{\lambda}|\Phi^\top(x) z_\theta|^2 + \frac{5}{\lambda}\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2 - \tilde{x}^\top L(x,r,\tilde{x})\tilde{x} \\
& + \left(\frac{1}{\lambda}|\Phi^\top(x) z_\theta|^2 + \frac{1}{4}\lambda r^2 |\tilde{x}|^2\right) \\
& + \left(\frac{1}{\lambda}\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2 + \frac{1}{4}\lambda r^2 |\tilde{x}|^2\right) \\
= & -\frac{1}{\lambda}|\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda r^2 |\tilde{x}|^2 - \tilde{x}^\top \bar{L}\tilde{x} + \frac{6}{\lambda}\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2.
\end{aligned} \tag{4.51}$$

Finally, consider the function $V_{z_\theta\tilde{x}r}(z_\theta,\tilde{x},r) = \lambda V_{z_\theta\tilde{x}}(z_\theta,\tilde{x}) + \frac{1}{2}\lambda\epsilon r^2$. Taking the time derivative along (4.44), (4.46), and (4.48) yields

$$\begin{aligned}
\dot{V}_{z_\theta\tilde{x}r} \leq & -|\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda^2 r^2 |\tilde{x}|^2 \\
& + 6\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2 - \lambda\big(\tilde{x}^\top \bar{L}\tilde{x} - \epsilon\kappa n r^2 |D(x,\tilde{x})|_{\mathrm{F}}^2 |\tilde{x}|^2\big) \\
\leq & -|\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda^2 |\tilde{x}|^2 + 6\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2.
\end{aligned} \tag{4.52}$$

If θ is constant, one can select the congealed parameters as $\ell_\theta = \theta$, yielding $\delta_{\Delta_\theta} = 0$. Then $\dot{V}_{z_\theta\tilde{x}r} \leq -|\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda^2 |\tilde{x}|^2 \leq 0$. Due to the positive definiteness and radial unboundedness of $V_{z_\theta \tilde{x} r}$ one can conclude that $z_\theta \in \mathcal{L}_\infty$, $\tilde{x} \in \mathcal{L}_\infty$, and $r \in \mathcal{L}_\infty$. From the supply rate it can be proven that $\Phi^\top(x) z_\theta \in \mathcal{L}_2$ and $\tilde{x} \in \mathcal{L}_2$. Combining these properties together completes the proof. □

Remark 4.5. *Compared to its counterpart in [65], Lemma 4.1 does not use over-parametrization, which makes the result applicable to non-overparametrized controllers. In addition, Lemma 4.1 uses the Frobenius norm instead of the induced 2-norm, which can be turned into pre-computed expressions without the need for online norm computation.*

4.2.2 ISS Controller

Although in the presence of time-varying θ, the standalone estimator design previously discussed cannot guarantee boundedness of the estimator states, this problem can be solved by a joint estimator-controller design. To see this, consider a linearly parametrized, input affine, nonlinear system described by the equation

$$\dot{x} = f(x) + g(x)u + \Phi^\top(x)\theta, \tag{4.53}$$

which is a special form of (4.38) with $f_u(x, u) = f(x) + g(x)u$. Consider a nominal continuous control law $v(x, \ell_\theta)$ parametrized by the *congealed* parameters ℓ_θ (assumed to be known for the nominal control law, but this assumption can be relaxed by adaptation). The resulting closed-loop system is described by the equation

$$\dot{x} = f(x) + g(x)v(x, \ell_\theta) + \Phi^\top(x)(\ell_\theta + \Delta_\theta) = f_\ell(x). \tag{4.54}$$

To be able to conclude stability properties one typically needs to make a structural assumption based on the plant and the nominal controller.

Assumption 4.1. *The system* (4.54) *has a globally asymptotically stable equilibrium at* $x = x_*$.

Assumption 4.1 means that the system can be robustly stabilized in the presence of the time-varying perturbation $\Delta_\theta(t)$ only with the nominal information of a constant ℓ_θ (for example, one can select it as the geometric centre of the set Θ to which $\theta(t)$ belongs to). Up to this point, the control law is designed in the spirit of robust control, but since

ℓ_θ is not assumed to be known in the context of this thesis, parameter adaptation has to be considered. To do this, replace ℓ_θ with the parameter estimate $\hat{\theta}+\beta$, which yields an adaptive control law of the form[4] $v(x, \hat{\theta}+\beta)$ and the closed-loop dynamics

$$\dot{x} = f(x) + g(x)v(x, \hat{\theta}+\beta) + \Phi^\top(x)(\hat{\theta}+\beta - rz_\theta + \Delta_\theta) \triangleq f_{\hat{\theta}\beta}, \tag{4.55}$$

Proposition 4.3. *Consider the system* (4.53) *and the dynamic scaling estimator given by* (4.41), (4.42), (4.43), (4.48), (4.47). *Assume that Assumption 4.1 holds and that there exists a positive definite (centred at $x = x_*$) and radially unbounded function V_x and a control law $v(x, \hat{\theta}+\beta)$ such that the time derivative of V_x satisfies the dissipation inequality*

$$\dot{V}_x = \frac{\partial V_x}{\partial x} f_{\hat{\theta}\beta} \leq -W(x) - \sigma_\Phi |\Phi(x)|_{\mathrm{F}}^2 + |\Phi^\top(x) z_\theta|^2, \tag{4.56}$$

where $W : \mathbb{R}^n \to \mathbb{R}_+$ is a differentiable positive-definite function centred at $x = x_*$; and σ_Φ is a damping coefficient that can be adjusted in the interval $(0, \bar{\sigma}_\Phi]$, with $\bar{\sigma}_\Phi \geq 6\delta_{\Delta_\theta}$, by tuning the control law. Then, all trajectories of the closed-loop system are bounded and $\lim_{t\to\infty} x(t) = x_*$.

Proof. Recall Lemma 4.1 and consider the function $V = V_x + V_{z_\theta \tilde{x} r}$. Using (4.49) and (4.56), the time derivative of V along the trajectories of the system satisfies

$$\begin{aligned} \dot{V} =& \dot{V}_x + \dot{V}_{z_\theta \tilde{x} r} \\ \leq& -W(x) - \sigma_\Phi |\Phi(x)|_{\mathrm{F}}^2 + |\Phi^\top(x) z_\theta|^2 \\ & - |\Phi^\top(x) z_\theta|^2 - \frac{1}{2}\lambda^2 |\tilde{x}|^2 + 6\delta_{\Delta_\theta}^2 |\Phi(x)|_{\mathrm{F}}^2. \end{aligned} \tag{4.57}$$

Selecting $\sigma_\Phi = 6\delta_{\Delta_\theta}^2$ yields $\dot{V} \leq -W(x) \leq 0$. Due to the positive definiteness and radially unboundedness of V, x, z_θ, $\tilde{x}$, and r are bounded, which proves boundedness of the trajectories of the closed-loop system. Furthermore, both $\frac{\partial W}{\partial x}$ and $\dot{x}$ are bounded due to diffentiability of W, f, g, v and boundedness of the relevant signals. Finally, invoking Lemma A.5 yields $\lim_{t\to\infty} x(t) = x_*$. □

[4] Note that there is a slight abuse of notation since the parametrized control needed may not be precisely $v(x, \hat{\theta}+\beta)$, that is, the nominal control law evaluated with the parameter estimate. For example, in the backstepping design one has to add additional terms to compensate for the effect of $\frac{\mathrm{d}}{\mathrm{d}t}(\hat{\theta}+\beta)$ caused by the substitution of $\hat{\theta}+\beta$ for ℓ_θ.

From (4.56) we can see that the controller in Proposition 4.3 guarantees ISS of the plant-controller subsystem with respect to the input $\Phi^\top(x)z_\theta$. The aim of designing such a controller is to use strengthened damping terms to construct the stabilizing term $-\delta_{\Delta_\theta}^2|\Phi(x)|_{\mathrm{F}}^2$ in $\dot{V}_x$ to dominate the positive term $\delta_{\Delta_\theta}^2|\Phi(x)|_{\mathrm{F}}^2$ in $\dot{V}_{z_\theta\tilde{x}r}$ resulting from the estimator design, and to treat $\Phi^\top(x)z_\theta$ as an exogenous input, which is dominated, in turn, by the stabilizing term $-|\Phi^\top(x)z_\theta|^2$ in $\dot{V}_{z_\theta\tilde{x}r}$. Then, boundedness and convergence properties of the closed-loop system are guaranteed by the small-gain-like condition enforced by (4.49) and (4.56).

In practice, this ISS controller is applicable to at least three types of systems. The first type is given by systems satisfying the *matching condition*, which has been discussed in Section 2.2. The second type is given by systems satisfying the *extended matching condition*, which will be discussed in Section 6.1 via a practical example. The third type is given by systems in *parametric strict-feedback* form. This type of systems requires exploiting the estimator-controller structure recursively and overparametrization is needed. This is to be discussed in the next subsection.

4.2.3 Overparametrized Backstepping Design

Consider now the *lower triangular* system (4.1). This requires *backstepping* techniques as the parametric uncertainty is *unmatched.* Since the *backstepping* design requires the design variables to have *lower-triangular* dependency on the state variables, that is, the variables used at step i, only depends on $\underline{x_i}$. In this sense, the non-recursive design introduced previously does not fit in the *backstepping* scheme as the β function and the dynamic scaling estimator, in general, does not satisfy the *lower-triangular* dependency. To circumvent this limitation, consider the overparametrization scheme discussed in [5, Section 4.3], that is, consider n parameter estimates $\hat{\theta}_i + \beta_i$, $i = 1, \ldots, n$, for the same unknown parameter θ. To this end, define, for $i = 1, \ldots, n$, the scaled *off-the-manifold* variable

$$z_{\theta i} = \frac{\hat{\theta}_i - \ell_\theta + \beta_i(x_i, \underline{\hat{x}_i})}{r_i}, \tag{4.58}$$

where $\beta_i : \mathbb{R} \times \mathbb{R}^i \to \mathbb{R}^q$ is selected as

$$\beta_i(x_i, \underline{\hat{x}_i}) = \gamma_i \phi_i(\underline{\hat{x}_i}) x_i \tag{4.59}$$

and it is such that

$$\frac{\partial \beta_i}{\partial x_i} = \gamma_i \phi_i(\underline{\hat{x}_i}), \tag{4.60}$$

with $\gamma_i > 0$. Note that the auxiliary filter (4.42) satisfies the *lower-triangular* dependency on the state variables and therefore can be used for backstepping. The overparametrized structure suggests dividing the counterpart of (4.42) into n filters, that is, the filters

$$\dot{\hat{x}}_i = x_{i+1} + \phi_i^\top(\underline{x_i})(\hat{\theta}_i + \beta_i) - l_i(\underline{x_i}, r_i, \underline{\tilde{x}_i})\tilde{x}_i, \tag{4.61}$$

$i = 1, \ldots, n$, where $l_i : \mathbb{R}^i \times \mathbb{R} \times \mathbb{R}^i \to \mathbb{R}$ is the injection gain function to be determined. The update law for $\hat{\theta}_i$ is selected as

$$\dot{\hat{\theta}}_i = -\gamma_i \phi_i(\underline{\hat{x}_i})\big(x_{i+1} + \phi_i^\top(\underline{x_i})(\hat{\theta}_i + \beta_i)\big) - \sum_{j=1}^{i} \frac{\partial \beta_j}{\partial \hat{x}_j}\dot{\hat{x}}_j. \tag{4.62}$$

Then, the resulting $\tilde{x}_i$-dynamics and $z_{\theta i}$-dynamics are

$$\begin{aligned} \dot{\tilde{x}}_i &= \dot{\hat{x}}_i - \dot{x}_i \\ &= -l_i(\underline{x_i}, r_i, \underline{\tilde{x}_i})\tilde{x}_i + r_i \phi_i^\top(\underline{x_i})\left(z_{\theta_i} - \frac{\Delta_\theta}{r_i}\right) \end{aligned} \tag{4.63}$$

and

$$\begin{aligned} \dot{z}_{\theta i} &= \frac{1}{r_i}\left(\dot{\hat{\theta}}_i + \frac{\partial \beta_i}{\partial x_i}\dot{x}_i + \frac{\partial \beta_i}{\partial \underline{\hat{x}_i}}\dot{\underline{\hat{x}_i}}\right) - \frac{\dot{r}_i}{r_i^2}(\hat{\theta} - \ell_\theta + \beta) \\ &= -\gamma_i \phi_i(\underline{\hat{x}_i})\phi_i^\top(\underline{x_i})\left(z_\theta - \frac{\Delta_\theta}{r_i}\right) - \frac{\dot{r}_i}{r_i} z_{\theta i} \\ &= -\gamma_i\big(\phi_i(\underline{x_i}) + D_i(\underline{x_i}, \underline{\tilde{x}_i})\underline{\tilde{x}_i}\big)\phi_i^\top(\underline{x_i})\left(z_{\theta i} - \frac{\Delta_\theta}{r_i}\right) - \frac{\dot{r}_i}{r_i} z_{\theta i}, \end{aligned} \tag{4.64}$$

respectively, where $D_i : \mathbb{R}^i \times \mathbb{R}^i \to \mathbb{R}^{q \times i}$ is a smooth mapping that is well-defined and known due to smoothness of ϕ_i and Lemma A.3. For conciseness of expression, the value of the mapping ϕ_i at $\underline{x_i}$, namely $\phi_i(\underline{x_i})$ is written as ϕ_i where appropriate, while $\phi_i(\underline{\hat{x}_i})$

is always written in its complete form to avoid confusion. Due to the overparametrized design, the counterpart of Lemma 4.1 is decomposed and allocated to each parameter estimate and auxiliary filter state variable.

Lemma 4.2. *Consider the system* (4.1)*, the auxiliary state filters* (4.61)*, and the dynamic parameter estimates updated by* (4.62)*, with the filter injection gain defined by*

$$l_i(\underline{x_i}, r_i, \underline{\tilde{x}_i}) = \lambda_i + \bar{l}_i(\underline{x_i}, r_i, \underline{\tilde{x}_i}), \tag{4.65}$$

and the dynamic scaling coefficients updated by

$$\dot{r}_i = \kappa_i |D_i(\underline{x_i}, \underline{\tilde{x}_i})\underline{\tilde{x}_i}|^2 r_i, \tag{4.66}$$

where $\bar{l}_i(\underline{x_i}, r_i, \underline{\tilde{x}_i}) = \epsilon\kappa_i r_i^2 |D_i(\underline{x_i}, \underline{\tilde{x}_i})\underline{\tilde{x}_i}|^2$*,* $r_i(0) = 1$*,* $\kappa_i \geq 2\gamma_i$*,* $\lambda_i > 0$*, and* $\epsilon > 0$*. Then the function* $V_{z_{\theta i}\tilde{x}_i r_i}(z_{\theta i}, \tilde{x}_i, r_i) = 2\gamma_i^{-1}|z_{\theta_i}|^2 + \frac{1}{2}\lambda_i\tilde{x}_i^2 + \frac{1}{2}\lambda_i\epsilon r_i^2$ *satisfies the dissipation inequality*

$$\dot{V}_{z_{\theta i}\tilde{x}_i r_i} \leq -\left(\phi_i^\top(\underline{x_i})z_{\theta i}\right)^2 - \frac{1}{2}\lambda_i^2\tilde{x}_i^2 + 6\delta_{\Delta_\theta}^2|\phi_i(\underline{x_i})|^2. \tag{4.67}$$

Proof. Due to the update law (4.66) and the initial condition of r_i, $r_i(t) \geq r_i(0) = 1$ and therefore $|\frac{\Delta_\theta(t)}{r_i}| \leq \delta_{\Delta_\theta}$, for all $t \geq 0$. Consider now the function $V_{z_{\theta i}}(z_{\theta i}) = \frac{1}{2\gamma_i}|z_{\theta i}|^2$. The time derivative of $V_{z_{\theta_i}}$ along the trajectories of the closed-loop system satisfies

$$\begin{aligned}
\dot{V}_{z_{\theta i}} \leq& -(\phi_i^\top z_{\theta i})^2 - \kappa_i\gamma_i^{-1}|D_i\underline{\tilde{x}_i}|^2|z_{\theta i}|^2 + \left(\frac{1}{4}(\phi_i^\top z_{\theta i})^2 + \delta_{\Delta_\theta}^2|\phi_i|^2\right) \\
&+ \left(\frac{1}{4}(\phi_i^\top z_{\theta i})^2 + |D_i\underline{\tilde{x}_i}|^2|z_{\theta i}|^2\right) + \left(\frac{1}{4}\delta_{\Delta_\theta}^2|\phi_i|^2 + |D_i\underline{\tilde{x}_i}|^2|z_{\theta i}|^2\right) \\
\leq& -\frac{1}{2}(\phi_i^\top z_{\theta i})^2 + \frac{5}{4}\delta_{\Delta_\theta}^2|\phi_i|^2.
\end{aligned} \tag{4.68}$$

Then, consider the function $V_{z_{\theta i}\tilde{x}_i}(z_{\theta i}, \tilde{x}_i) = \frac{4}{\lambda_i}V_{z_{\theta i}}(z_{\theta i}) + \frac{1}{2}\tilde{x}_i^2$. The time derivative of each $V_{z_{\theta i}\tilde{x}_i}$ along the system trajectories satisfies

$$\begin{aligned}
\dot{V}_{z_{\theta i}\tilde{x}_i} \leq& -\frac{2}{\lambda_i}(\phi_i^\top z_{\theta i})^2 + \frac{5}{\lambda_i}\delta_{\Delta_\theta}^2|\phi_i|^2 - l_i^2\tilde{x}_i^2 \\
&+ \left(\frac{1}{\lambda_i}(\phi_i^\top z_{\theta i})^2 + \frac{1}{4}\lambda_i r_i^2\tilde{x}_i^2\right) + \left(\frac{1}{\lambda_i}\delta_{\Delta_\theta}^2|\phi_i|^2 + \frac{1}{4}\lambda_i r_i^2\tilde{x}_i^2\right) \\
=& -\frac{1}{\lambda_i}|\phi_i^\top z_{\theta_i}|^2 - \frac{1}{2}\lambda_i r_i^2\tilde{x}_i^2 - \bar{l}_i\tilde{x}_i^2 + \frac{6}{\lambda_i}\delta_{\Delta_\theta}^2|\phi_i|^2.
\end{aligned} \tag{4.69}$$

Finally, consider the function $V_{z_{\theta i}\tilde{x}_i r_i}(z_{\theta i}, \tilde{x}_i, r_i) = \lambda_i V_{z_{\theta i}\tilde{x}_i}(z_{\theta i}, \tilde{x}_i) + \frac{1}{2}\epsilon\lambda_i r_i^2$. Taking the time derivative along the system trajectories yields

$$\begin{aligned}\dot{V}_{z_{\theta i}\tilde{x}_i r_i} &\leq -(\phi_i^\top z_{\theta i})^2 - \frac{1}{2}\lambda_i^2 r_i^2 \tilde{x}_i^2 + 6\delta_{\Delta_\theta}^2 |\phi_i|^2 - \lambda_i\big(\bar{l}_i \tilde{x}_i^2 - \epsilon\kappa_i r_i^2 |D_i \underline{\tilde{x}_i}|^2\big)\\ &\leq -(\phi_i^\top z_{\theta i})^2 - \frac{1}{2}\lambda_i^2 \tilde{x}_i^2 + 6\delta_{\Delta_\theta}^2 |\phi_i|^2. \end{aligned} \tag{4.70}$$

□

To introduce the overparametrized backstepping controller, for $i = 1, \ldots, n$, consider the variable definitions:

- the error variables

$$z_0 = 0, \tag{4.71}$$

$$z_i = x_i - \alpha_{i-1}, \tag{4.72}$$

- the virtual control laws

$$\alpha_0 = 0, \tag{4.73}$$

$$\begin{aligned}\alpha_i(\underline{x_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i}) = &-z_{i-1} - (c_i + \zeta_i) z_i - \phi_i^\top(\underline{x_i})(\hat{\theta}_i + \beta_i)\\ &+ \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j}\big(x_{j+1} + \phi_j^\top(\underline{x_j})(\hat{\theta}_j + \beta_j)\big)\\ &+ \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}_j}\dot{\hat{\theta}}_j + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \hat{x}_j}\dot{\hat{x}}_j + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial r_j}\dot{r}_j, \end{aligned} \tag{4.74}$$

where $c_i > 0$ is the constant feedback gain, and $\zeta_i(\underline{x_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i})$ is the nonlinear feedback gain to be determined. To understand the design of the control laws, we note a structural property that is essential for enforcing the $\mathcal{L}_2$ property of the regressor ϕ_i.

Lemma 4.3. *The regressor* $\psi_i(\underline{z_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i}) \triangleq \phi_i(\underline{x_i})$ *can be expressed as* $\psi_i(\underline{z_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i}) = \sum_{j=1}^{i} \bar{\Psi}_{i,j}(\underline{z_j}, \underline{\hat{\theta}_j}, \underline{\hat{x}_j}, \underline{r_j}) z_j \triangleq \bar{\Psi}_i(\underline{z_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i})\underline{z_i}$, *where* $\bar{\Psi}_{i,j}$*'s are smooth mappings.*

Proof. First note that $\phi_i(\underline{x_i})$ can be expressed as $\phi_i(\underline{x_i}) = \sum_{j=1}^{i} \bar{\Phi}_{i,j}(\underline{x_j}) x_j$. To see this, for $j = 1, \ldots, i$, define a smooth mapping $\bar{\Phi}_{i,j}$, such that $\bar{\Phi}_{i,j}(\underline{x_j}) x_j = \phi_i([\underline{x_j}, 0]^\top) - \phi_i([\underline{x_{j-1}}, 0]^\top)$, which is feasible due to Assumption 1.4 and Lemma A.3. Similar to what is discussed in Remark 4.2, z_j depends on $\underline{x_j}$, $\underline{\hat{\theta}_j}$, $\underline{\hat{x}_j}$, $\underline{r_j}$, and $z_j = 0 \Leftrightarrow x_j = 0$. Therefore the equation $\psi_i(\underline{z_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i}) = \phi_i(\underline{x_i}) = \sum_{j=1}^{i} \bar{\Phi}_{i,j}(\underline{x_j})x_j = \sum_{j=1}^{i} \bar{\Psi}_{i,j}(\underline{z_j}, \underline{\hat{\theta}_j}, \underline{\hat{x}_j}, \underline{r_j})z_j$ holds, which completes the proof. □

Lemma 4.3 shows the feasibility of decomposing the regressor $\phi_i(\underline{x_i})$ into the components $\bar{\Psi}_{i,j}z_j$ that have lower-triangular dependency on the closed-loop signals. This allows distributing the $\delta_{\Delta_\theta}^2|\phi_i|^2$ terms (from Lemma 4.2) to be dominated to the recursive steps of the backstepping design, as is explained in the following result.

Proposition 4.4. *Consider system* (4.1), *the dynamic scaling estimator described by* (4.61), (4.62), (4.66), *and the virtual control laws* (4.74) *with the nonlinear damping gains*

$$\zeta_i(\underline{x_i}, \underline{\hat{\theta}_i}, \underline{\hat{x}_i}, \underline{r_i}) = \zeta_{i1} + \zeta_{i2} + \zeta_{i3} + \zeta_{i4}, \tag{4.75}$$

where

$$\zeta_{i1} \triangleq (n^2 + n)\frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}, \tag{4.76}$$

$$\zeta_{i2} \triangleq \frac{1}{2\epsilon_\phi}\left(r_i^2 + \sum_{j=1}^{i-1} r_j^2 \left(\frac{\partial \alpha_j}{\partial x_i}\right)^2\right), \tag{4.77}$$

$$\zeta_{i3} \triangleq \frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}|\bar{\Psi}_i|_{\mathrm{F}}^2\left(1 + \sum_{j=1}^{i-1}\left(\frac{\partial \alpha_j}{\partial x_i}\right)^2\right), \tag{4.78}$$

$$\zeta_{i4} \triangleq \sigma_\Psi \sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2, \tag{4.79}$$

with $\sigma_\Psi \triangleq 3\epsilon_\phi n^2 \delta_{\Delta_\theta}^2$, $\epsilon_{(\cdot)} > 0$, *and* $\bar{\Psi}_{(\cdot)}$, $\bar{\Psi}_{(\cdot,\cdot)}$ *defined as in Lemma 4.3. Let the actual control law be* $u = \alpha_n$. *Then, all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} x(t) = 0$.

Proof. First note that, for $i = 1, \ldots, n$, one has

$$\begin{aligned}\dot{z}_i &= \dot{x}_i - \dot{\alpha}_i \\ &= z_{i+1} + \alpha_i + \phi_i^\top \theta - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j}(x_{j+1} + \phi_j^\top \theta) \\ &\quad - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}_j}\dot{\hat{\theta}}_j - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \hat{x}_j}\dot{\hat{x}}_j - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial r_j}\dot{r}_j \\ &= z_{i+1} - z_{i-1} - (c_i + \zeta_i)z_i - \phi_i^\top (r_i z_{\theta i} - \Delta_\theta) + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j}\phi_j^\top (r_j z_{\theta j} - \Delta_\theta),\end{aligned} \tag{4.80}$$

where $z_0 = z_{n+1} = 0$. Consider the function $V_{z_i} = \frac{1}{2}z_i^2$. Its time derivative along the trajectories of the closed-loop system satisfies

$$\begin{aligned}\dot{V}_{z_i} &= z_i z_{i+1} - z_{i-1}z_i - (c_i + \zeta_i)z_i^2 - z_i\phi_i^\top (r_i z_{\theta i} - \Delta_\theta) + z_i \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j}\phi_j^\top (r_j z_{\theta j} - \Delta_\theta) \\ &\le z_i z_{i+1} - z_{i-1}z_i - (c_i + \zeta_i)z_i^2 \\ &\quad + \frac{\epsilon_\phi}{2}(\phi_i^\top z_{\theta i})^2 + \frac{1}{2\epsilon_\phi} r_i^2 z_i^2 + \sum_{j=1}^{i-1}\left(\frac{\epsilon_\phi}{2}(\phi_j^\top z_{\theta j})^2 + \frac{1}{2\epsilon_\phi} r_j^2 \left(\frac{\partial \alpha_{i-1}}{\partial x_j}\right)^2 z_i^2\right) \\ &\quad + \frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}|\bar{\Psi}_i|_{\mathrm{F}}^2 z_i^2 + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}|\underline{z_i}|^2 + \sum_{j=1}^{i-1}\left(\frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}|\bar{\Psi}_i|_{\mathrm{F}}^2 \left(\frac{\partial \alpha_{i-1}}{\partial x_j}\right)^2 z_i^2 + \frac{\delta_{\Delta_\theta}}{2\epsilon_{\Delta_\theta}}|\underline{z_j}|^2\right) \\ &\le z_i z_{i+1} - z_{i-1}z_i - c_i z_i^2 - \sigma_\Psi \sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2 z_i^2 + \frac{\epsilon_\phi}{2}\sum_{j=1}^{i}(\phi_j^\top z_{\theta j})^2.\end{aligned} \tag{4.81}$$

Consider now the function $V_z(z) = \frac{1}{2}|z|^2 = \sum_{i=1}^n V_{z_i}(z_i)$. Its time derivative along the trajectories of the closed-loop system satisfies

$$\begin{aligned}\dot{V}_z &= \sum_{i=1}^{n} \dot{V}_{z_i} \\ &\le \sum_{i=1}^{n} c_i z_i^2 - \sigma_\Psi \sum_{i=1}^{n}\sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2 z_j^2 + \frac{\epsilon_\phi}{2}\sum_{i=1}^{n}\sum_{j=1}^{i}(\phi_j^\top z_{\theta j})^2 \\ &\le \sum_{i=1}^{n} c_i z_i^2 - \sigma_\Psi \sum_{i=1}^{n}\sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2 z_i^2 + \frac{\epsilon_\phi}{2} n \sum_{i=1}^{n}(\phi_i^\top z_{\theta i})^2.\end{aligned} \tag{4.82}$$

Recall that, by Lemma 4.2, $\dot{V}_{z_{\theta i}\tilde{x}_i r_i} \le -(\phi_i^\top z_{\theta i})^2 + 6\delta_{\Delta_\theta}^2|\phi_i|^2$. Consider the function $V = V_z + \frac{\epsilon_\phi}{2} n \sum_{i=1}^n V_{z_{\theta i}\tilde{x}_i r_i}$ and take its time derivative along the system trajectories. This yields

$$\begin{aligned}\dot{V} &= \dot{V}_z + \dot{V}_{z_{\theta i}\tilde{x}_i r_i} \\ &\leq \sum_{i=1}^{n} c_i z_i^2 - \sigma_\Psi \sum_{i=1}^{n}\sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2 z_i^2 + 3\epsilon_\phi n^2 \delta_{\Delta_\theta}^2 \sum_{i=1}^{n} |\phi_i|^2. \end{aligned} \tag{4.83}$$

By Lemma 4.3, $\phi_i = \sum_{j=1}^{i} \bar{\Psi}_{i,j} z_j$. Then, invoking the triangle inequality yields $|\phi_i| \leq \sum_{j=1}^{i} |\bar{\Psi}_{i,j}||z_j|$ and therefore $|\phi_i|^2 \leq i \sum_{j=1}^{i} |\bar{\Psi}_{i,j}|^2 z_j^2$. Note that $\sum_{i=1}^{n} |\phi_i|^2 \leq \sum_{i=1}^{n} \left(i \sum_{j=1}^{i} |\bar{\Psi}_{i,j}|^2 z_j^2 \right) \leq n \sum_{i=1}^{n} \sum_{j=1}^{n} |\bar{\Psi}_{j,i}|^2 z_i^2$ and $\sigma_\Psi = 3\epsilon_\phi n^2 \delta_{\Delta_\theta}^2$. This implies that $\dot{V} \leq \sum_{i=1}^{n} c_i z_i^2 \leq 0$.

Boundedness. Since the function V is positive definite and radially unbounded in z_i, $z_{\theta i}$, $\tilde{x}_i$, r_i, $i = 1, \ldots, n$, the inequality $\dot{V} \leq 0$ implies that these variables are bounded. Due to the recursive coupling between z_i and x_i, and $\hat{x}_i = x_i + \tilde{x}_i$, one can conclude, recursively, that x_i, $\hat{x}_i$, and $\hat{\theta}_i$ are also bounded, for $i = 1 \ldots, n$. This completes the proof of boundedness.

Convergence. Due to boundedness of the closed-loop signals, $\dot{z}$ is bounded. To perform an invariance-like analysis, invoking Lemma A.5 yields $\lim_{t\to+\infty} z(t) = 0$. This further indicates that $\lim_{t\to+\infty} x(t) = 0$, since $z = 0 \Leftrightarrow x = 0$. □

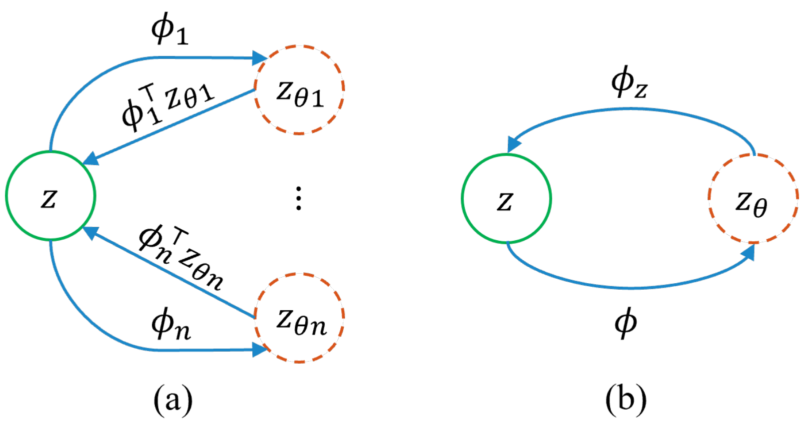

Figure 4.2: Schematic interpretation of the interconnected z-subsystem and the $(z_{\theta i}, \tilde{x}_i, r_i)$-subsystems (represented by $z_{\theta i}$ for conciseness): (a) the original structure of the underlying directed graph and (b) the graph after merging $z_{\theta i}$-nodes into the z_θ-node.

Note that it is possible to visualize the proof of Proposition 4.4 using the convention developed in Chapter 3, as shown in Fig. 4.2(a). Similar to the schematic interpretation of the scalar system shown in Fig. 2.2, the *congelation of variables* method removes $\dot{\theta}$ from the analysis and adds cyclic couplings between the z-subsystem and the $z_{\theta i}$-subsystems, though in a more complex structure (containing n cycles instead of one cycle) due to the overparametrization. Since the z-node and each of the $z_{\theta i}$-nodes are connected in a similar way except from being coupled by a different regressor ϕ_i, a simpler interpretation can be obtained by defining $\phi \triangleq [\phi_1^\top, \ldots, \phi_n^\top]^\top$ and $\phi_z = [\phi_1^\top z_{\theta 1}, \ldots, \phi_n^\top z_{\theta n}]^\top$, as the concatenated output and input, respectively, of the z-node, and by merging the n $z_{\theta i}$-nodes into one z_θ-node. Then, the overall graph reduces to a single cycle consisting of the z-node and the z_θ-node as shown in Fig. 4.2(b). If we regard the design parameter σ_Ψ in (4.79) as an adjustable parameter (which is the damping coefficient of the $|\phi_i|^2$-terms), the z-node is an *active node*. Furthermore, it is contained in every directed cycle of the underlying graph depicted in Fig. 4.2(a) and it is contained in the only remaining cycle of Fig. 4.2(b). Therefore, by Theorem 3.3 we can guarantee the existence of σ_Ψ such that Proposition 4.4 holds.

4.3 Identification-Based Scheme with Backstepping

In this section we consider the identification-based adaptive control problem for system (4.1). Before diving into the technical details, it is beneficial to first recall the overall picture of the identification-based adaptive control scheme that we have seen in Section 2.3. Its design procedure can be summarized in the following steps.

1. Design a parametric control law for the plant such that the closed-loop dynamics is ISS.

2. Build a set of stable filters to obtain an algebraic model parametrized by a vector of constant parameters ℓ_θ, with the parametrization error dynamics driven by a time-varying perturbation term related to Δ_θ. Use the classical method to design the state prediction and the parameter update law based on the model parametrized in ℓ_θ.

3. Select a storage function and compute the dissipation inequality for each subsystem.

4. Design the additional damping terms via a small-gain-like control synthesis that takes all subsystems into account.

For clarity, the rest of the section is organized into four parts according to this overall picture.

4.3.1 ISS Error Dynamics

We begin with by applying the *backstepping* techniques in Section 4.1, but since the identifier is designed in a modular way (that is, decoupled from the control law design), we do not cancel the $\dot{\hat{\theta}}$-terms in the recursive design procedures to avoid adding couplings between the control laws and the identifier which may "corrupt" the modularity. As a result, there is no need for designing the "tuning functions" τ_i. Despite this difference, one could still treat $x_2, \ldots, x_n$ as virtual control inputs and u as the actual control input, to recursively derive a smooth and invertible change of coordinates between the plant state x and the error state $z \triangleq [z_1, \ldots, z_n]^\top$, defined (with $z_0 \triangleq 0$ and $\alpha_0 \triangleq 0$) as

$$
\begin{aligned}
z_i(\underline{x_i}, \hat{\theta}) &= x_i - \alpha_i, && (4.84)\\
\alpha_i(\underline{x_i}, \hat{\theta}) &= \, - z_{i-1} - k_i z_i - w_i^\top \hat{\theta} + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} x_{j+1}, && (4.85)\\
k_i(\underline{x_i}, \hat{\theta}) &= k_{iL} + k_{iw}|w_i|^2 + k_{ig}|g_i|^2 + \zeta_i && (4.86)\\
w_i(\underline{x_i}, \hat{\theta}) &= \phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j, && (4.87)\\
q_i^\top(\underline{x_i}, \hat{\theta}) &= \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}}, \quad i = 1, \ldots, n && (4.88)
\end{aligned}
$$

where $\zeta_i(\underline{x_i}, \hat{\theta}) > 0$ is a strictly positive nonlinear damping term to be defined. The actual control law is selected as

$$
u = \alpha_n(x, \hat{\theta}). \tag{4.89}
$$

In this subsection, we focus on the stabilization problem and therefore the excitation signal μ, discussed in Section 2.3, is omitted.

Remark 4.6. *Observing the recursive change of coordinates described by the equations*

(4.84) to (4.88) and noting that $\phi_i(0) = 0$, it is not difficult to see that there is a smooth and invertible relation between $\underline{x_i}$ and $\underline{z_i}$, thus $\underline{x_i} = 0 \Leftrightarrow \underline{z_i} = 0$. Moreover, $w_i(\underline{x_i}, \hat{\theta})$ satisfies the condition $w_i(0, \hat{\theta}) = 0$. Then, by Lemma A.3 and with a slight abuse of notation, one could write $w_i(\underline{z_i}, \hat{\theta}) = \bar{W}_i(\underline{z_i}, \hat{\theta})\underline{z_i}$, where $W_i : \mathbb{R}^i \times \mathbb{R}^q \to \mathbb{R}^{q \times i}$ is a smooth mapping.

Consider now the system (4.1) driven by the feedback control law (4.89), and express the closed-loop dynamics in terms of the error state z which yields the z-dynamics described by the equation

$$\dot{z} = Az + W^\top(\theta - \hat{\theta}) - G^\top \dot{\hat{\theta}}, \tag{4.90}$$

where[5]

$$A(z, \hat{\theta}) \triangleq -\operatorname{diag}(k_1, \ldots, k_n) + S - S^\top, \tag{4.91}$$

$$W(z, \hat{\theta}) \triangleq [w_1, \ldots, w_n], \tag{4.92}$$

$$G(z, \hat{\theta}) \triangleq [0, g_2, \ldots, g_n], \tag{4.93}$$

$A(z, \hat{\theta}) = -\operatorname{diag}(k_1, \ldots, k_n) + S - S^\top$, $W(z, \hat{\theta}) = [w_1, \ldots, w_n]$, $G(z, \hat{\theta}) = [0, g_2, \ldots, g_n]$. Compared with the x-dynamics (2.48) in the scalar case, equation (4.90), which describes the z-dynamics, has an additional input $\dot{\hat{\theta}}$ when $n \geq 2$, which is dominated by the $k_{ig}|g_i|^2$ term in (4.86). Similar to Lemma 2.1, ISS of the z-dynamics can be established as follows.

Lemma 4.4. *System* (4.90) *is ISS with respect to the inputs $\theta - \hat{\theta}$ and $\dot{\hat{\theta}}$.*

Proof. First note that $k_i > k_i - \zeta_i = k_{iL} + k_{iw}|w_i|^2 + k_{ig}|g_i|^2$, since $\zeta_i > 0$ by definition. Note that $k_i - \zeta_i$ is the damping term used in the classical scheme presented in [75]. The rest of the proof is identical to the proof of [75, Lemma 5.8], which is straight-forward and therefore omitted. □

[5] Note that there is a slight abuse of notation in (4.92) and (4.93) as we replace the arguments $(x, \hat{\theta})$ with $(z, \hat{\theta})$. It is not difficult to see that there is a $\hat{\theta}$-dependent smooth and invertible change of coordinates between $\underline{x_i}$ and $\underline{z_i}$, for $i = 1, \ldots, n$, and therefore any term depending on $(\underline{x_i}, \hat{\theta})$ can be equivalently expressed in term of $(\underline{z_i}, \hat{\theta})$.

4.3.2 Identifier Subsystems

Similar to Section 2.3, rewrite the z-dynamics with the congealed parameter ℓ_θ and the time-varying perturbation Δ_θ as

$$\dot{z} = Az + W^\top(\ell_\theta - \hat{\theta}) - G^\top\dot{\hat{\theta}} + W^\top\Delta_\theta, \tag{4.94}$$

implement the filters

$$\dot{\Omega}_0 = A\Omega_0 - W^\top\hat{\theta} - G^\top\dot{\hat{\theta}}, \tag{4.95}$$

$$\dot{\Omega}^\top = A\Omega^\top + W^\top, \tag{4.96}$$

and consider the parametric model

$$z = \Omega_0 + \Omega^\top\ell_\theta + \varepsilon. \tag{4.97}$$

The resulting parametrization error dynamics are then described by

$$\dot{\varepsilon} = \dot{z} - \dot{\Omega}_0 - \dot{\Omega}^\top\ell_\theta = A\varepsilon + W^\top\Delta_\theta. \tag{4.98}$$

Based on the parametric model (4.97) one computes the *certainty-equivalence* prediction $\hat{z} = \Omega_0 + \Omega^\top\hat{\theta}$ for z, which yields the prediction dynamics

$$\dot{\hat{z}} = \dot{\Omega}_0 + \dot{\Omega}^\top\hat{\theta} + \Omega^\top\dot{\hat{\theta}} = A\hat{z} + (\Omega - G)^\top\dot{\hat{\theta}}. \tag{4.99}$$

Defining the prediction error $\tilde{z} = z - \hat{z}$, another model parametrized in $(\ell_\theta - \hat{\theta})$ can be obtained, namely

$$\tilde{z} = \Omega^\top(\ell_\theta - \hat{\theta}) + \varepsilon. \tag{4.100}$$

Finally, select the normalized gradient parameter update law defined as

$$\dot{\hat{\theta}} = \Gamma\frac{\Omega\tilde{z}}{1 + \nu|\Omega|_{\mathrm{F}}^2} = \Gamma\frac{\Omega\tilde{z}_N}{\sqrt{1 + \nu|\Omega|_{\mathrm{F}}^2}}, \tag{4.101}$$

where $\tilde{z}_N$ is the normalized prediction error defined as $\tilde{z}_N \triangleq \frac{\tilde{z}}{m}$, $m(\Omega) = \sqrt{1 + \nu|\Omega|_{\mathrm{F}}^2} > 1$, $\Gamma = \Gamma^\top \succ 0$ is the adaptation gain, and $\nu > 0$ is a coefficient that adjusts the strength of the normalization. Up to this point, the dynamics of the subsystems of interest have been computed and we are now ready to proceed with the control synthesis.

4.3.3 Small-Gain-Like Synthesis

First we establish boundedness of Ω by considering the function $V_\Omega(\Omega) = \frac{1}{2}\mathrm{tr}(\Omega^\top\Omega)$ and its time derivative along the solutions of (4.96), that is,

$$\begin{aligned}
\dot{V}_\Omega &= \mathrm{tr}\Big(\frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{1}{2}\Omega\Omega^\top\big)\Big) = \mathrm{tr}\Big(\frac{1}{2}\Omega(A + A^\top)\Omega^\top\Big) + \mathrm{tr}(W^\top\Omega) \\
&\leq \; -\sum_{i=1}^{n} k_{iL}|(\Omega)_i|^2 - \sum_{i=1}^{n} k_{iw}|w_i|^2|(\Omega)_i|^2 + \sum_{i=1}^{n} w_i^\top(\Omega)_i \\
&\leq \; -k_L|\Omega|_{\mathrm{F}}^2 + \frac{1}{4k_W} = 2k_L V_\Omega + \frac{1}{4k_W},
\end{aligned} \tag{4.102}$$

where $k_L = \min_i k_{iL}$ and $k_W = (\sum_{i=1}^n k_{iw}^{-1})^{-1}$. Similar to the analysis in Section 2.3, this guarantees that Ω is bounded and $|\Omega|_F \leq \max\{|\Omega(0)|_{\mathrm{F}}, \frac{1}{2\sqrt{k_L k_W}}\} \triangleq \delta_\Omega$, where the constant δ_Ω can be computed from known constants, namely $|\Omega(0)|_{\mathrm{F}}$, k_L, and k_W.

For the next step, compute the dissipation inequalities of the $\hat{z}$, $\hat{\theta}$, and ε-subsystems. Consider the function $V_{\hat{z}}(\hat{z}) = \frac{1}{2}|\hat{z}|^2$. Taking the time derivative of $V_{\hat{z}}$ along the solutions of (4.99) yields

$$\begin{aligned}
\dot{V}_{\hat{z}} &= \frac{1}{2}\hat{z}^\top(A + A^\top)\hat{z} + \hat{z}^\top\frac{\Omega^\top\Gamma\Omega}{m}\tilde{z}_N - \hat{z}^\top\frac{G^\top\Gamma\Omega}{m}\tilde{z}_N \\
&\leq \; -\sum_{i=1}^{n} k_{iL}|\hat{z}_i|^2 - \sum_{i=2}^{n} k_{ig}|g_i|^2|\hat{z}_i|^2 + \frac{1}{2}k_L|\hat{z}|^2 \\
&\quad + \frac{\bar{\gamma}^2\delta_\Omega^2}{2\nu^2 k_L}|\tilde{z}_N|^2 + \sum_{i=1}^{n} k_{ig}|g_i|^2|\hat{z}_i|^2 + \frac{\bar{\gamma}^2}{4\nu^2 k_G}|\tilde{z}_N|^2 \\
&\leq \; -\frac{1}{2}k_L|\hat{z}|^2 + \frac{\bar{\gamma}^2}{2\nu^2}\left(\frac{\delta_\Omega^2}{k_L} + \frac{1}{2k_G}\right)|\tilde{z}_N|^2,
\end{aligned} \tag{4.103}$$

where $\bar{\gamma}$ is the largest eigenvalue of Γ and $k_G = (\sum_{i=2}^n k_{ig}^{-1})^{-1}$. One could see that the use of normalization reduces the order of δ_Ω if compared with (2.63). Consider now the function $V_{\hat\theta}(\hat\theta) = \frac{1}{2}|\ell_\theta - \hat\theta|^2_{\Gamma^{-1}}$ and its time derivative along (4.101), which yields

$$\begin{aligned}\dot V_{\hat\theta} =& -(\ell_\theta - \hat\theta)^\top \Gamma^{-1}\Gamma \frac{\Omega \tilde z_N}{m} = -(m\tilde z_N - \varepsilon)^\top \frac{\tilde z_N}{m} \\ \leq& -\frac{1}{2}|\tilde z_N|^2 + \frac{1}{2m^2}|\varepsilon|^2 \leq -\frac{1}{2}|\tilde z_N|^2 + \frac{1}{2}|\varepsilon|^2. \end{aligned} \tag{4.104}$$

For the ε-subsystem, consider the function $V_\varepsilon(\varepsilon) = \frac{1}{2}|\varepsilon|^2$, the time derivative of which along the trajectories of (4.98) is

$$\dot V_\varepsilon = \frac{1}{2}\varepsilon^\top (A + A^\top)\varepsilon + \varepsilon^\top W^\top \Delta_\theta. \tag{4.105}$$

Recalling Remark 4.6, and substituting $\bar W_i \underline{z_i}$ for w_i in W, yields

$$\begin{aligned}\varepsilon^\top W^\top \Delta_\theta = \varepsilon^\top \begin{bmatrix} z_1^\top \bar W_1^\top \\ \underline{z_2}^\top \bar W_2^\top \\ \vdots \\ \underline{z_{n-1}}^\top \bar W_{n-1}^\top \\ z^\top \bar W_n^\top \end{bmatrix} \Delta_\theta =& \sum_{i=1}^n \varepsilon_i \underline{z_i}^\top \bar W_i^\top \Delta_\theta \\ \leq& \frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}\sum_{i=1}^n |\bar W_i|^2_{\mathrm F}\varepsilon_i^2 + \frac{\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}\sum_{i=1}^n (|\underline{\hat z_i}|^2 + |\underline{\tilde z_i}|^2) \\ \leq& \frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}\sum_{i=1}^n |\bar W_i|^2_{\mathrm F}\varepsilon_i^2 + \frac{n\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}(|\hat z|^2 + |\tilde z|^2). \end{aligned} \tag{4.106}$$

Let the additional nonlinear damping terms be defined as

$$\zeta_i = \zeta_{iL} + \frac{\epsilon_{\Delta_\theta}\delta_{\Delta_\theta}}{2}|\bar W_i|^2_{\mathrm F}, \quad i = 1, \ldots, n, \tag{4.107}$$

where $\zeta_{iL} \geq \zeta_L$ and $\zeta_L > 0$ is a constant to be defined. This yields

$$\dot V_\varepsilon \leq -(k_L + \zeta_L)|\varepsilon|^2 + \frac{n\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}|\hat z|^2 + \frac{n\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}(1 + \nu\delta_\Omega^2)|\tilde z_N|^2. \tag{4.108}$$

With the obtained dissipation inequalities, the following lemma (as in Section 2.3) guarantees the existence of a feasible ζ_L.

Lemma 4.5. *Consider the dissipation inequalities* (4.103), (4.104) *and* (4.108)*, and the additional nonlinear damping terms* (4.107)*. For all given positive constants $a_{\hat z}$, $a_{\tilde z_N}$,* *and a_ε, there exist constant scaling coefficients $c_{\hat{z}}$, $c_{\hat{\theta}}$, c_ε, and a damping gain $\zeta_L > 0$, depending on $a_{(\cdot)}$, such that the function $V(\hat{z}, \hat{\theta}, \varepsilon) = c_{\hat{z}} V_{\hat{z}}(\hat{z}) + c_{\hat{\theta}} V_{\hat{\theta}}(\hat{\theta}) + c_\varepsilon V_\varepsilon(\varepsilon)$ satisfies the dissipation inequality*

$$\dot{V} \leq -a_{\hat{z}} |\hat{z}|^2 - a_{\tilde{z}_N} |\tilde{z}_N|^2 - a_\varepsilon |\varepsilon|^2. \tag{4.109}$$

Proof. The dissipation inequalities (4.103), (4.104), and (4.108) can be written as

$$\begin{bmatrix} \dot{V}_{\hat{z}} \\ \dot{V}_{\hat{\theta}} \\ \dot{V}_\varepsilon \end{bmatrix} \leq - \underbrace{\begin{bmatrix} \frac{1}{2}k_L & (E)_{12} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ (E)_{31} & (E)_{32} & k_L + \zeta_L \end{bmatrix}}_{\triangleq E} \underbrace{\begin{bmatrix} |\hat{z}|^2 \\ |\tilde{z}_N|^2 \\ |\varepsilon|^2 \end{bmatrix}}_{\triangleq \varphi}, \tag{4.110}$$

where $(E)_{12} = -\frac{\bar{\gamma}^2}{2\nu^2}\left(\frac{\delta_\Omega^2}{k_L} + \frac{1}{2k_G}\right)$, $(E)_{31} = -\frac{n\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}$, and $(E)_{32} = -\frac{n\delta_{\Delta_\theta}}{\epsilon_{\Delta_\theta}}(1 + \nu\delta_\Omega^2)$.

In the spirit of Section 3.1 we can obtain the underlying directed graph of E as illustrated in Fig. 4.3. Note that ζ_L can be regarded as an adjustable parameter, hence according to Definition 3.3, the ε-subsystem is an *active node*. It is not difficult to see that each directed cycle in the underlying directed graph of E contains the vertex associated with the ε-subsystem and therefore satisfies the condition of Theorem 3.3. As a result, there exists a sufficiently large $\zeta_L > 0$ such that the matrix E is a non-singular *M-matrix*. Then invoking Theorem 3.2 yields that for any given $a \triangleq [a_{\hat{z}}, a_{\tilde{z}_N}, a_\varepsilon]^\top > 0$, there exists a vector $c \triangleq [c_{\hat{z}}, c_{\hat{\theta}}, c_\varepsilon]^\top = (E^{-1})^\top a > 0$ such that

$$\dot{V} \leq -c^\top E \varphi = -a^\top \varphi, \tag{4.111}$$

which is (4.109) and completes the proof. Finally, ζ_L and the scaling coefficients can be computed using Proposition 3.2. □

Remark 4.7. *Although Lemma 2.2 and Lemma 4.5 are both used for establishing the small-gain-like property of the interconnected system, compared with Lemma 2.2, Lemma 4.5 is proven by exploiting the notion of active nodes, which does not require an explicit construction of the scaling coefficients and meanwhile, it guarantees the existence of the design parameter ζ_L and the scaling coefficients for an arbitrary dissipation*

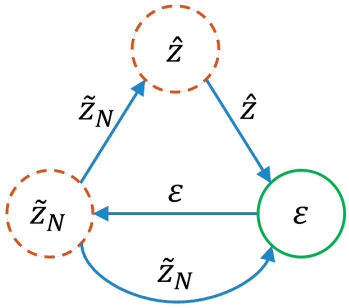

Figure 4.3: A schematic interpretation of the interconnected $\hat{z}$, $\tilde{z}_N$, and ε-subsystems. Note that this is also the underlying directed graph of the matrix E defined in (4.110) and has the same structure as the graph Fig. 2.3(b).

rate specified by $a_{(\cdot)}$, whereas in Lemma 2.2 the dissipation rate is fixed. This highlights the flexibility brought by the notion of active nodes in control synthesis.

The dissipation inequality (4.109) allows concluding boundedness and convergence properties for the closed-loop system, as stated hereafter.

Proposition 4.5. *Consider the closed-loop system consisting of the plant* (4.1) *and the adaptive controller described by the equations* (4.84) *to* (4.89)*,* (4.95)*,* (4.96)*,* (4.101)*, and with the additional nonlinear damping terms* (4.107)*. Then, all closed-loop signals are bounded and* $\lim_{t\to+\infty} x(t) = 0$*.*

Proof. We break the proof into two parts.

Boundedness. Using similar arguments as in the proof of Proposition 2.5 we can establish boundedness of Ω, $\hat{z}$, $\hat{\theta}$, ε, and $\tilde{z}$. Note that $\tilde{z}_N$ is also bounded due to boundedness of Ω and m, and this leads to boundedness of $\dot{\hat{\theta}}$. Invoking Lemma 4.4 and Assumption 1.1, we can conclude boundedness of z, and therefore boundedness of x, due to the smooth and invertible relation between x and z. Since both $W(z, \hat{\theta})$ and $Q(z, \hat{\theta})$ are smooth in z and $\hat{\theta}$ and both z and $\hat{\theta}$ are bounded, then Ω_0 is bounded, because the autonomous part of (4.95) is exponentially stable. Hence all closed-loop signals are bounded.

Convergence. First rewrite the dissipation inequality (4.109) as

$$\dot{V} \leq -a_{\hat{z}}|\hat{z}|^2 - \frac{a_{\tilde{z}_N}}{1+\nu\delta_\Omega^2}|\tilde{z}|^2 - a_\varepsilon|\varepsilon|^2, \tag{4.112}$$

and recall boundedness of z, $\hat{z}$, $\tilde{z}$, as well as boundedness of their time derivatives. By

Lemma A.5, we can conclude that $\lim_{t\to+\infty} z(t) = 0$. Finally, as pointed out in Remark 4.6, $x = 0 \Leftrightarrow z = 0$, hence it is clear that $\lim_{t\to+\infty} z(t) = 0 \Leftrightarrow \lim_{t\to+\infty} x(t) = 0$, which completes the proof. $\square$

Remark 4.8. *As long as M in* (4.110) *is a non-singular M-matrix, or equivalently, the condition*

$$k_L + \zeta_L > \frac{2(E)_{31}(E)_{12}}{k_L} - (E)_{32} \tag{4.113}$$

is satisfied, the scaling coefficients $c_{(\cdot)} > 0$ can be computed from any given constants $a_{(\cdot)}$ and E^{-1}. Due to the definitions of $(E)_{31}$ and $(E)_{32}$, when θ is constant, or equivalently, $\delta_{\Delta_\theta} = 0$, we have $(E)_{31} = (E)_{32} = 0$, and E is a non-singular M-matrix for any $k_L > 0$ and $\zeta_L = 0$. Then, by (4.107)*, we can set $\zeta_i = 0$, $i = 1, \ldots, n$, in which case the proposed scheme reduces to the classical scheme in [74] and [75], while still guaranteeing all the properties claimed in Proposition 4.5.*

Remark 4.9. *Proposition 4.5, in contrast to the integrated design in Proposition 4.1, provides a modular design, in the sense that the control law design is independent of the parameter update law design. The advantage of a modular design is that one can change the identifier (say, using a least-square identifier instead of the gradient-descent one* (4.101)*), without repeating the backstepping procedure for the control law design.*

4.4 Simulations

This section provides a numerical example for the identification-based adaptive scheme discussed in Section 4.3. Consider the nonlinear system described by the equations

$$\dot{x}_1 = \theta_1 x_1^2 + x_2, \quad \dot{x}_2 = \theta_2 x_1^2 + \theta_3 x_2^2 + u, \tag{4.114}$$

with the time-varying parameter vector $\theta(t) \triangleq [\theta_1(t), \theta_2(t), \theta_3(t)]^\top$ defined by

$$\theta(t) = \theta_c + \theta_{sw}(t) + \frac{W\varepsilon}{|W\varepsilon|}, \tag{4.115}$$

where $\theta_c = [1, 1, 1]^\top$ is the constant component of the parameter, and θ_{sw} consists of switching signals composed of three square waves of amplitudes equal to 1 and frequencies

0.5 Hz, 0.3 Hz, and 1 Hz, respectively. The last term $\frac{W\varepsilon}{|W\varepsilon|}$ in (4.115) can be treated as a unit vector controlled by a player who intends to maximize the supply rate in (4.105) and to destabilize the system by manipulating Δ_θ. Although ε cannot be implemented in the controller, it can be computed in the simulation by setting $\ell_\theta = \theta_c$ due to (4.97) and therefore can be used for testing robustness.

Consider three scenarios. In the "Baseline" scenario, we use constant parameters by setting $\theta = \theta_c$, and the classical identification-based backstepping controller proposed in [74]. In the "Controller 1" scenario, the parameter vector is set to (4.115) and the controller is the classical controller used in the "Baseline" scenario modified by a projection update law similar to the one used in [86], which confines the parameter vector within a ball centred at the origin and with a radius of 5. (note that $|\theta(t)| \leq 2\sqrt{3} + 1 < 5$, for all $t \geq 0$). In the "Controller 2" scenario, the proposed adaptive control scheme is implemented. In all scenarios, the initial state is set to $x(0) = [1, -1]^\top$, and all other state variables are initialized to 0. The design parameters are defined as $k_{1L} = k_{2L} = 1$, $k_{1w} = k_{2w} = k_{2q} = 5$, $\Gamma = I$, $\nu = 1$, $\epsilon_{\Delta_\theta} = 1$, $\delta_{\Delta_\theta} = 3$, and $\zeta_{1L} = \zeta_{2L} = 1$.

It is worth noting that W and z are state-dependent. Meanwhile, the performance comparison between the "Controller 1" scenario and the "Controller 2" scenario can only be fair if the two scenarios are driven by exactly the same time-varying parameters ("same" in the sense of time histories). One way to achieve this is to run the two scenarios "Controller 1" and "Controller 2" in parallel and run the simulation twice, collecting two sets of simulation data. In the first set, W and z are provided by the system controlled by Controller 1, and in the second set, by Controller 2.

As shown in Fig. 4.4 and Fig. 4.5, in the presence of time-varying parameters, the classical adaptive controller, namely Controller 1, still yields boundedness of the system state yet the transient performance is significantly degraded compared with the "Baseline" scenario. Oscillations at 0.5 Hz can be observed in both figures, which are triggered by the 0.5 Hz switching signal component in $\theta_1(t)$. The proposed Controller 2, in contrast, demonstrates robustness to parameter variations in both simulation sets and maintains the transient performance of the "Baseline" scenario.

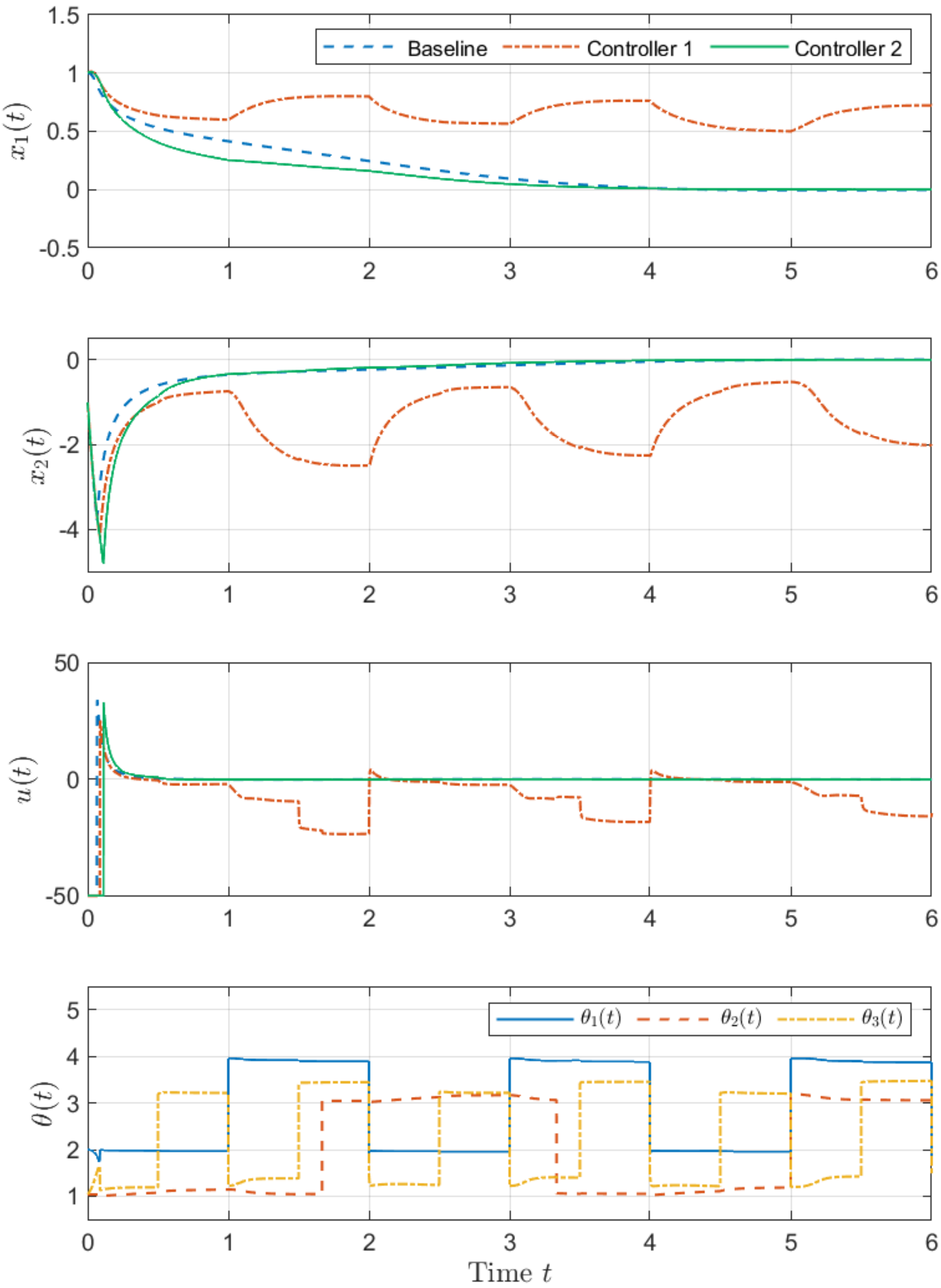

Figure 4.4: Simulation set 1: time histories of the system state, control effort, and state-dependent time-varying parameters for different controllers.

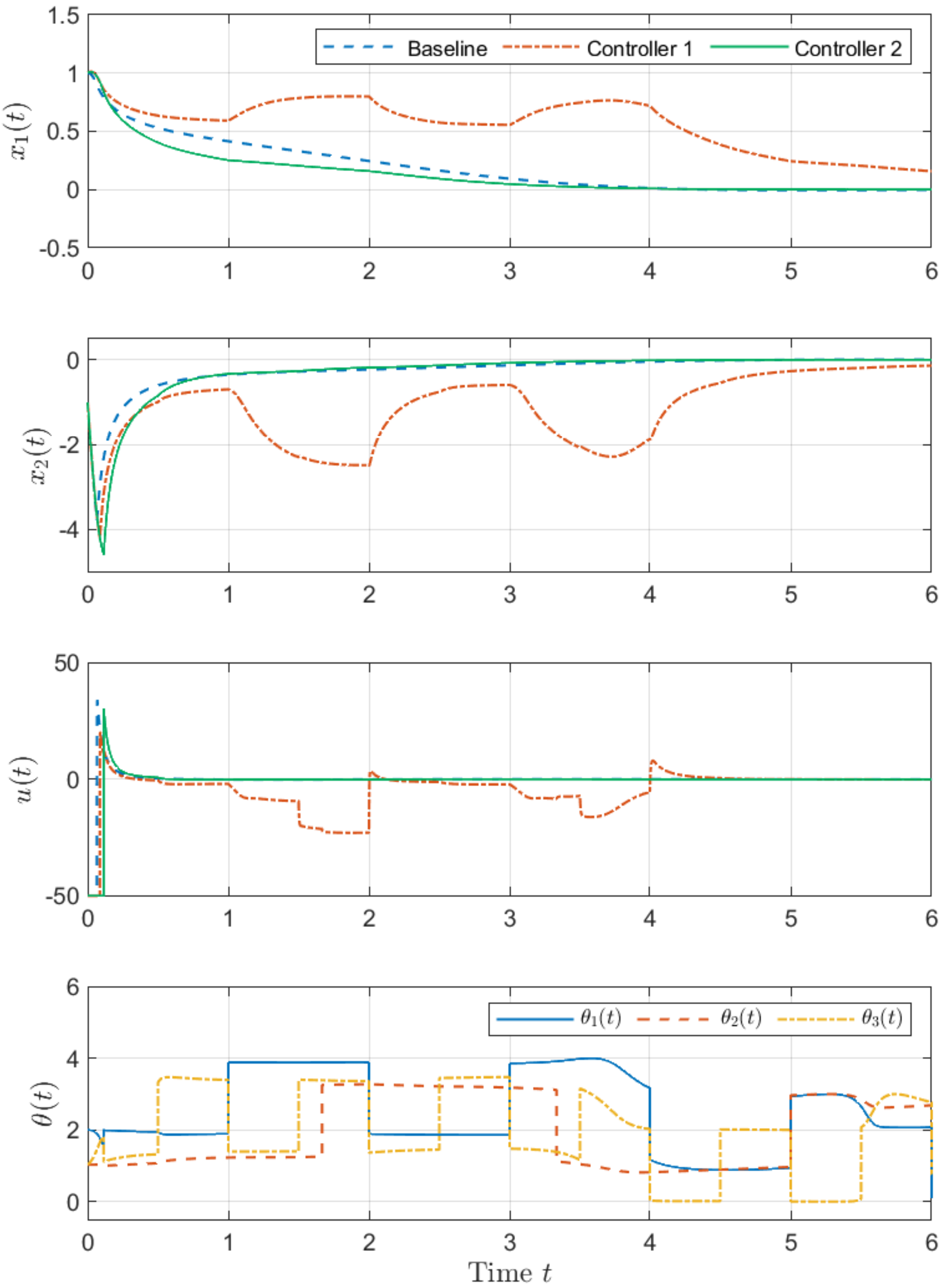

Figure 4.5: Simulation set 2: time histories of the system state, control effort, and state-dependent time-varying parameters for different controllers.

Chapter 5

Adaptive Regulation via Output Feedback

The methods discussed in Chapter 4 generalize the results of the *congelation of variables* method for scalar systems to multi-dimensional systems yet still require full-state feedback. Aside from the requirement of the signals that may not be measured in practice, the *lower triangular* structure puts an additional structural restriction. This is not a disquieting issue for systems with known parameters, as one can transform a given system into the *lower-triangular form* via a parameter-dependent change of coordinates. For systems with unknown parameters, this is more complicated as the new state variables obtained from the parameter-dependent change of coordinates are no longer known even if the old state variables are measured. It is in general not possible to obtain a global parameter-independent change of coordinates[1] that transforms a general nonlinear system into a system in *lower-triangular form*. Therefore, both the restrictions on the availability of signal measurement and on the system structure call for an output-feedback counterpart of the schemes discussed in Chapter 4. The output-feedback scheme, on one hand, only requires the output signal instead of all state variables to be measured and, on the other hand, does not require a parameter-dependent change of coordinates as in this case, only the input-output mapping matters and one can build up a reparametrized system with an equivalent input-output mapping using a set of filters with known state variables, for

[1] This is, however, possible if one either assumes some geometric conditions for the original system or simply seeks for a local parameter-independent change of coordinates instead of a global one. See [75, Section G.3] for a detailed discussion.

which the state-feedback schemes can be applied.

To this end, two output-feedback schemes for systems with time-varying parameters are presented in this chapter. One is the I&I scheme for linear single-input single-output (SISO) systems, and the other is the passivity-based scheme for a class of SISO nonlinear systems.

5.1 I&I Design for Linear SISO System

Consider an n-dimensional SISO linear system in observable canonical form with $n \geq 2$ and relative degree ρ, described by the equations

$$\begin{aligned}
\dot{x}_1 &= -a_1(t)x_1 + x_2, \\
&\vdots \\
\dot{x}_\rho &= -a_\rho(t)x_1 + x_{\rho+1} + b_\rho(t)u, \\
&\vdots \\
\dot{x}_n &= -a_n(t)x_1 + b_n(t)u, \\
y &= x_1,
\end{aligned} \tag{5.1}$$

or, in compact form, by the equations

$$\begin{aligned}
\dot{x} &= S_n x - a(t)y + \begin{bmatrix} 0_{(\rho-1)\times 1} \\ b(t) \end{bmatrix} u, \\
y &= e_1^\top x,
\end{aligned} \tag{5.2}$$

where $x(t) \in \mathbb{R}^n$ is the state vector; $u(t) \in \mathbb{R}$ is the input; $y(t) \in \mathbb{R}$ is the output; S_n is the $n \times n$ upper-shift matrix; and $e_1 = [1, 0, \ldots, 0]^\top \in \mathbb{R}^n$. The unknown time-varying parameters are denoted by the vector $\theta(t) = [b^\top(t), a^\top(t)]^\top$, $a(t) = [a_1(t), \ldots, a_n(t)]^\top \in \mathbb{R}^n$, $b(t) = [b_\rho(t), \ldots, b_n(t)]^\top \in \mathbb{R}^{n-\rho+1}$. Both b and a satisfy Assumption 1.2 and in addition, b_ρ satisfies Assumption 1.3.

5.1.1 System Reparametrization

As the formulation given by (5.1) is not in the *lower-triangular* form to which the backstepping techniques are applicable, one needs to reparametrize the original system by exploiting a set of input/output filters, similar to the ones adopted in [5] yet with special care for the time-varying parameters. In the spirit of the *congelation of variables* method the system parameters θ can be regarded as the sum of a vector of constant "congealed" parameters ℓ_θ and a vector of differences Δ_θ between the actual time-varying parameters and the "congealed" parameters, that is, $\theta = \ell_\theta + \Delta_\theta$. To avoid introducing $\dot{\theta}$-related terms in the analysis, consider first the system parametrized by the constant parameters $\ell_\theta \triangleq [\ell_{b_\rho}, \dots, \ell_{b_n}, \ell_{a_1}, \dots, \ell_{a_n}]^\top$. The input-output relation of the system parametrized by ℓ_θ is described by the differential equation

$$y^{(n)} = [u^{(n-\rho)}, u^{(n-\rho-1)}, \dots, u, -y^{(n-1)}, -y^{(n-2)}, \dots, -y]\ell_\theta. \tag{5.3}$$

Since the time derivatives of u and y are not directly measured, one needs to apply a stable filter $\Lambda(s) = s^{n-1} + \lambda_{n-1}s^{n-2} + \dots + \lambda_2 s + \lambda_1$ to both sides of (5.3). This yields

$$\frac{s^{n-1}}{\Lambda(s)}[\dot{y}] = \left[\frac{s^{n-\rho}}{\Lambda(s)}[u], \dots, \frac{1}{\Lambda(s)}[u], \frac{-s^{n-1}}{\Lambda(s)}[y], \dots, \frac{-1}{\Lambda(s)}[y]\right]\ell_\theta. \tag{5.4}$$

Note now that $\dfrac{s^{n-1}}{\Lambda(s)} = 1 - \dfrac{\lambda_{n-1}s^{n-2} + \dots + \lambda_2 s + \lambda_1}{\Lambda(s)}$, which yields

$$\begin{aligned}\dot{y} = {} & s\frac{\lambda_{n-1}s^{n-2} + \dots + \lambda_2 s + \lambda_1}{\Lambda(s)}[y] \\ & + \left[\frac{s^{n-\rho}}{\Lambda(s)}[u], \dots, \frac{1}{\Lambda(s)}[u], \frac{-s^{n-1}}{\Lambda(s)}[y], \dots, \frac{-1}{\Lambda(s)}[y]\right]\ell_\theta.\end{aligned} \tag{5.5}$$

Consider the state-space realization of the filters described by

$$\begin{aligned}\dot{\zeta} &= A_\lambda \zeta + e_{n-1}u, \\ \dot{\xi} &= A_\lambda \xi - e_{n-1}y,\end{aligned} \tag{5.6}$$

where $\zeta(t) \in \mathbb{R}^{n-1}$ and $\xi(t) \in \mathbb{R}^{n-1}$ are the filter states; $A_\lambda = S_{n-1} - e_{n-1}\lambda^\top$ is Hurwitz; S_{n-1} is the $(n-1) \times (n-1)$ upper-shift matrix; $e_{n-1} = [0, \dots, 0, 1]^\top \in \mathbb{R}^{n-1}$; and

$\lambda \triangleq [\lambda_1, \ldots, \lambda_{n-1}]^\top$ is the vector of filter gains. Equation (5.5) can then be written as

$$\begin{aligned}\dot{y} &= -\lambda^\top \dot{\xi} + \phi^\top \ell_\theta + \eta_0 \\ &= -\lambda^\top (A_\lambda \xi - e_{n-1} y) + \phi^\top \ell_\theta + \eta_0,\end{aligned} \tag{5.7}$$

where $\phi \triangleq [\zeta_{n-\rho+1}, \ldots, \zeta_1, -\lambda^\top \xi - y, \xi_{n-1}, \ldots, \xi_1]^\top$ and $\zeta_n \triangleq -\lambda^\top \zeta + u$. This is also the parametrization used in [5, Section 4.4.1]. For linear time-invariant systems considered in classical adaptive schemes, η_0 is an exponentially decaying error term because of the Hurwitz property of A_λ and it is typically ignored in analysis and design. However, this is not the case when the system parameters are time-varying. To see this, rearranging (5.7) yields

$$\begin{aligned}\eta_0 = & -(\ell_{a_1} + \Delta_{a_1}) y + x_2 + \lambda^\top (A_\lambda \xi - e_{n-1} y) \\ & - [\zeta_{n-\rho+1}, \ldots, \zeta_1, -\lambda^\top \xi - y, \xi_{n-1}, \ldots, \xi_1] \ell_\theta \\ = & \, \lambda^\top A_\lambda \xi - \lambda^\top e_{n-1} y + x_2 - \Delta_{a_1} y \\ & - [\zeta_{n-\rho+1}, \ldots, \zeta_1, -\lambda^\top \xi, \xi_{n-1}, \ldots, \xi_1] \ell_\theta.\end{aligned} \tag{5.8}$$

Define $\eta_1 = \eta_0 + \Delta_{a_1} y$ to separate the perturbation term $-\Delta_{a_1} y$ from the expression of η_1: this allows avoiding differentiating unknown time-varying parameters when deriving the error dynamics and yields

$$\begin{aligned}\dot{\eta}_1 = & \, \lambda^\top A_\lambda^2 \xi - \lambda^\top A_\lambda e_{n-1} y - \lambda^\top e_{n-1} x_2 + x_3 \\ & - [\zeta_{n-\rho+2}, \ldots, \zeta_2, -\lambda^\top A_\lambda \xi, -\lambda^\top \xi, \xi_{n-1} \ldots, \xi_2] \ell_\theta \\ & + (\lambda^\top e_{n-1} \Delta_{a_1} - \Delta_{a_2}) y.\end{aligned} \tag{5.9}$$

Define $\eta_2 = \dot{\eta}_1 + (-\lambda^\top e_{n-1} \Delta_{a_1} + \Delta_{a_2}) y$, which yields

$$\begin{aligned}\dot{\eta}_2 = & \, \lambda^\top A_\lambda^3 \xi - \lambda^\top A_\lambda^2 e_{n-1} y - \cdots - \lambda^\top e_{n-1} x_3 + x_4 \\ & - [\zeta_{n-\rho+3}, \ldots, \zeta_3, -\lambda^\top A_\lambda^2 \xi, -\lambda^\top A_\lambda \xi, -\lambda^\top \xi, \xi_{n-1} \ldots, \xi_3] \ell_\theta \\ & - (-\lambda^\top A_\lambda e_{n-1} \Delta_{a_1} - \lambda^\top e_{n-1} \Delta_{a_2} + \Delta_{a_3}) y.\end{aligned} \tag{5.10}$$

Repeating the procedures above for $\eta_3, \ldots, \eta_{\rho-2}$ and then defining $\eta_{\rho-1} = \dot{\eta}_{\rho-2} + (-\lambda^\top A_\lambda^{\rho-3} e_{n-1}\Delta_{a_1} - \cdots - \lambda^\top e_{n-1}\Delta_{a_{\rho-2}} + \Delta_{a_{\rho-1}})y$ yields

$$\begin{aligned}
\dot{\eta}_{\rho-1} = {} & \lambda^\top A_\lambda^\rho \xi - \lambda^\top A_\lambda^{\rho-1} e_{n-1} y - \cdots - \lambda^\top e_{n-1} x_\rho + x_{\rho+1} \\
& - [-\lambda^\top \zeta, \ldots, \zeta_\rho, -\lambda^\top A_\lambda^{\rho-1}\xi, \ldots, \xi_\rho]\ell_\theta \\
& - (-\lambda^\top A_\lambda^{\rho-2} e_{n-1}\Delta_{a_1} - \cdots - \lambda^\top e_{n-1}\Delta_{a_{\rho-1}} + \Delta_{a_\rho})y \\
& + \Delta_{b_\rho} u. \qquad (5.11)
\end{aligned}$$

Repeating again the procedures for $\eta_\rho \ldots, \eta_{n-2}$ and, finally, defining $\eta_{n-1} = \dot{\eta}_{n-2} + (-\lambda^\top A_\lambda^{n-3} e_{n-1}\Delta_{a_1} - \cdots - \lambda^\top e_{n-1}\Delta_{a_{n-2}} + \Delta_{a_{n-1}})y - (-\lambda^\top A_\lambda^{n-\rho-2} e_{n-1}\Delta_{b_\rho} - \cdots - \lambda^\top e_{n-1}\Delta_{b_{n-2}} + \Delta_{b_{n-1}})u$ yields

$$\begin{aligned}
\dot{\eta}_{n-1} = {} & \lambda^\top A_\lambda^n \xi - \lambda^\top A_\lambda^{n-1} e_{n-1} y - \cdots - \lambda^\top e_{n-1} x_n - \\
& [-\lambda^\top A_\lambda^{n-\rho}\zeta, \ldots, -\lambda^\top \zeta, -\lambda^\top A_\lambda^{n-1}\xi, \ldots, -\lambda^\top \xi]\ell_\theta \\
& - (-\lambda^\top A_\lambda^{n-2} e_{n-1}\Delta_{a_1} - \cdots - \lambda^\top e_{n-1}\Delta_{a_{n-1}} + \Delta_{a_n})y \\
& + (-\lambda^\top A_\lambda^{n-\rho-1} e_{n-1}\Delta_{b_\rho} \cdots - \lambda^\top e_{n-1}\Delta_{b_{n-1}} + \Delta_{b_n})u. \qquad (5.12)
\end{aligned}$$

Theorem 5.1. *The dynamics of the reparametrization error η_0 are described by the equations*

$$\dot{\eta} = A_\lambda \eta + B_\lambda \left(\begin{bmatrix} 0_{(\rho-1)\times 1} \\ \Delta_b(t) \end{bmatrix} u - \Delta_a(t) y \right), \qquad (5.13)$$

$$\eta_0 = e_1^\top \eta - \Delta_{a_1}(t) y, \qquad (5.14)$$

where

$$B_\lambda \triangleq \begin{bmatrix} \lambda^\top e_{n-1} & 1 & 0 & \cdots & 0 \\ \lambda^\top A_\lambda e_{n-1} & \lambda^\top e_{n-1} & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \lambda^\top A_\lambda^{n-2} e_{n-1} & \cdots & \cdots & \lambda^\top e_{n-1} & 1 \end{bmatrix}, \qquad (5.15)$$

$\eta \triangleq [\eta_1, \ldots, \eta_{n-1}]^\top$ and

$$\begin{aligned}\eta_i \triangleq \lambda^\top A_\lambda^i \xi - \lambda^\top \sum_{j=1}^{i} A_\lambda^{i-j} e_{n-1} x_j + x_{i+1} \\ - e_1^\top [A_\lambda^{i+n-\rho-1}\zeta, \ldots, A_\lambda^{i-1}\zeta, A_\lambda^{i+n-2}\xi, \ldots, A_\lambda^{i-1}\xi]\ell_\theta,\end{aligned} \tag{5.16}$$

for $i = 1, \ldots, n-1$.

Proof. We first prove the classical part of the result, that is, the case in which $\Delta_a(t) = 0$ and $\Delta_b(t) = 0$. for all $t \geq 0$. In this case the equations (5.13) and (5.14) reduce to $\dot{\eta} = A_\lambda \eta$ and $\eta_0 = \eta_1$, respectively. It should be noted that the definitions of the η_i's in (5.16) are the same as the ones previously defined from η_0 and $\dot{\eta}_1, \ldots, \dot{\eta}_{n-2}$, and therefore one can directly conclude that $\dot{\eta}_1 = \eta_2$, $\dot{\eta}_2 = \eta_3$, ..., $\dot{\eta}_{n-2} = \eta_{n-1}$. It only remains to prove that

$$\dot{\eta}_{n-1} = -\lambda_1 \eta_1 - \cdots - \lambda_{n-1}\eta_{n-1}. \tag{5.17}$$

Note that $e_1^\top A_\lambda^{n-1} = -\lambda^\top$, and that when $\Delta_a = 0$ and $\Delta_b = 0$ the equation

$$\begin{aligned}\dot{\eta}_{n-1} &= \lambda^\top A_\lambda^n \xi - \lambda^\top \sum_{j=1}^{n} A_\lambda^{n-j} e_{n-1} x_j \\ &\quad - e_1^\top [A_\lambda^{(2n-\rho-1)}\zeta, \ldots, A_\lambda^{n-1}\zeta, A_\lambda^{2n-2}\xi, \ldots, A_\lambda^{n-1}\xi]\ell_\theta \\ &= e_1^\top A_\lambda^{n-1} \underbrace{\Big(- A_\lambda^n \xi + \sum_{j=1}^{n} A_\lambda^{n-j} e_{n-1} x_j - [A_\lambda^{n-\rho}\zeta, \ldots, \zeta, A_\lambda^{n-1}\xi, \ldots, \xi]\ell_\theta \Big)}_{\triangleq Y}\end{aligned} \tag{5.18}$$

holds. Note also that $e_1^\top A_\lambda^{n-2} e_{n-1} = 1$, and that one can write the standalone x_{i+1}-term in (5.16) as $e_1^\top A_\lambda^{i-1} A_\lambda^{n-i-1} e_{n-1} x_{i+1}$. Thus, one can rewrite (5.16) as

$$\eta_i = e_1^\top A_\lambda^{i-1} \Big(- A_\lambda^n \xi + \sum_{j=1}^{i+1} A_\lambda^{n-j} e_{n-1} x_j - [A_\lambda^{n-\rho}\zeta, \ldots, \zeta, A_\lambda^{n-1}\xi, \ldots, \xi]\ell_\theta \Big). \tag{5.19}$$

Note now that $e_1^\top A_\lambda^{i-1} A_\lambda^{n-j} e_{n-1} = 0$, for $j = i+2, \ldots, n$. This yields

$$\eta_i = e_1^\top A_\lambda^{i-1} Y. \tag{5.20}$$

Applying the Cayley-Hamilton theorem to (5.18) yields

$$\begin{aligned}\dot{\eta}_{n-1} &= e_1^\top A_\lambda^{n-1} Y = e_1^\top(-\lambda_1 I_{n-1} - \cdots - \lambda_{n-1} A_\lambda^{n-2})Y \\ &= -\lambda_1\eta_1 - \cdots - \lambda_{n-1}\eta_{n-1}.\end{aligned} \tag{5.21}$$

Thus we have proven the classical part of the theorem.

Consider now the time-varying part of the proof by observing the equations (5.9)–(5.12). It is not difficult to see that the time-varying perturbations, namely the Δ_a-terms and the Δ_b-terms, are added to the η-dynamics exactly in the way described by the structure of B_λ. Hence the proof of the theorem is complete. □

Remark 5.1. *The filters* (5.6) *are reduced-order filters (containing $n-1$, instead of n, state variables each) because in the regressor ϕ in* (5.7)*, ξ_n, the additional filter state variable in the full-order case, is replaced by $-\sum_{i=1}^{n-1}\lambda_i\xi_i - y$, exploiting the fact that y is measured.*

As is shown by Theorem 5.1, when θ is time-varying the perturbation terms coupled with y and with u appear in the dynamics of η, which is a result of the substitution of ℓ_θ for θ. Because of these perturbation terms coupled with y and with u, we cannot directly determine the stability properties of the η-dynamics by exploiting the Hurwitz property of A_λ, like in classical adaptive control schemes, whereas the convergence of η_0 is necessary to make the reparametrized model (5.7) valid. The perturbations coupled with y can be dominated by a strengthened filter and controller design, as shown in [16]; however, the results therein merely circumvent the perturbations coupled with u by assuming a constant vector of input coefficients b, i.e. $\Delta_b(t) = 0$, $\forall t \geq 0$. This restriction can be removed by exploiting a minimum-phase property, which merges the perturbations coupled with u into the perturbations coupled with y.

5.1.2 Inverse Dynamics

Consider the inverse dynamics of system (5.1), *i.e.* considering the system as driven by y and its time-derivatives instead of u, which yields

$$\begin{aligned} x_2 &= y^{(1)} + a_1 y, \\ &\vdots \\ x_\rho &= y^{(\rho-1)} + (a_1 y)^{(\rho-2)} + \cdots + a_{\rho-1} y. \end{aligned} \tag{5.22}$$

Setting $u = \frac{1}{b_\rho}(-x_{\rho+1} + y^{(\rho)} + (a_1 y)^{(\rho-1)} + \cdots + a_\rho y)$ yields

$$\begin{aligned} \dot{x}_{\rho+1} &= -\frac{b_{\rho+1}}{b_\rho} x_{\rho+1} + x_{\rho+2} + \frac{b_{\rho+1}}{b_\rho}\Big(y^{(\rho)} + (a_1 y)^{(\rho-1)} + \cdots + a_\rho y\Big), \\ &\vdots \\ \dot{x}_n &= -\frac{b_n}{b_\rho} x_{\rho+1} + \frac{b_n}{b_\rho}\Big(y^{(\rho)} + (a_1 y)^{(\rho-1)} + \cdots + a_\rho y\Big). \end{aligned} \tag{5.23}$$

We now perform a change of coordinates to eliminate the time derivatives of y, which are not desirable in the design and the analysis. To this end, note that the identity[2]

$$s_1 s_2^{(i)} = (-1)^i s_1^{(i)} s_2 + \left(\sum_{j=0}^{i-1} (-1)^j s_1^{(j)} s_2^{(i-1-j)}\right)^{(1)} \tag{5.24}$$

holds for any pair of smooth signals s_1 and s_2. Using (5.24) and defining the new coordinate

$$\begin{aligned} \bar{x}_n = x_n &- \sum_{j=0}^{\rho-1} (-1)^j \left(\frac{b_n}{b_\rho}\right)^{(j)} y^{(\rho-1-j)} \\ &- \sum_{i=1}^{\rho-1} \sum_{j=0}^{\rho-i-1} (-1)^{(j)} \left(\frac{b_n}{b_\rho}\right)^{(j)} (a_i y)^{(\rho-i-1-j)} \end{aligned} \tag{5.25}$$

yields

$$\dot{\bar{x}}_n = -\frac{b_n}{b_\rho} x_{\rho+1} - a_n y + (-1)^\rho \left(\frac{b_n}{b_\rho}\right)^{(\rho)} y + \sum_{i=1}^{\rho} (-1)^{\rho-i} \left(\frac{b_n}{b_\rho}\right)^{(\rho-i)} a_i y, \tag{5.26}$$

[2] This identity is derived by recursively applying the Leibniz product rule for differentiation. The notation $(\cdot)^{(i)}$ is used to denote the ith time derivative, assuming it exists.

Algorithm 5.1 Change of coordinates $x_{\rho+1}, \dots, x_n$.

Input: $x_{\rho+1}, \dots, x_n, \dot{x}_{\rho+1}, \dots, \dot{x}_n$.
Output: $\bar{x}_{\rho+1}, \dots, \bar{x}_n, \dot{\bar{x}}_{\rho+1}, \dots, \dot{\bar{x}}_n$.

while time derivatives of y appear in the expression of $\dot{x}_{\rho+1}, \dots, \dot{x}_n$ **do** ▷ This while-loop iterates for ρ times as it reduces the order of $y^{(\rho)}$ by one each iteration.
  for $i = n \to \rho + 2$ **do**
    Update $\bar{x}_i$ and $\dot{\bar{x}}_i$ using (5.24).
    Rewrite x_i in terms of $\bar{x}_i$ in the expression of $\dot{x}_{i-1}$ and leave the feedback term $-\frac{b_i}{b_\rho} x_{\rho+1}$ unchanged.
  end for
  Update $\bar{x}_{\rho+1}$ and $\dot{\bar{x}}_{\rho+1}$ using (5.24).
  Rewrite $x_{\rho+1}$ in terms of $\bar{x}_{\rho+1}$ in the expressions of $\dot{\bar{x}}_{\rho+1}, \dots, \dot{\bar{x}}_n$, respectively. ▷ This brings back the time derivatives of y, but with the order reduced by one.
  $x_{\rho+1} \leftarrow \bar{x}_{\rho+1}, \ \dots, \ x_n \leftarrow \bar{x}_n, \ \dot{x}_{\rho+1} \leftarrow \dot{\bar{x}}_{\rho+1}, \ \dots, \ \dot{x}_n \leftarrow \dot{\bar{x}}_n$. ▷ Update the old coordinates before the next iteration.
end while

which does not contain time derivatives of y. In the spirit of the above procedure, we proceed with the change of coordinates using Algorithm 5.1, which yields the inverse dynamics in the new coordinates described by the equations

$$\dot{\bar{x}} = A_{\bar{b}}(t)\bar{x} + b_y(t)y, \tag{5.27}$$

$$u = \frac{1}{b_\rho(t)} \left(-\bar{x}_{\rho+1} + y^{(\rho)} + \sum_{j=0}^{\rho-1} a_{y^{(j)}}(t) y^{(j)} \right), \tag{5.28}$$

where $\bar{x}(t) = [\bar{x}_{\rho+1}, \dots, \bar{x}_n]^\top \in \mathbb{R}^{n-\rho}$, $A_{\bar{b}} = S_{n-\rho} - \bar{b} e_1^\top$, $\bar{b}(t) = [\frac{b_{\rho+1}}{b_\rho}, \dots, \frac{b_n}{b_\rho}]^\top \in \mathbb{R}^{n-\rho}$, and $S_{n-\rho}$ is the $(n-\rho) \times (n-\rho)$ upper-shift matrix. b_y, $a_{y^{(i)}}$ are unknown due to the unknown θ, but bounded for all $t \geq 0$ due to Assumption 1.2.

Assumption 5.1 (Strong minimum-phase property)**.** *System* (5.1) *has a strong minimum-phase property in the sense that the inverse dynamics* (5.27) *are input-to-state stable (ISS) with respect to the input y. Moreover, there exists an ISS Lyapunov function $V_{\bar{x}}(\bar{x}) \triangleq \bar{x}^\top P_{\bar{x}} \bar{x}$, with a constant $P_{\bar{x}} = P_{\bar{x}}^\top \succ 0$ and the time derivative of $V_{\bar{x}}$ along the trajectories of the inverse dynamics satisfies the inequality*

$$\dot{V}_{\bar{x}} \leq -\bar{x}^\top \bar{x} + \epsilon_{b_y}^2 \delta_{b_y}^2 y^2, \tag{5.29}$$

where $\epsilon_{b_y} > 0$ is constant and $\delta_{b_y} = \sup_{t \geq 0} |b_y(t)|$.

Remark 5.2. *Due to the linearity of the inverse dynamics* (5.27)*, the exponential stability of the origin of the zero dynamics (which is* (5.27) *with* $y(t) = 0$ *for all* $t \geq 0$*) is equivalent to the ISS of the inverse dynamics. For linear time-invariant systems the ISS property can be equivalently replaced by the condition that* $A_{\ell_{\bar{b}}}$ *is Hurwitz, which is the typical assumption made in classical adaptive control schemes.*

5.1.3 Filter Design

In the previous subsection we have shown that the control input u can be equivalently written in terms of the inverse dynamics state variable $\bar{x}_{\rho+1}$ and the time derivatives of y of order up to ρ. Since the time derivatives of y is not desirable in the design, in this subsection we continue to exploit the low-pass characteristics of the dynamics of the reparametrization error η to eliminate the time derivatives of y. Substituting (5.28) into (5.13) yields

$$\dot{\eta} = A_\lambda \eta - B_\lambda \Delta_a(t) y + \bar{\Delta}_b(t) \left(y^{(\rho)} + \sum_{j=0}^{\rho-1} a_{y^{(j)}}(t) y^{(j)} - \bar{x}_{\rho+1} \right), \tag{5.30}$$

where $\bar{\Delta}_b(t) = B_\lambda \begin{bmatrix} 0_{(\rho-1)\times 1} \\ \Delta_b(t) \end{bmatrix} \dfrac{1}{b_\rho(t)}$, which is unknown, but bounded due to Assumption 1.1. It should be noted that due to the structure of B_λ, the first $\bar{\rho} - 1$ elements of $\bar{\Delta}_b(t)$ are 0, or equivalently, u is separated from η_1 (or η_0) by $\bar{\rho}$ integrators, where $\bar{\rho}$ is defined by

$$\bar{\rho} = \begin{cases} 1, & \text{if } \rho = 1; \\ \rho - 1, & \text{if } \rho \geq 2 \text{ and } \Delta_{b_\rho} \neq 0; \\ \rho, & \text{if } \rho \geq 2 \text{ and } \Delta_{b_\rho} = 0. \end{cases} \tag{5.31}$$

To guarantee that no time derivative of y appears in η_0, at least ρ integrators between u and η_1 are required, as u contains $y^{(\rho)}$. According to (5.31), the assumption that follows is introduced to deal with this requirement.

Assumption 5.2. *The relative degree of system* (5.1) *is either* $\rho = 1$ *or* $\rho \geq 2$*, in which case* b_ρ *is constant (that is,* $\Delta_{b_\rho}(t) = 0$, $\forall t \geq 0$*).*

Remark 5.3. *The restriction imposed by Assumption 5.2 is only related to the relative*

Algorithm 5.2 Change of coordinates $\eta_1, \ldots, \eta_{n-1}$.

Input: $\eta_1, \ldots, \eta_{n-1}, \dot{\eta}_1, \ldots, \dot{\eta}_{n-1}$.
Output: $\bar{\eta}_1, \ldots, \bar{\eta}_{n-1}, \dot{\bar{\eta}}_1, \ldots, \dot{\bar{\eta}}_{n-1}$.

1: **while** time derivatives of y appear in the expression of $\dot{\eta}_1, \ldots, \dot{\eta}_{n-1}$ **do** ▷ This while-loop should only iterate for once if Assumption 5.2 is satisfied.
2: **for** $i = n-1 \to 2$ **do**
3: Update $\bar{\eta}_i$ and $\dot{\bar{\eta}}_i$ using (5.24).
4: Rewrite η_i in terms of $\bar{\eta}_i$ in the expression of $\dot{\eta}_{i-1}$.
5: **end for**
6: Update $\bar{\eta}_1$ and $\dot{\bar{\eta}}_1$ using (5.24).
7: Rewrite $\eta_1, \ldots, \eta_{n-1}$ in terms of $\bar{\eta}_1, \ldots, \bar{\eta}_{n-1}$, respectively, in the expressions of $\dot{\bar{\eta}}_{n-1}$. ▷ This should not bring back any time derivatives of y if Assumption 5.2 is satisfied.
8: $\eta_1 \leftarrow \bar{\eta}_1, \ldots, \eta_{n-1} \leftarrow \bar{\eta}_{n-1}, \dot{\eta}_1 \leftarrow \dot{\bar{\eta}}_1, \ldots, \dot{\eta}_{n-1} \leftarrow \dot{\bar{\eta}}_{n-1}$. ▷ Update the old coordinates before the next iteration.
9: **end while**

degree of the original system (5.1) *and to* $\Delta_b(t)$ *yet it is independent of whether we use reduced-order filters or full-order filters (as implemented in [18]). Using filters of different order can only provide a different non-minimal realization of the system while cannot change the relative degree.*

Similar to what implemented in the previous subsection, we use a change of coordinates to eliminate the time derivatives of y. Applying Algorithm 5.2 yields the dynamics (5.30) in the new coordinates, namely

$$\dot{\bar{\eta}} = A_\lambda \bar{\eta} + \bar{b}_y(t)y - \bar{\Delta}_b(t)\bar{x}_{\rho+1}, \tag{5.32}$$

$$\eta_0 = \bar{\eta}_1 + \bar{a}_y(t)y, \tag{5.33}$$

where, by Assumption 1.2, $\bar{b}_y$ and $\bar{a}_y$ are unknown but bounded.

Proposition 5.1. *The reparametrization error subsystem* (5.32) *is ISS with respect to the inputs* $\bar{x}_{\rho+1}$ *and* y *if the vector of filter gains is given by* $\lambda \triangleq \frac{1}{2}e_{n-1}^\top P_{\bar{\eta}}$, *where* $P_{\bar{\eta}} = P_{\bar{\eta}}^\top \succ 0$ *satisfies the algebraic Riccati inequality*

$$S_{n-1}^\top P_{\bar{\eta}} + P_{\bar{\eta}} S_{n-1} - P_{\bar{\eta}} e_{n-1} e_{n-1}^\top P_{\bar{\eta}} + Q_{\bar{\eta}} \preceq 0, \tag{5.34}$$

with

$$Q_{\bar{\eta}} = \left(\frac{1}{\epsilon^2_{P_{\bar{\eta}}\bar{\Delta}_b}} + \frac{1}{\epsilon^2_{P_{\bar{\eta}}\bar{b}_y}} + 1\right) I_{n-1}, \tag{5.35}$$

$\epsilon_{P_{\bar{\eta}}\bar{\Delta}_b} > 0$ *and* $\epsilon_{P_{\bar{\eta}}\bar{b}_y} > 0$. *Moreover, there exists an ISS Lyapunov function* $V_{\bar{\eta}}(\bar{\eta}) \triangleq \bar{\eta}^\top P_{\bar{\eta}}\bar{\eta}$ *and the time derivative of* $V_{\bar{\eta}}$ *along the trajectories of the reparametrization error dynamics* (5.32) *satisfies the inequality*

$$\dot{V}_{\bar{\eta}} \leq \; -\bar{\eta}^\top\bar{\eta} + \epsilon^2_{P_{\bar{\eta}}\bar{b}_y}\delta^2_{P_{\bar{\eta}}\bar{b}_y}y^2 + \epsilon^2_{P_{\bar{\eta}}\bar{\Delta}_b}\delta^2_{P_{\bar{\eta}}\bar{\Delta}_b}\bar{x}^2_{\rho+1}, \tag{5.36}$$

where $\delta_{P_{\bar{\eta}}\bar{b}_y} = \sup_{t\geq 0}|P_{\bar{\eta}}\bar{b}_y(t)|$ *and* $\delta_{P_{\bar{\eta}}\bar{\Delta}_b} = \sup_{t\geq 0}|P_{\bar{\eta}}\bar{\Delta}_b(t)|$.

Proof. Taking the time derivative of $V_{\bar{\eta}} = \bar{\eta}^\top P_{\bar{\eta}}\bar{\eta}$ along the trajectories of the system (5.32) and recalling that $A_\lambda = S_{n-1} - e_{n-1}\lambda^\top$ yields

$$\begin{aligned}
\dot{V}_{\bar{\eta}} &= 2\bar{\eta}^\top P_{\bar{\eta}}(A_\lambda\bar{\eta} + \bar{b}_y y - \bar{\Delta}_b\bar{x}_2)\\
&\leq \bar{\eta}^\top(S^\top_{n-1}P_{\bar{\eta}} + P_{\bar{\eta}}S_{n-1} - P_{\bar{\eta}}e_{n-1}e^\top_{n-1}P_{\bar{\eta}} + Q_{\bar{\eta}})\bar{\eta}\\
&\quad - \bar{\eta}^\top\bar{\eta} + \epsilon^2_{P_{\bar{\eta}}\bar{b}_y}\delta^2_{P_{\bar{\eta}}\bar{b}_y}y^2 + \epsilon^2_{P_{\bar{\eta}}\bar{\Delta}_b}\delta^2_{P_{\bar{\eta}}\bar{\Delta}_b}\bar{x}^2_{\rho+1}.
\end{aligned} \tag{5.37}$$

Substituting (5.34) into (5.37) yields (5.36), which completes the proof. □

5.1.4 Controller Design

To proceed with the controller design first note that the reparametrization error η_0 can be rewritten in terms of the new coordinate η_1, that is,

$$\eta_0 = \bar{\eta}_1 + \bar{a}_y(t)y, \tag{5.38}$$

where $\bar{a}_y(t)$ is unknown yet bounded due to Assumption 1.2. As discussed in Section 5.1.1, the input-output relation can be reparametrized using (5.7) and the filters (5.6). For the sake of implementing an I&I controller conveniently, rewrite the reparametrized system in the equivalent form[3]

$$
\begin{aligned}
\dot{y} &= \vartheta_3 \nu_1 + \varphi(y,d)^\top \vartheta_1 + \eta_0, \\
\dot{\nu}_1 &= \nu_2, \\
&\vdots \\
\dot{\nu}_{\rho-2} &= \nu_{\rho-1}, \\
\dot{\nu}_{\rho-1} &= -\lambda^\top \zeta + u,
\end{aligned}
\tag{5.39}
$$

where $\vartheta_1 = [\ell_{b_{\rho+1}}, \ldots, \ell_{b_n}, \ell_{a_1} - \lambda_{n-1}, \ldots, \ell_{a_{n-1}} - \lambda_1, \ell_{a_n}]^\top \in \mathbb{R}^{2n-\rho}$; $\vartheta_3 = \ell_{b_\rho} \in \mathbb{R}$; $\nu_1 = \zeta_{n-\rho+1}$; $\nu_2 = \zeta_{n-\rho+2}, \ldots$; $\nu_{\rho-1} = \zeta_{n-1}$; $\varphi(y,d) = [\zeta_{n-\rho}, \ldots, \zeta_1, -\lambda^\top \xi - y, \xi_{n-1}, \ldots, \xi_1]^\top$; and d stands for $\xi, \zeta_1, \ldots, \zeta_{n-\rho}$. Define $\vartheta_2 = \ell_{b_\rho}^{-1} \in \mathbb{R}$ and $\vartheta = [\vartheta_1^\top, \vartheta_2, \vartheta_3]^\top \in \mathbb{R}^q$, with $q \triangleq 2n - \rho + 2$. Similarly to the I&I approach with state feedback, a dynamic (integral) parameter estimate $\hat{\vartheta}$ and a static (proportional) parameter estimate β are exploited together for parameter estimation, that is, to use $\hat{\vartheta} + \beta$ for estimating ϑ. Also similarly to the state-feedback case, one can treat ν_i as a virtual input, governed by a virtual control law α_i, to control the dynamics of ν_{i-1}, for $i = 1, \ldots, \rho$, with $\nu_0 \triangleq y$. In the light of this, define now the error variables

$$z_1 = y, \tag{5.40}$$

$$z_i = \nu_{i-1} - \alpha_{i-1}; \tag{5.41}$$

[3] This subsection only discusses the case in which the relative degree $\rho \geq 2$. The case in which $\rho = 1$ is discussed in Section 5.1.5 via a simulation example. Note that the results in that case can be obtained without using backstepping.

the virtual control laws

$$\bar{\alpha}_1 = \sigma_1 + \varphi^\top(\hat{\vartheta}_1 + \beta_1), \tag{5.42}$$

$$\alpha_1 = -(\hat{\vartheta}_2 + \beta_2)\bar{\alpha}_1, \tag{5.43}$$

$$\begin{aligned}\alpha_2 = & -\sigma_2 - (\hat{\vartheta}_3 + \beta_3)y + \frac{\partial \alpha_1}{\partial \hat{\vartheta}}\dot{\hat{\vartheta}} + \frac{\partial \alpha_1}{\partial y}(-\sigma_1 + (\hat{\vartheta}_3 + \beta_3)z_2) \\ & + \frac{\partial \alpha_1}{\partial \xi}(A_\lambda \xi - e_{n-1}y) + \sum_{j=1}^{n-\rho}\frac{\partial \alpha_1}{\partial \zeta_j}\zeta_{j+1},\end{aligned} \tag{5.44}$$

$$\begin{aligned}\alpha_i = & -\sigma_i - z_{i-1} + \frac{\partial \alpha_{i-1}}{\partial \hat{\vartheta}}\dot{\hat{\vartheta}} + \frac{\partial \alpha_{i-1}}{\partial y}(-\sigma_1 + (\hat{\vartheta}_3 + \beta_3)z_2) \\ & + \sum_{j=1}^{i-2}\frac{\partial \alpha_{i-1}}{\partial \nu_j}\nu_{j+1} + \frac{\partial \alpha_{i-1}}{\partial \xi}(A_\lambda \xi - e_{n-1}y) + \sum_{j=1}^{n-\rho}\frac{\partial \alpha_{i-1}}{\partial \zeta_j}\zeta_{j+1},\end{aligned} \tag{5.45}$$

for $i = 2, \ldots, \rho$; the actual control law

$$u = \alpha_\rho + \lambda^\top \zeta; \tag{5.46}$$

and the update law for the dynamic parameter estimate

$$\begin{aligned}\dot{\hat{\vartheta}} = & -\left(I_q + \frac{\partial \beta}{\partial \hat{\vartheta}}\right)^{-1}\left(\frac{\partial \beta}{\partial y}\big(-\sigma_1 + (\hat{\vartheta}_3 + \beta_3)z_2\big)\right. \\ & \left. + \frac{\partial \beta}{\partial \nu_1}\nu_2 + \frac{\partial \beta}{\partial \xi}(A_\lambda \xi - e_{n-1}y) + \sum_{j=1}^{n-\rho}\frac{\partial \beta}{\partial \zeta_j}\zeta_{j+1}\right),\end{aligned} \tag{5.47}$$

where $\beta \triangleq [\beta_1^\top(y,d), \beta_2(y,\hat{\vartheta}_1,d), \beta_3(y,\hat{\vartheta}_1,\hat{\vartheta}_2,d)]^\top$ is the static parameter estimate, with

$$\beta_1 = \gamma_1 \int_0^y \varphi(\chi, d)\mathrm{d}\chi, \tag{5.48}$$

$$\beta_2 = \gamma_2 \mathrm{sgn}(\vartheta_3)\left(\frac{1}{2}\sigma_1 y + \int_0^y \varphi^\top(\chi,d)(\hat{\vartheta}_1 + \beta_1(\chi,d)\mathrm{d}\chi)\right), \tag{5.49}$$

$$\beta_3 = \gamma_3\left(\nu_1 y - \int_0^y \alpha_1(\chi, \hat{\vartheta}_1, \hat{\vartheta}_2, d)\mathrm{d}\chi\right); \tag{5.50}$$

$\gamma_1 > 0$, $\gamma_2 > 0$, $\gamma_3 > 0$ are constant adaptation gains; and $\sigma_1, \ldots, \sigma_\rho$ are damping terms to be defined.

Remark 5.4. *Equations* (5.48)–(5.50) *show that β_1 does not depend on $\hat{\vartheta}$; β_2 depends only on $\hat{\vartheta}_1$; and β_3 depends only on $\hat{\vartheta}_1$, $\hat{\vartheta}_2$. These indicate that $\dfrac{\partial \beta}{\partial \hat{\vartheta}}$ is a lower triangular* *matrix with all-zero diagonal terms. Thus* $\left(I + \dfrac{\partial \beta}{\partial \hat{\vartheta}}\right)^{-1}$ *exists and can be computed using the formula*[4]

$$\left(I + \frac{\partial \beta}{\partial \hat{\vartheta}}\right)^{-1} = \sum_{i=0}^{q-1} (-1)^i \left(\frac{\partial \beta}{\partial \hat{\vartheta}}\right)^i, \tag{5.51}$$

without explicitly computing the inverse online.

With the structure of the controller introduced, the design of the damping terms σ_i's is now ready to be presented.

Theorem 5.2. *Consider the system* (5.1)*, the filters specified by Proposition 5.1 and the adaptive controller described by* (5.42)-(5.50) *with the damping terms*

$$\sigma_1 = \left(c_1 + 2(\rho - 1)\epsilon^2 (\delta_{\bar{a}_y}^2 + \bar{\epsilon}^{-2} \epsilon_{b_y}^2 \delta_{b_y}^2 + \epsilon_{P_{\bar{\eta}} \bar{b}_y}^2 \delta_{P_{\bar{\eta}} \bar{b}_y}^2)\right) y, \tag{5.52}$$

$$\sigma_2 = \left(c_2 + \frac{3}{2\epsilon^2}\left(\frac{\partial \alpha_1}{\partial y}\right)^2\right) z_2, \tag{5.53}$$

$$\sigma_i = \left(c_i + \frac{3}{2\epsilon^2}\left(\frac{\partial \alpha_{i-1}}{\partial y}\right)^2\right) z_i, \quad i = 3, \ldots, \rho, \tag{5.54}$$

where $c_1 > 0, \ldots, c_\rho > 0$, $\epsilon > 0$, $\bar{\epsilon} > \epsilon_{P_{\bar{\eta}} \bar{\Delta}_b} \delta_{P_{\bar{\eta}} \bar{\Delta}_b}$*. Then, all trajectories of the closed-loop system are bounded and* $\lim\limits_{t \to +\infty} y(t) = 0$.

Proof. First, consider the dynamics of the parameter estimation error $z_\vartheta \triangleq \hat{\vartheta} - \vartheta + \beta$, which are described by the equation

$$\begin{aligned} \dot{z}_\vartheta = & \left(I + \frac{\partial \beta}{\partial \hat{\vartheta}}\right) \dot{\hat{\vartheta}} + \frac{\partial \beta}{\partial y} (-\sigma_1 - \varphi^\top z_{\vartheta_1} - \vartheta_3 z_{\vartheta_2} \bar{\alpha}_1 + \vartheta_3 z_2 + \eta_0) \\ & + \frac{\partial \beta}{\partial \nu_1} \nu_2 + \frac{\partial \beta}{\partial \xi} (A_\lambda \xi - e_{n-1} y) + \sum_{j=1}^{n-\rho} \frac{\partial \beta}{\partial \zeta_j} \zeta_{j+1}. \end{aligned} \tag{5.55}$$

Substituting $\dot{\hat{\vartheta}}$ into (5.55) yields

$$\dot{z}_\vartheta = -\frac{\partial \beta}{\partial y} (\Phi^\top z_\vartheta - \eta_0) = -\Gamma \Phi (\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y), \tag{5.56}$$

[4]The formula (5.51) is derived as follows. For $A \in \mathbb{R}^{q \times q}$, $1 - A^q = (1 - A) \sum_{i=0}^{q-1} A^i$ and in particular, if $A \triangleq -\frac{\partial \beta}{\partial \hat{\vartheta}}$, the characteristic polynomial of which is $p_A(s) = s^q$, one has $A^q = 0$ by Cayley-Hamilton Theorem, and $(1 - A)^{-1} = \sum_{i=0}^{q-1} A^i$, namely (5.51) holds.

where $\Gamma \triangleq \mathrm{diag}(\gamma_1, \gamma_2|\vartheta_2|, \gamma_3)$, with $\gamma_{(\cdot)} > 0$, and

$$\Phi \triangleq \begin{bmatrix} \varphi \\ \vartheta_3\bar{\alpha}_1 \\ z_2 \end{bmatrix}. \tag{5.57}$$

Consider the function $V_{z_\vartheta}(z_\vartheta) = \frac{1}{2} z_\vartheta^\top \Gamma^{-1} z_\vartheta$. Taking the time derivative of V_{z_ϑ} along the trajectories of (5.56) yields

$$\dot{V}_{z_\vartheta} = -z_\vartheta^\top \Phi(\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) \le -\frac{1}{2}(\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 + \delta_{\bar{a}_y}^2 y^2. \tag{5.58}$$

With the virtual control laws (5.42)–(5.45) and the actual control law (5.46), the dynamics of y can be rewritten as

$$\dot{y} = -\sigma_1 - (\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) + (\hat{\vartheta}_3 + \beta_3) z_2, \tag{5.59}$$

and the dynamics of the backstepping variables are described by

$$\begin{aligned} \dot{z}_2 &= -\sigma_2 - (\hat{\vartheta}_3 + \beta_3) y + z_3 + \frac{\partial \alpha_1}{\partial y}(\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y), \\ \dot{z}_i &= -\sigma_i - z_{i-1} + z_{i+1} + \frac{\partial \alpha_{i-1}}{\partial y}(\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y), \end{aligned} \tag{5.60}$$

for $i = 3, \ldots, \rho$, with $z_{\rho+1} \triangleq 0$. Define $z = [z_1, \ldots, z_\rho]^\top$ and consider the function $V_z(z) = \frac{1}{2}|z|^2$. The time derivative of V_z along the trajectories of the closed-loop system satisfies

$$
\begin{aligned}
\dot{V}_z = & -\sigma_1 y - y(\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) \\
& + \sum_{i=2}^{\rho} \bigg(-\sigma_i z_i + z_i \frac{\partial \alpha_{i-1}}{\partial y} (\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) \bigg) \\
\le & -\sum_{i=1}^{\rho} c_i z_i^2 - 2(\rho-1)\epsilon^2 (\delta_{\bar{a}_y}^2 + \bar{\epsilon}^2 \epsilon_{\bar{b}_y}^2 \delta_{\bar{b}_y}^2 + \epsilon_{P_{\bar{\eta}} \bar{b}_y}^2 \delta_{P_{\bar{\eta}} \bar{b}_y}^2) y^2 \\
& - \frac{3}{2\epsilon^2} \sum_{i=2}^{\rho} \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 z_i^2 + \frac{1 + \frac{1}{4} + \frac{1}{4}}{\epsilon^2} \sum_{i=2}^{\rho} \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 z_i^2 \\
& + (\rho-1)\epsilon^2 \bigg(\frac{1}{4} (\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 + \delta_{\bar{a}_y}^2 y^2 \bigg) \\
= & -\sum_{i=1}^{\rho} c_i z_i^2 - (\rho-1)\epsilon^2 (\delta_{\bar{a}_y}^2 + 2\bar{\epsilon}^2 \epsilon_{\bar{b}_y}^2 \delta_{\bar{b}_y}^2 + 2\epsilon_{P_{\bar{\eta}} \bar{b}_y}^2 \delta_{P_{\bar{\eta}} \bar{b}_y}^2) y^2 \\
& + (\rho-1)\epsilon^2 \bigg(\frac{1}{4} (\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 \bigg).
\end{aligned}
\tag{5.61}
$$

Finally, consider the function

$$
V(z, z_\vartheta, \bar{\eta}, \bar{x}) = V_z + (\rho-1)\epsilon^2 \bigg(V_{z_\vartheta} + 2(V_{\bar{\eta}} + \bar{\epsilon}^2 V_{\bar{x}}) \bigg). \tag{5.62}
$$

Recalling Assumption 5.1 and Proposition 5.1 and taking the time derivative of V along the trajectories of the closed-loop system yields

$$
\begin{aligned}
\dot{V} &= \dot{V}_z + (\rho-1)\epsilon^2 \bigg(\dot{V}_{z_\vartheta} + 2(\dot{V}_{\bar{\eta}} + \bar{\epsilon}^2 \dot{V}_{\bar{x}}) \bigg) \\
&\le -\sum_{i=1}^{\rho} c_i z_i^2 - \frac{(\rho-1)\epsilon^2}{4} (\Phi^\top z_\vartheta)^2 \le 0.
\end{aligned}
\tag{5.63}
$$

Boundedness. $\dot{V} \le 0$ guarantees that z, z_ϑ, $\bar{\eta}$, $\bar{x}$ are bounded. Note that ξ is bounded since $y = z_1$ is bounded and the bounded-input bounded-state property of the ξ-dynamics (5.6). By (5.28), for $i = 1, \ldots, n-\rho$, one has

$$
\zeta_i = \frac{s^{i-1}}{\Lambda(s)}[u] = \frac{s^{i-1}}{\Lambda(s)} \left[\frac{1}{b_\rho} \bigg(-\bar{x}_{\rho+1} + y^{(\rho)} + \sum_{j=0}^{\rho-1} a_{y^{(j)}} y^{(j)} \bigg) \right], \tag{5.64}
$$

which indicates that $\zeta_1, \ldots, \zeta_{n-\rho}$ are bounded, since $\Lambda(s)$ is Hurwitz and $\bar{x}_{\rho+1}$, y are bounded. Then α_1 is bounded. By using (5.41) recursively, one can establish boundedness for ν_i and α_i, $i = 2, \ldots, \rho$, which leads to boundedness of $\zeta_{n-\rho+1}, \ldots, \zeta_{n-1}$ and boundedness of u. By (5.13) and (5.14), η and η_0 are also bounded due to the fact that u and y are bounded. Then by (5.39) and Assumption 1.2, it can be concluded that $y^{(1)}, \ldots, y^{(\rho)}$ are bounded. Furthermore, equations (5.22) imply that $x_1, \ldots, x_\rho$ are bounded. Meanwhile, note that boundedness of $\bar{x}$ and y, along with Assumption 1.2, implies boundedness of $x_{\rho+1}, \ldots, x_n$. Hence the system state x is bounded. Finally, β is bounded, which yields that $\hat{\vartheta}$ is bounded. This completes the proof of boundedness of the closed-loop signals.

Convergence. In addition to boundedness of z, by boundedness of the closed-loop signals, $\dot{z}$ is also bounded. Therefore, by Lemma A.5 we can conclude from (5.63) that $\lim_{t\to+\infty} z(t) = 0$, which further indicates that $\lim_{t\to+\infty} y(t) = 0$. The proof is then complete. □

Theorem 5.3. *Consider the system* (5.1), *the filters specified by Proposition 5.1 and the adaptive controller described by* (5.42)-(5.50) *with the damping terms*

$$\sigma_1 = \left(c_1 + (\rho - 1)\epsilon^2 \delta_{\bar{a}_y}^2\right) y, \tag{5.65}$$

$$\sigma_2 = \left(c_2 + \frac{3}{2\epsilon^2}\left(\frac{\partial \alpha_1}{\partial y}\right)^2\right) z_2, \tag{5.66}$$

$$\sigma_i = \left(c_i + \frac{3}{2\epsilon^2}\left(\frac{\partial \alpha_i}{\partial y}\right)^2\right) z_i, \quad i = 3, \ldots, \rho, \tag{5.67}$$

where $c_2 > 0$, ... , $c_\rho > 0$, and $\epsilon > 0$. Then, there exists $c_1 > 0$, such that all trajectories of the closed-loop system are bounded and $\lim_{t\to+\infty} y(t) = 0$. Moreover, if the system (5.1) *is time-invariant, any $c_1 > 0$ with $\delta_{\bar{a}_y} = 0$ guarantees the same boundedness and convergence properties.*

Proof. Following the same path of the proof of Theorem 5.2 but with σ_1 defined by (5.65) yields

$$\begin{aligned}
\dot{V}_z = & -\sigma_1 y - y(\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) \\
& + \sum_{i=2}^{\rho} \left(-\sigma_i z_i + z_i \frac{\partial \alpha_{i-1}}{\partial y} (\Phi^\top z_\vartheta - \bar{\eta}_1 - \bar{a}_y y) \right) \\
\leq & -\sum_{i=1}^{\rho} c_i z_i^2 - (\rho - 1)\epsilon^2 \delta_{\bar{a}_y}^2 y^2 \\
& - \frac{3}{2\epsilon^2} \sum_{i=2}^{\rho} \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 z_i^2 + \frac{1 + \frac{1}{4} + \frac{1}{4}}{\epsilon^2} \sum_{i=2}^{\rho} \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 z_i^2 \\
& + (\rho - 1)\epsilon^2 \left(\frac{1}{4} (\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 + \delta_{\bar{a}_y}^2 y^2 \right) \\
= & -c_1 y^2 + (\rho - 1)\epsilon^2 \left(\frac{1}{4} (\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 \right) - \sum_{i=2}^{\rho} c_i z_i^2.
\end{aligned} \tag{5.68}$$

Furthermore, by Assumption 5.1, Proposition 5.1, and (5.58), we have

$$\dot{V}_{\bar{x}} \leq -\bar{x}_{\rho+1}^2 + \epsilon_{b_y}^2 \delta_{b_y}^2 y^2 - \sum_{i=\rho+2}^{n} \bar{x}_i^2, \tag{5.69}$$

$$\dot{V}_{\bar{\eta}} \leq -\bar{\eta}_1^2 + \epsilon_{P_{\bar{\eta}}\bar{b}_y}^2 \delta_{P_{\bar{\eta}}\bar{b}_y}^2 y^2 + \epsilon_{P_{\bar{\eta}}\bar{\Delta}_b}^2 \delta_{P_{\bar{\eta}}\bar{\Delta}_b}^2 \bar{x}_{\rho+1}^2 - \sum_{i=2}^{n-1} \bar{\eta}_i^2, \tag{5.70}$$

$$\dot{V}_{z_\vartheta} \leq -\frac{1}{2} (\Phi^\top z_\vartheta)^2 + \bar{\eta}_1^2 + \delta_{\bar{a}_y}^2 y^2, \tag{5.71}$$

respectively. This allows writing

$$\dot{\bar{V}} \leq -E\psi - e_1 \sum_{i=2}^{\rho} c_i z_i^2 - e_2 \sum_{i=\rho+2}^{n} \bar{x}_i^2 - e_3 \sum_{i=2}^{n-1} \bar{\eta}_i^2 \tag{5.72}$$

with $\bar{V} \triangleq [V_z, V_{\bar{x}}, V_{\bar{\eta}}, V_{z_\vartheta}]^\top$, $\psi \triangleq [y^2, \bar{x}_{\rho+1}^2, \bar{\eta}_1^2, |z_\vartheta|^2]^\top$, e_i the ith unit vector, and

$$E \triangleq \begin{bmatrix} c_1 & 0 & -(\rho-1)\epsilon^2 & -\frac{(\rho-1)\epsilon^2}{4} \\ -\epsilon_{b_y}^2 \delta_{b_y}^2 & 1 & 0 & 0 \\ -\epsilon_{P_{\bar{\eta}}\bar{b}_y} & -\epsilon_{P_{\bar{\eta}}\bar{\Delta}_b}^2 \delta_{P_{\bar{\eta}}\bar{\Delta}_b}^2 & 1 & 0 \\ \delta_{\bar{a}_y}^2 & 0 & 1 & -\frac{1}{2} \end{bmatrix}. \tag{5.73}$$

The underlying graph of E is depicted in Fig. 5.1(a). Clearly, the z-node is contained in every directed cycle and if we treat it as an *active node* by letting c_1 adjustable, the condition of Theorem 3.3 is satisfied. Then, applying Theorem 3.2 and Theorem 3.3 guarantees that there exists $c_1 > 0$ such that E is a non-singular *M-matrix*, which further indicates that for all $\varpi > 0$, there exists a vector scaling coefficients $\varsigma > 0$, depending on ϖ, such that the time derivative of $V \triangleq \varsigma^\top \bar{V}$ satisfies the dissipation inequality

$$\dot{V} = \varsigma^\top E\psi \leq \varpi^\top \psi + \varsigma_1 \sum_{i=2}^{\rho} z_i^2 + \varsigma_2 \sum_{i=\rho+2}^{n} \bar{x}_i^2 + \varsigma_3 \sum_{i=2}^{n-1} \bar{\eta}_i^2 \leq 0. \tag{5.74}$$

Finally, following the same routine as in the proof of Theorem 5.2, one can prove boundedness of all closed-loop signals and $\lim_{t\to+\infty} y(t) = 0$.

Consider now the classical time-invariant case, in which the system parameters a and b are constant. It is not difficult to observe that in this case the dynamics of η described by (5.13) are not coupled with y or u. Therefore, one can define $\bar{\eta} = \eta$ (as there is no $\bar{x}_{\rho+1}$ signal or y signal injected in the η-dynamics) and the edges $\bar{x} \to \bar{\eta}$ and $z \to \bar{\eta}$ are removed. Furthermore, since η_0 is no longer perturbed by y, the y-term in the z_ϑ-dynamics (which was brought by η_0) does not appear and the edge $z \to z_\vartheta$ is removed. These reduce the underlying graph from the one in Fig. 5.1(a) to the one in Fig.5.1(b), which is acyclic. Therefore the condition of Theorem 3.3 is naturally satisfied for any $c_1 > 0$, and this guarantees the same boundedness and convergence properties using the analysis argument used for the time-varying case. □

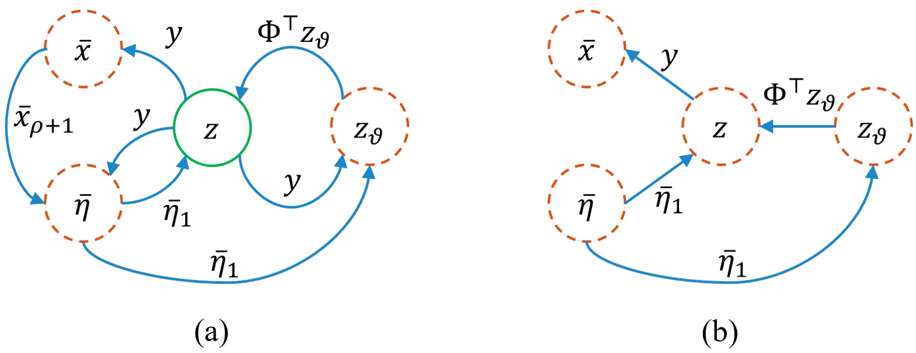

Figure 5.1: Schematic interpretation of the interconnected z, $\bar{x}$, $\bar{\eta}$ and z_ϑ subsystems: (a) the interconnection of the case in which the system parameters are time-varying and (b) the interconnection of the classical time-invariant case.

Remark 5.5. *Compared with the constructive method used in the proof of Theorem 5.2, the active-node-based method used to prove Theorem 5.3 reveals better the effect of the* *time-varying parameters: the time-varying perturbations introduce couplings between subsystems and cyclic interconnections, which do not appear in the classical time-invariant case. In the classical case, the directed graph describing the subsystem interconnection is acyclic. The classical analysis, say, the one in Section 4.2.4 of [5], separately establishes the boundedness and convergence properties of individual subsystems, and the same properties can be extended to the overall system due to the acyclic interconnection. The time-varying case, however, requires a small-gain-like analysis exploiting Theorem 3.3, in which a strengthened damping design (parametrized by c_1) using y is performed, which dominates the cyclic interconnection. In other words, the classical scheme becomes a special case of the proposed scheme for the time-varying case, as claimed by Theorem 5.3.*

5.1.5 Simulations

To verify the effectiveness of the I&I scheme, consider the 2-dimensional system with relative degree $\rho = 1$, described by the equations

$$\begin{aligned} \dot{x_1} &= -a_1(t) + b_1(t)u, \\ \dot{x_2} &= -a_2(t) + b_2(t)u, \\ y &= x_1, \end{aligned} \tag{5.75}$$

where b_1, b_2, $a_)$, a_2 are time-varying parameters to be defined. Since system (5.75) has relative degree $\rho = 1$, the reparametrized system (5.39) reduces to

$$\dot{y} = \vartheta_3(u - \lambda\zeta) + \varphi(y, \zeta, \xi)\vartheta_1 + \eta_0, \tag{5.76}$$

where $\varphi \triangleq [\zeta, -\lambda\xi - y, \xi]^\top$ and ζ, ξ are the states of the filters

$$\begin{aligned} \dot{\zeta} &= -\lambda\zeta + u, \\ \dot{\xi} &= -\lambda\xi - y. \end{aligned} \tag{5.77}$$

The time-varying parameters are given by

$$b_1(t) = 1 + 0.2\sin(3t), \quad b_2(t) = 2 + 1.5\cos(5t),$$
$$a_1(t) = 1 + \frac{5\mathrm{sgn}(y)\varphi_2}{\sqrt{\varphi_2^2+\varphi_3^2}}, \quad a_2(t) = 1 + \frac{5\mathrm{sgn}(y)\varphi_3}{\sqrt{\varphi_2^2+\varphi_3^2}}, \tag{5.78}$$

where φ_2, φ_3 are the second and the third element of φ, respectively. a_1 and a_2 are state-dependent time-varying parameters designed to destabilize the system. One can easily verify that Assumption 5.1 holds since the zero dynamics of (5.75) is exponentially stable due to the fact that $\frac{b_2}{b_1}$ is strictly positive. Also note that the reparametrized model (5.76) is one-dimensional and the $\hat{\vartheta}_3 + \beta_3$ term does not appear. Thus, the update law (5.47) is reduced to

$$\dot{\hat{\vartheta}} = -\left(I + \frac{\partial\beta}{\partial\hat{\vartheta}}\right)^{-1}\left(\frac{\partial\beta}{\partial y}(-\sigma_1 + \frac{\partial\beta}{\partial\xi}(-\lambda\xi - y) + \frac{\partial\beta}{\partial\zeta}(-\lambda\zeta + u)\right). \tag{5.79}$$

The control law design is not affected by relative degree and therefore (5.46) is considered.

For comparison, consider three scenarios as follows. In the "Baseline" scenario, only the constant components of the system parameters are considered, namely $b_1 = 1$, $b_2 = 2$, $a_1 = 1$, $a_2 = 1$, and the system is time-invariant, controlled by a classical I&I adaptive controller with $\sigma_1 = y$. In the "Controller 1" scenario, the same classical I&I controller is adopted whereas the system parameters are time-varying, specified by (5.77). In the "Controller 2" scenario, the proposed adaptive controller with a strengthened damping term $\sigma_1 = 8y$ is used for the time-varying system. In the three scenarios, the rest of the controller parameters are set the same, namely $\gamma_1 = 0.5$, $\gamma_2 = 0.1$, $\lambda = 10$, and all filter states and parameter estimates are initialized at 0.

Similar to the simulation study in Section 4.4, the scenarios "Controller 1" and "Controller 2" are simulated in parallel twice to keep the comparison fair considering the use of state-dependent parameters. In the first simulation set, both scenarios are driven by the state-dependent parameters generated by the "Controller 1" scenario and in the second simulation set, the state-dependent parameters are generated by the "Controller 2" scenario.

The results of both simulation sets are presented in Fig. 5.2 and Fig. 5.3. One can see that in both simulation sets, the "Controller 2" scenario outperforms the "Controller 1" scenario (with fast convergence and fewer oscillations in the transient stage). The proposed controller also restores the performance of the "Baseline" scenario in terms of settling time.

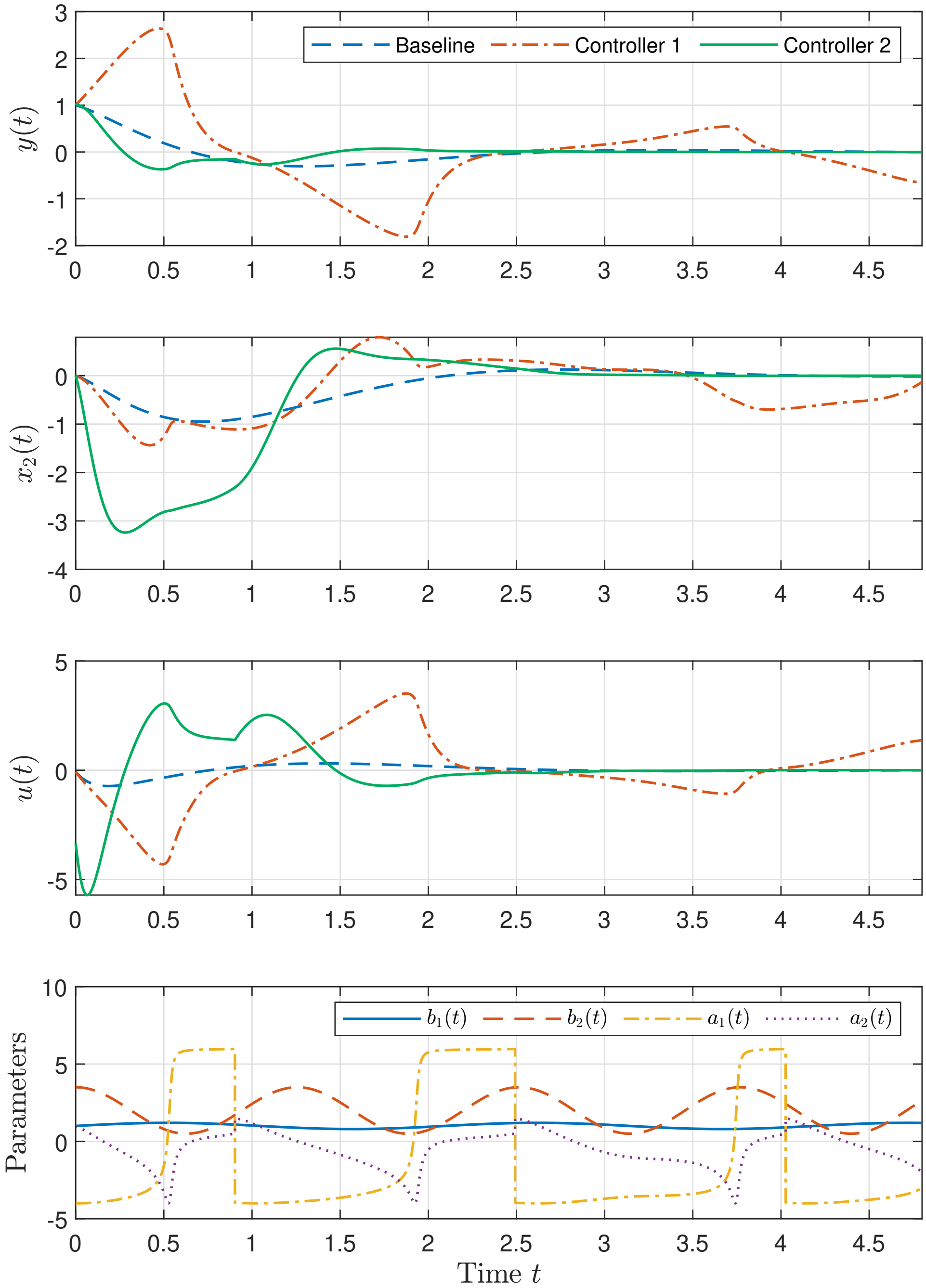

Figure 5.2: Simulation set 1: time histories of the system output, state, control effort, and state-dependent time-varying parameters for different controllers.

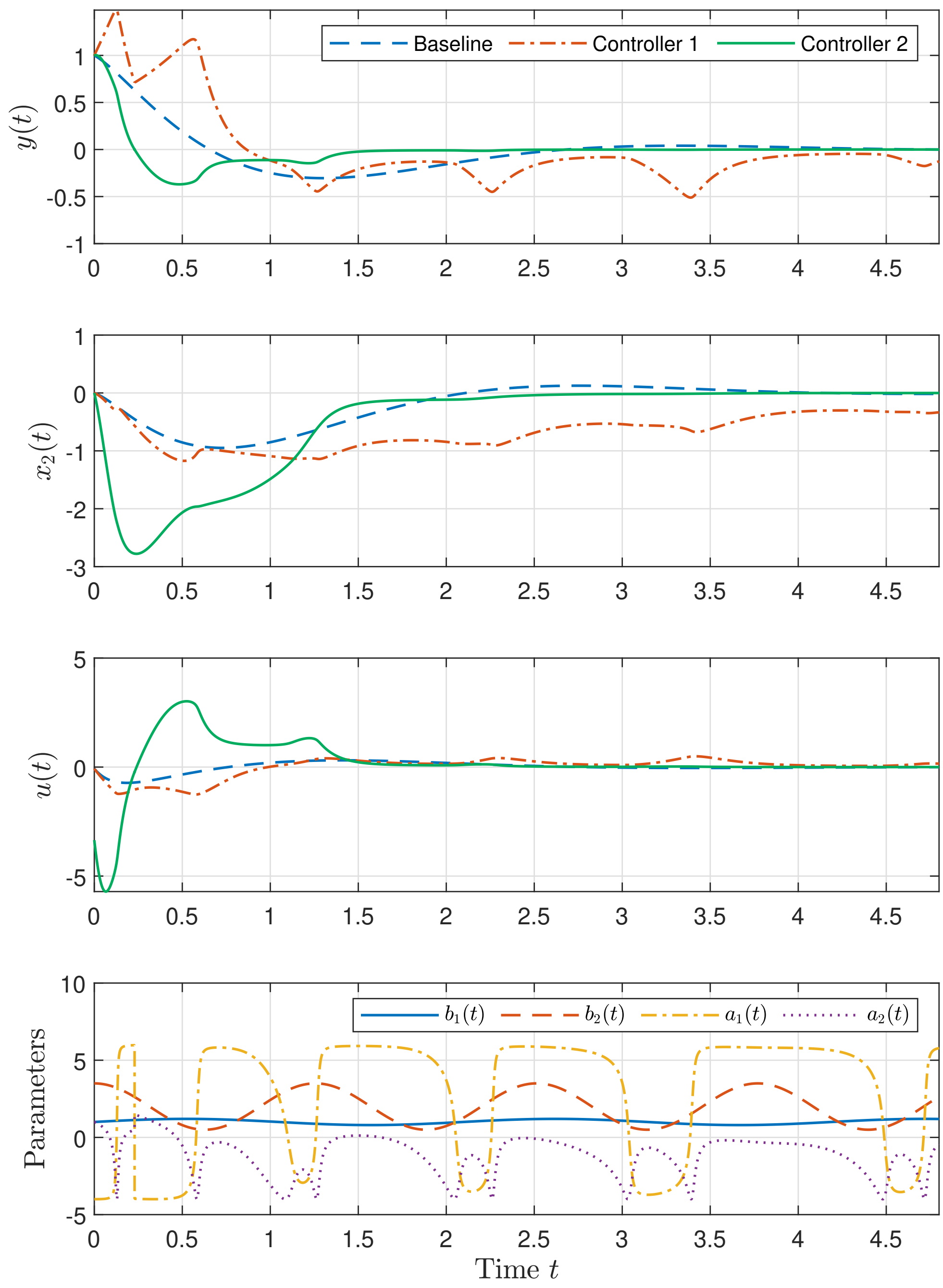

Figure 5.3: Simulation set 2: time histories of the system output, state, control effort, and state-dependent time-varying parameters for different controllers.

5.2 Passivity-Based Design for Nonlinear Systems in Observer Form

In this section, we proceed to investigate the output-feedback adaptive control problem for a class of nonlinear systems, and at the same time, relax the restriction imposed by Assumption 5.2 to the I&I scheme introduced in Section 5.1. Due to the increased complexity of the design, the passivity-based scheme is adopted throughout the section. Consider an n-dimensional system in output feedback form with $n \geq 2$ and relative degree ρ, described by the equations

$$
\begin{aligned}
\dot{x}_1 &= x_2 + \phi_{0,1}(y) + \sum_{j=1}^{q} \phi_{1,j}(y) a_j(t), \\
&\vdots \\
\dot{x}_\rho &= x_{\rho+1} + \phi_{0,\rho}(y) + \sum_{j=1}^{q} \phi_{\rho,j}(y) a_j(t) + b_m(t) g(y) u, \\
&\vdots \\
\dot{x}_n &= \phi_{0,n}(y) + \sum_{j=1}^{q} \phi_{n,j}(y) a_j(t) + b_0(t) g(y) u, \\
y &= x_1,
\end{aligned}
\tag{5.80}
$$

or, in compact form, by the equations

$$
\begin{aligned}
\dot{x} &= Sx + \phi_0(y) + F^\top(y,u)\theta, \\
y &= e_1^\top x,
\end{aligned}
\tag{5.81}
$$

where $x(t) = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ is the state; $u(t) \in \mathbb{R}$ is the input; $y(t) \in \mathbb{R}$ is the output; $\theta(t) \triangleq [b^\top(t), a^\top(t)]^\top$ is the vector of unknown time-varying parameters, with $a(t) \triangleq [a_1(t), \ldots, a_q(t)]^\top \in \mathbb{R}^q$, $b(t) \triangleq [b_m(t), \ldots, b_0(t)]^\top \in \mathbb{R}^{m+1}$, $m \triangleq n - \rho$;

$$
F^\top(y,u) \triangleq \left[\begin{bmatrix} 0_{(\rho-1)\times(m+1)} \\ I_{m+1} \end{bmatrix} g(y)u, \Phi^\top(y) \right],
\tag{5.82}
$$

with $(\Phi^\top(y))_{ij} = \phi_{i,j}(y)$, and $g : \mathbb{R} \to \mathbb{R}$ a smooth mapping satisfying $g(y) \neq 0$, for all $y \in \mathbb{R}$. In addition, θ satisfies Assumption 1.2, and, in particular, b_m also satisfies Assumption 1.3. The mappings $\phi_{0,i} : \mathbb{R} \to \mathbb{R}$ and $\phi_{i,j} : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, $j = 1, \ldots, q$, satisfy Assumption 1.4.

Remark 5.6. *It is easy to see, by invoking Lemma A.3, that there exist smooth mappings* $\bar{\phi}_{0,i}$ *and* $\bar{\phi}_{i,j}$ *such that* $\phi_{0,i}(y) = \bar{\phi}_{0,i}(y)y$, $\phi_{i,j}(y) = \bar{\phi}_{i,j}(y)y$.

5.2.1 System Reparametrization

To cope with the unmeasured state variables, *Kreisselmeier filters* (*K-filters*) [72] are applied to reparametrize the system with the filter state variables (which are known) into a new form that is favourable for the backstepping design. The filters are given by the equations

$$\dot{\xi} = A_k \xi + k y + \phi_0(y), \tag{5.83}$$

$$\dot{\Xi}^\top = A_k \Xi^\top + \Phi^\top(y), \tag{5.84}$$

$$\dot{\lambda} = A_k \lambda + e_n g(y) u, \tag{5.85}$$

where $A_k = S - k e_1^\top$ and $k \in \mathbb{R}^n$ is the vector of filter gains. These filters are equivalent, see [75], to the filters

$$\dot{\xi} = A_k \xi + k y + \phi_0(y), \tag{5.86}$$

$$\dot{\Omega}^\top = A_k \Omega^\top + F^\top(y, u), \tag{5.87}$$

where

$$\Omega^\top \triangleq [v_m, \ldots, v_0, \Xi^\top], \tag{5.88}$$

$$v_i \triangleq A_k^i \lambda, \quad i = 0, \ldots, m. \tag{5.89}$$

Define now the non-implementable state estimate

$$\hat{x} = \xi + \Omega^\top \ell_\theta. \tag{5.90}$$

The state estimation error dynamics are then described by the equation

$$\begin{aligned}\dot{\varepsilon} &= A_k\varepsilon + F^\top(y,u)\Delta_\theta \\ &= A_k\varepsilon + \Phi^\top(y)\Delta_a + \begin{bmatrix} 0_{(\rho-1)\times 1} \\ \Delta_b \end{bmatrix} g(y)u,\end{aligned} \tag{5.91}$$

where $\varepsilon = x - \hat{x}$. After using the *K-filters* (5.83)–(5.85) with the *congelation of variables* method, the original n-dimensional system with time-varying parameters can be reparametrized as a ρ-dimensional system with constant parameters ℓ_θ and some auxiliary systems to be defined. The substitution of ℓ_θ for θ prevents $\dot{\theta}$ terms from appearing in the ε-dynamics. For $\rho > 1$ one has the system described by the equations

$$\begin{aligned}\dot{y} &= \omega_0 + \bar{\omega}^\top \ell_\theta + \varepsilon_2 + \ell_{b_m} v_{m,2} \\ &\vdots \\ \dot{v}_{m,i} &= -k_i v_{m,1} + v_{m,i+1}, \quad i = 2, \ldots, \rho - 1, \\ &\vdots \\ \dot{v}_{m,\rho} &= -k_\rho v_{m,1} + v_{m,\rho+1} + g(y)u,\end{aligned} \tag{5.92}$$

and, for $\rho = 1$ one has

$$\dot{y} = \omega_0 + \omega^\top \ell_\theta + \varepsilon_2 + \ell_{b_m} g(y)u, \tag{5.93}$$

where $\omega_0 \triangleq \phi_{0,1} + \xi_2$, $\bar{\omega} \triangleq [0, v_{m-1,2}, \ldots, v_{0,2}, (\Phi)_1^\top + (\Xi)_2^\top]^\top$, and $\omega \triangleq \bar{\omega} + e_1 v_{m,2}$.

Similarly to the classical adaptive backstepping scheme we consider the ρth order system (5.92) (or (5.93) if $\rho = 1$) to exploit its *lower-triangular* structure. It should be noted that the reparametrized models (5.92) and (5.93) are only valid if the estimation error ε_2 converges to 0 (this is similar to the requirement that η_0 converges to 0 in Section 5.1). In classical schemes this is not a problem since the Δ_a and Δ_b terms are not present in the ε-dynamics and ε converges to 0 exponentially provided A_k is Hurwitz. The difference in the time-varying case is that the estimation error ε_2 may not converge to 0 due to the presence of the Δ_a and Δ_b terms, which may make the reparametrized models (5.92) and (5.93) behave differently from the original system and therefore not appropriate to be used for analysis or design. The effect of Δ_a can be dominated via a strengthened damping design, as proposed in [16]. However, the dominance method cannot be directly applied to (5.91) since Δ_b is coupled with the input u, and any additional damping terms added (to u) for dominance, in turn, alter the perturbation term $\Delta_b u$ itself. To avoid this issue, one can transform the perturbation terms coupled with u into new perturbation terms coupled with y by exploiting the inverse dynamics of system (5.80), in a similar spirit of the solution presented in Section 5.1.

5.2.2 Inverse Dynamics

The inverse dynamics of (5.80) are more complex than that of (5.1) considered in Section 5.1 due to the presence of nonlinearities, whereas the spirit of their derivation is essentially the same. First, pretend that the system is "driven" by the signals y, $\phi_{0,i}(y)$, $\phi_i(y)$, and their time derivatives. Then one could write

$$\begin{aligned} x_2 &= y^{(1)} - (\phi_1^\top a + \phi_{0,1}), \\ &\vdots \\ x_\rho &= y^{(\rho-1)} - (\phi_1^\top a + \phi_{0,1})^{(\rho-2)} - \cdots - (\phi_{\rho-1}^\top + \phi_{0,\rho-1}). \end{aligned} \tag{5.94}$$

Defining $y_i = \phi_i^\top a + \phi_{0,i}$, $i = 1, \dots, n$ and $u_g = g(y)u$, yields

$$u_g = \frac{1}{b_m}(-x_{\rho+1} + y^{(\rho)} - y_1^{(\rho-1)} - \cdots - y_\rho). \tag{5.95}$$

The resulting inverse dynamics are then described by

$$\begin{aligned} \dot{x}_{\rho+1} = &\ -\frac{b_{m-1}}{b_m}x_{\rho+1} + x_{\rho+2} + y_{\rho+1} + \frac{b_{m-1}}{b_m}(y^{(\rho)} - y_1^{(\rho-1)} - \cdots - y_\rho), \\ &\vdots \\ \dot{x}_n = &\ -\frac{b_0}{b_m}x_{\rho+1} + y_n + \frac{b_0}{b_m}(y^{(\rho)} - y_1^{(\rho-1)} - \cdots - y_\rho). \end{aligned} \tag{5.96}$$

Since it is difficult to use backstepping techniques to establish stability, or convergence, properties for the time derivatives of y or y_i, we need to perform a change of coordinates, or more intuitively, propagating the time derivatives of the output-related signals along the chain of integrators to eliminate them from the inverse dynamics. Recall the identity (5.24)

Algorithm 5.3 Change of coordinates $x_{\rho+1}, \ldots, x_n$.

Input: $x_{\rho+1}, \ldots, x_n, \dot{x}_{\rho+1}, \ldots, \dot{x}_n$.
Output: $\bar{x}_{\rho+1}, \ldots, \bar{x}_n, \dot{\bar{x}}_{\rho+1}, \ldots, \dot{\bar{x}}_n$.

while time derivatives of y appear in the expression of $\dot{x}_{\rho+1}, \ldots, \dot{x}_n$ **do** ▷ This while-loop iterates for ρ times as it reduces the order of $y^{(\rho)}$ by one each iteration.
  for $i = n \to \rho + 2$ **do**
    Update $\bar{x}_i$ and $\dot{\bar{x}}_i$ using (5.24).
    Rewrite x_i in terms of $\bar{x}_i$ in the expression of $\dot{x}_{i-1}$ and leave the feedback term $-\frac{b_{n-i}}{b_m} x_{\rho+1}$ unchanged.
  end for
  Update $\bar{x}_{\rho+1}$ and $\dot{\bar{x}}_{\rho+1}$ using (5.24).
  Rewrite $x_{\rho+1}$ in terms of $\bar{x}_{\rho+1}$ in the expressions of $\dot{\bar{x}}_{\rho+1}, \ldots, \dot{\bar{x}}_n$, respectively. ▷ This brings back the time derivatives of $y, y_1, \ldots, y_\rho$, but with the order reduced by one.
  $x_{\rho+1} \leftarrow \bar{x}_{\rho+1}$, … , $x_n \leftarrow \bar{x}_n$, $\dot{x}_{\rho+1} \leftarrow \dot{\bar{x}}_{\rho+1}$, … , $\dot{x}_n \leftarrow \dot{\bar{x}}_n$. ▷ Update the old coordinates before the next iteration.
end while

in Section 5.1. With this fact, the change of coordinate

$$\bar{x}_n = x_n - \sum_{j=0}^{\rho-1} (-1)^j \left(\frac{b_0}{b_m}\right)^{(j)} y^{(\rho-1-j)} + \sum_{i=1}^{\rho-1} \sum_{j=0}^{\rho-i-1} (-1)^j \left(\frac{b_0}{b_m}\right)^{(j)} y_i^{(\rho-i-1-j)} \tag{5.97}$$

yields

$$\dot{\bar{x}}_n = -\frac{b_0}{b_m} x_{\rho+1} + y_n + (-1)^\rho \left(\frac{b_0}{b_m}\right)^{(\rho)} y - \sum_{i=1}^{\rho} (-1)^{\rho-i} \left(\frac{b_0}{b_m}\right)^{(\rho-i)} y_i, \tag{5.98}$$

which does not contain time derivatives of y and y_i's. In the same spirit, applying the change of coordinates specified by Algorithm 5.3, we are able to remove the terms containing the time derivatives of y and y_i's in each equation describing the inverse dynamics. The resulting inverse dynamics in the new coordinates (denoted by $\bar{x}_i$, $i = \rho + 1, \ldots, n$, with a slight abuse of notation) are described by the equations

$$\dot{\bar{x}} = A_{\bar{b}}(t)\bar{x} + b_{\bar{x}y}(t)y + \sum_{i=1}^{n} b_{\bar{x}\phi,0,i}(t)\phi_{0,i}(y) + \sum_{i=1}^{n} \sum_{j=1}^{q} b_{\bar{x}\phi,i,j}(t)\phi_{i,j}(y), \tag{5.99}$$

$$u_g = \frac{1}{b_m(t)} \left(-\bar{x}_{\rho+1} + y^{(\rho)} + \sum_{j=0}^{\rho-1} a_{u_g y^{(j)}}(t) y^{(j)} + \sum_{i=1}^{\rho} \sum_{j=0}^{\rho-i} a_{u_g y_i^{(j)}}(t) y_i^{(j)} \right), \tag{5.100}$$

where $\bar{x}(t) \triangleq [\bar{x}_{\rho+1}(t), \ldots, \bar{x}_n(t)]^\top \in \mathbb{R}^m$, $A_{\bar{b}} \triangleq S - \bar{b}e_1^\top$, $\bar{b}(t) \triangleq \frac{1}{b_m(t)}[b_{m-1}(t), \ldots, b_0(t)]^\top \in \mathbb{R}^m$.

Remark 5.7. *The time-varying vectors $b_{\bar{x}y}$, $b_{\bar{x}\phi,0,i}$, $b_{\bar{x}\phi,i,j}$ and the time-varying scalars $a_{u_g y^{(j)}}$, $a_{u_g y_i^{(j)}}$ are unknown as they depend on the unknown system parameter vector θ. However, as a consequence of Assumption 1.2, they are bounded.*

Assumption 5.3 (Strong minimum-phase property)**.** *The time-varying system* (5.80) *has a strong minimum-phase property in the sense that the inverse dynamics* (5.99) *are input-to-state stable (ISS) with respect to the inputs y, $\phi_{0,i}(y)$, $\phi_{i,j}(y)$, $i = 1, \ldots, n$, $j = 1, \ldots, q$. Moreover, there exists an ISS Lyapunov function $\underline{\gamma}_{\bar{x}}|\bar{x}|^2 \leq V_{\bar{x}}(\bar{x}, t) \leq \bar{\gamma}_{\bar{x}}|\bar{x}|^2$, $0 \leq \underline{\gamma}_{\bar{x}} \leq \bar{\gamma}_{\bar{x}}$, and the time derivative of $V_{\bar{x}}$ along the trajectories of the inverse dynamics satisfies the inequality*

$$\dot{V}_{\bar{x}} \leq -|\bar{x}|^2 + \sigma_{\bar{x}y}y^2 + \sigma_{\bar{x}\phi_0}|\phi_0(y)|^2 + \sigma_{\bar{x}\Phi}|\Phi(y)|_{\mathrm{F}}^2, \tag{5.101}$$

for some constant $\sigma_{(\cdot)} > 0$.

Remark 5.8. *Assumption 5.3 holds if $\bar{x} = 0$ is a globally exponentially stable equilibrium of the zero dynamics described by $\dot{\bar{x}} = A_{\bar{b}}(t)\bar{x}$, see* e.g. *Lemma 4.6 in [70]. Some works (*e.g. *[129] and [86]) exploit this exponential stability property as a substitute for the classical minimum-phase assumption. Note, finally, that Assumption 5.3 is not more restrictive than the classical minimum-phase assumption because for time-invariant systems Assumption 5.3 reduces to minimum-phaseness.*

5.2.3 Filter Design

Consider now the state estimation error dynamics (5.91) with u_g given by (5.100), which yields

$$\begin{aligned}\dot{\varepsilon} = A_k\varepsilon + \Phi^\top(y)\Delta_a + &\begin{bmatrix} 0_{(\rho-1)\times 1} \\ \Delta_b \end{bmatrix} \frac{1}{b_m} \times \\ &\left(-\bar{x}_{\rho+1} + y^{(\rho)} + \sum_{j=0}^{\rho-1} a_{u_g y^{(j)}}(t) y^{(j)} + \sum_{i=1}^{\rho}\sum_{j=0}^{\rho-i} a_{u_g y_i^{(j)}}(t) y_i^{(j)} \right).\end{aligned} \tag{5.102}$$

Similarly to what is done in Section 5.2.2, we need to use a change of coordinates to remove the time derivative terms brought by u_g. Implementing a change of coordinates in the same spirit of Algorithm 5.3, the state estimation error dynamics in the new coordinates

$\bar{\varepsilon}$ are described by the equations

$$\begin{aligned}\dot{\bar{\varepsilon}} &= A_k\bar{\varepsilon} - \bar{\Delta}_b\bar{x}_{\rho+1} + b_{\bar{\varepsilon}y}(t)y \\ &\quad + \sum_{i=1}^{n} b_{\bar{\varepsilon}\phi,0,i}(t)\phi_{0,i}(y) + \sum_{i=1}^{n}\sum_{j=1}^{q} b_{\bar{\varepsilon}\phi,i,j}(t)\phi_{i,j}(y),\end{aligned} \tag{5.103}$$

where $\bar{\Delta}_b = [0_{1\times(\rho-1)}, \Delta_b^\top]^\top \dfrac{1}{b_m}$.

Remark 5.9. *The time derivative terms are injected into the ε-dynamics via the vector of time-varying gains Δ_b. Similarly to Remark 5.7, the time-varying vectors $\bar{\Delta}_b$, $b_{\bar{\varepsilon}y}$, $b_{\bar{\varepsilon}\phi,0,i}$, $b_{\bar{\varepsilon}\phi,i,j}$ are unknown, yet bounded, due to Assumption 1.2. We will see that as long as these parameters are bounded they do not affect the controller design. In particular, when b is constant, $\Delta_b(t) = 0$, for all $t \geq 0$, provided $\ell_b = b$. Thus $\bar{\Delta}_b(t)$, $b_{\bar{\varepsilon}y}(t)$, $b_{\bar{\varepsilon}\phi,0,i}(t)$, $b_{\bar{\varepsilon}\phi,i,j}(t)$ remain 0, for all $t \geq 0$, and $\bar{\varepsilon} = \varepsilon$. This yields $\dot{\varepsilon} = A_k\varepsilon + \Phi^\top(y)\Delta_a$, which is the simplified case that has been dealt with in [16].*

Similarly to the description of the ISS inverse dynamics, it is desirable for the subsequent design that the state estimation error dynamics be ISS. Moreover, instead of assuming such a property, one can enforce ISS of the $\bar{\varepsilon}$-dynamics by designing the *K-filters*.

Proposition 5.2. *The state estimation error dynamics are ISS with respect to the inputs $\bar{x}_{\rho+1}$, y, $\phi_{0,i}(y)$, $\phi_{i,j}(y)$, $i = 1, \dots, n$, $j = 1, \dots, q$, if the vector of filter gains is given by $k \triangleq \frac{1}{2}X_{\bar{\varepsilon}}e_1$, where $X_{\bar{\varepsilon}} = X_{\bar{\varepsilon}}^\top \succ 0$ satisfies the Riccati inequality*

$$SX_{\bar{\varepsilon}} + X_{\bar{\varepsilon}}S^\top - X_{\bar{\varepsilon}}(e_1e_1^\top - \gamma_{\bar{\varepsilon}}^{-1}I)X_{\bar{\varepsilon}} + Q_{\bar{\varepsilon}} \preceq 0, \tag{5.104}$$

where

$$Q_{\bar{\varepsilon}} = \left(\frac{\delta_{\bar{\Delta}_b}}{\epsilon_{\bar{\Delta}_b}} + \frac{\delta_{b_{\bar{\varepsilon}y}}}{\epsilon_{b_{\bar{\varepsilon}y}}} + \sum_{i=1}^{n}\frac{\delta_{b_{\bar{\varepsilon}\phi,0,i}}}{\epsilon_{b_{\bar{\varepsilon}\phi,0,i}}} + \sum_{i=1}^{n}\sum_{j=1}^{q}\frac{\delta_{b_{\bar{\varepsilon}\phi,i,j}}}{\epsilon_{b_{\bar{\varepsilon}\phi,i,j}}}\right)I, \tag{5.105}$$

and $\epsilon_{(\cdot)} > 0$. Moreover, there exists an ISS Lyapunov function $V_{\bar{\varepsilon}} \triangleq \gamma_{\bar{\varepsilon}}|\bar{\varepsilon}|_{P_{\bar{\varepsilon}}}^2$, with $P_{\bar{\varepsilon}} \triangleq X_{\bar{\varepsilon}}^{-1}$ and the time derivative of $V_{\bar{\varepsilon}}$ along the trajectories of the state estimation error dynamics *satisfies the inequality*

$$\begin{aligned}\dot{V}_{\bar{\varepsilon}} \leq\ & -|\varepsilon|^2 + \epsilon_{b_{\bar{\varepsilon}y}}\delta_{b_{\bar{\varepsilon}y}} y^2 + \sum_{i=1}^{n} \epsilon_{b_{\bar{\varepsilon}\phi,0,i}}\delta_{b_{\bar{\varepsilon}\phi,0,i}}\phi_{0,i}^2(y) \\ & + \sum_{i=1}^{n}\sum_{j=1}^{q} \epsilon_{b_{\bar{\varepsilon}\phi,i,j}}\delta_{b_{\bar{\varepsilon}\phi,i,j}}\phi_{i,j}^2(y) + \epsilon_{\bar{\Delta}_b}\delta_{\bar{\Delta}_b}\bar{x}_{\rho+1}^2, \end{aligned} \tag{5.106}$$

where $\epsilon_{(\cdot)} > 0$*, or in a more compact (yet more conservative) form,*

$$\begin{aligned}\dot{V}_{\bar{\varepsilon}} \leq\ & -|\bar{\varepsilon}|^2 + \sigma_{\bar{\varepsilon}y} y^2 + \sigma_{\bar{\varepsilon}\phi_0}|\phi_0(y)|^2 + \sigma_{\bar{\varepsilon}\Phi}|\Phi(y)|_{\mathrm{F}}^2 \\ & + \sigma_{\bar{\varepsilon}\bar{x}_{\rho+1}}\bar{x}_{\rho+1}^2, \end{aligned} \tag{5.107}$$

for some constant $\sigma_{(\cdot)} > 0$*.*

Proof. Taking the time derivative of $V_{\bar{\varepsilon}} = \gamma_{\bar{\varepsilon}}|\bar{\varepsilon}|_{P_{\bar{\varepsilon}}}^2$ along the trajectories of system (5.103) yields

$$\begin{aligned}\dot{V}_{\bar{\varepsilon}} =\ & 2\gamma_{\bar{\varepsilon}}\bar{\varepsilon}^\top P_{\bar{\varepsilon}}\Bigg(A_k\bar{\varepsilon} + b_{\bar{\varepsilon}y(t)}y + \sum_{i=1}^{n} b_{\bar{\varepsilon}\phi,0,i}(t)\phi_{0,i}(y) \\ & + \sum_{i=1}^{n}\sum_{j=1}^{q} b_{\bar{\varepsilon}\phi,i,j}(t)\phi_{i,j}(y) - \bar{\Delta}_b\bar{x}_{\rho+1}\Bigg) \\ \leq\ & \gamma_{\bar{\varepsilon}}\Big(\bar{\varepsilon}^\top(P_{\bar{\varepsilon}}S + S^\top P_{\bar{\varepsilon}} - e_1e_1^\top + P_{\bar{\varepsilon}}Q_{\bar{\varepsilon}}P_{\bar{\varepsilon}})\bar{\varepsilon}\Big) \\ & + \epsilon_{b_{\bar{\varepsilon}y}}\delta_{b_{\bar{\varepsilon}y}} y^2 + \sum_{i=1}^{n} \epsilon_{b_{\bar{\varepsilon}\phi,0,i}}\delta_{b_{\bar{\varepsilon}\phi,0,i}}\phi_{0,i}^2(y) \\ & + \sum_{i=1}^{n}\sum_{j=1}^{q} \epsilon_{b_{\bar{\varepsilon}\phi,i,j}}\delta_{b_{\bar{\varepsilon}\phi,i,j}}\phi_{i,j}^2(y) + \epsilon_{\bar{\Delta}_b}\delta_{\bar{\Delta}_b}\bar{x}_{\rho+1}^2. \end{aligned} \tag{5.108}$$

Left-multiplying and right-multiplying by $P_{\bar{\varepsilon}}$ on both sides of (5.104) yields

$$P_{\bar{\varepsilon}}S + S^\top P_{\bar{\varepsilon}} - (e_1e_1^\top - \gamma_{\bar{\varepsilon}}^{-1}I) + P_{\bar{\varepsilon}}Q_{\bar{\varepsilon}}P_{\bar{\varepsilon}} \preceq 0 \tag{5.109}$$

or, equivalently,

$$P_{\bar{\varepsilon}}S + S^\top P_{\bar{\varepsilon}} - e_1e_1^\top + P_{\bar{\varepsilon}}Q_{\bar{\varepsilon}}P_{\bar{\varepsilon}} \preceq -\gamma_{\bar{\varepsilon}}^{-1}I. \tag{5.110}$$

Substituting (5.110) into (5.108) yields (5.106). Finally, defining $\sigma_{\bar{\varepsilon}y} = \epsilon_{b_{\bar{\varepsilon}y}}\delta_{b_{\bar{\varepsilon}y}}$, $\sigma_{\bar{\varepsilon}\phi_0} =$

$\max_i \epsilon_{b_{\bar{\varepsilon}\phi,0,i}} \delta_{b_{\bar{\varepsilon}\phi,0,i}}$, $\sigma_{\bar{\varepsilon}\phi_0} = \max_{i,j} \epsilon_{b_{\bar{\varepsilon}\phi,i,j}} \delta_{b_{\bar{\varepsilon}\phi,i,j}}$, and $\sigma_{\bar{\varepsilon}\bar{x}_{\rho+1}} = \epsilon_{\bar{\Delta}_b} \delta_{\bar{\Delta}_b}$ yields (5.107), which completes the proof. □

Remark 5.10. *In practice $Q_{\bar{\varepsilon}}$ is tuned, to achieve better filtering performance, rather than computed analytically. This is feasible since there exist $\epsilon_{(\cdot)}$ for any bounded $\delta_{(\cdot)}$ such that $Q_{\bar{\varepsilon}}$ can be set to an arbitrary positive multiple of I, due to* (5.105)*. Moreover, $\epsilon_{(\cdot)}$ and $\delta_{(\cdot)}$ do not affect the controller design, as the $\sigma_{(\cdot)}$-related terms in* (5.107) *are dominated adaptively as shown in the next subsection. In this sense, neither $\epsilon_{(\cdot)}$ nor $\delta_{(\cdot)}$ are implemented or need to be known.*

5.2.4 Controller Design

In Sections 5.2.2 and 5.2.3 we have established the ISS property of the inverse dynamics and the state estimation error dynamics. However, before proceeding to design the controller, we have to consider (5.92) in the new coordinates. Note that ε_2 can be written as

$$\varepsilon_2 = \bar{\varepsilon}_2 + a_{\varepsilon_2 y^{(1)}}(t)\dot{y} + Y_{\varepsilon_2}(y), \tag{5.111}$$

where $Y_{\varepsilon_2}(y) \triangleq a_{\varepsilon_2 y}(t)y + \sum_{i=1}^n a_{\varepsilon_2\phi,0,i}(t)\phi_{0,i}(y) + \sum_{i=1}^n \sum_{j=1}^q a_{\varepsilon_2\phi,i,j}(t)\phi_{i,j}(y)$ and $a_{\varepsilon_2 y^{(1)}}(t) = \frac{\Delta_{b_m}(t)}{b_m(t)}$. If the counterpart of Assumption 5.2 holds, that is, either $\rho = 1$ or $\rho \geq 2$ and b_m is constant, then $a_{\varepsilon_2\phi,0,i}(t) = 0$, for all $t \geq 0$. These two special cases have been discussed in [18] as well as in Section 5.1 of the thesis. In general, $a_{\varepsilon_2 y^{(1)}}(t) \neq 0$ and, as a result, ε_2 contains $\dot{y}$. Substituting (5.111) into the first equation of (5.92) yields

$$(1 - a_{\varepsilon_2 y^{(1)}})\dot{y} = \omega_0 + \bar{\omega}^\top \ell_\theta + \ell_{b_m} v_{m,2} + \bar{\varepsilon}_2 + Y_{\varepsilon_2}. \tag{5.112}$$

Noting that $\frac{1}{1-a_{\varepsilon_2 y^{(1)}}} = \frac{b_m}{b_m - \Delta_{b_m}} = \frac{b_m}{\ell_{b_m}}$, we can write the dynamics of y as

$$\dot{y} = \frac{b_m(t)}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top \left(\frac{b_m(t)}{\ell_{b_m}} \ell_\theta \right) + b_m(t) v_{m,2}. \tag{5.113}$$

Observe that the effect of the $a_{\varepsilon_2 y^{(1)}}(t)\dot{y}$ term is to bring the time-varying parameters back to the dynamics of y, which requires the *congelation of variables* method again. To do this, we need first to augment the system (5.92) with the ξ, Ξ and v-dynamics, which are not needed in the classical constant parameter scenarios but necessary in the current setup.

It turns out that the extended system is in the so-called *parametric block-strict-feedback* form [75, Section 2.3.3], described by the equations

$$\dot{\xi} = A_k\xi + ky + \phi_0(y), \tag{5.114}$$

$$\dot{\Xi}^\top = A_k\Xi^\top + \Phi^\top(y), \tag{5.115}$$

$$\begin{aligned} \dot{y} &= \frac{b_m(t)}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top\left(\frac{b_m(t)}{\ell_{b_m}}\ell_\theta\right) + b_m(t)v_{m,2}, \\ \dot{v}_{0,2} &= v_{1,2}, \\ &\vdots \\ \dot{v}_{m-1,2} &= v_{m,2}, \end{aligned} \tag{5.116}$$

$$\begin{aligned} \dot{v}_{m,2} &= -\,k_1 v_{m,1} + v_{m,3}, \\ &\vdots \\ \dot{v}_{m,\rho-1} &= -\,k_{\rho-1}v_{m,1} + v_{m,\rho}, \\ \dot{v}_{m,\rho} &= -\,k_\rho v_{m,1} + v_{m,\rho+1} + g(y)u, \end{aligned} \tag{5.117}$$

and recall that $\omega_0 = \phi_{0,1} + \xi_2$ and $\bar{\omega} = [0, v_{m-1,2}, \ldots, v_{0,2}, (\Phi)_1^\top + (\Xi)_2^\top]^\top$. In these equations, (5.114) and (5.115) describe the state evolution of the filters of the regressors; equations (5.117) give the integrator-chain structure used for backstepping; and equations (5.116) are the key part of the design that contains the dynamics of the output y. As demonstrated in the preceding discussions, the *congelation of variables* method requires an ISS or ISS-like property of each subsystem in the analysis (described by their dissipation inequalities), and then applying a small-gain-like analysis for the overall system to conclude boundedness and convergence properties. As a preparation for the backstepping design on the chain of integrators described by (5.117), one first needs to establish ISS properties for the filter dynamics (5.114), (5.115), and the zero dynamics of (5.116). For the subsystems described by (5.114) and (5.115) we have the following result.

Lemma 5.1. *Let the filter gain k be as in Proposition 5.2. Then, the system* (5.114) *is ISS with respect to the inputs y, $\phi_{0,i}(y)$ and the system* (5.115) *is ISS with respect to the inputs $\phi_{i,j}(y)$, where $i = 1, \ldots, n$, $j = 1, \ldots, q$. Moreover, there exist two ISS Lyapunov functions $V_\xi \triangleq |\xi|^2_{P_\xi}$, $V_\Xi \triangleq \mathrm{tr}(\Xi P_\Xi \Xi^\top)$, with $P_\xi = P_\Xi \triangleq \gamma_{\bar{\varepsilon}} P_{\bar{\varepsilon}} \succ 0$ and $\gamma_{\bar{\varepsilon}} > 0$, such that* *the time derivative of* V_ξ *along the trajectories of* (5.114) *satisfies*

$$\dot{V}_\xi \leq \; -|\xi|^2 + \sigma_{\xi y} y^2 + \sigma_{\xi \phi_0} |\phi_0(y)|^2 \tag{5.118}$$

and the time derivative of V_Ξ *along the trajectories of* (5.115) *satisfies*

$$\dot{V}_\Xi \leq \; -|\Xi|_{\mathrm{F}}^2 + \sigma_{\Xi\Phi} |\Phi(y)|_{\mathrm{F}}^2, \tag{5.119}$$

for some constant $\sigma_{(\cdot)} > 0$.

Proof. Noting (5.104) and the fact that $P_\xi = \gamma_{\bar{\varepsilon}} P_{\bar{\varepsilon}}$ yields

$$A_k^\top P_\xi + P_\xi A_k \preceq -I - \gamma_{\bar{\varepsilon}} P_{\bar{\varepsilon}} Q_{\bar{\varepsilon}} P_{\bar{\varepsilon}}. \tag{5.120}$$

Define $\bar{Q}_\xi = \gamma_{\bar{\varepsilon}} P_{\bar{\varepsilon}} Q_{\bar{\varepsilon}} P_{\bar{\varepsilon}} \succ 0$, take the time derivative of $V_\xi = |\xi|_{P_\xi}^2$ along the trajectories of (5.115), and invoke Young's inequality to obtain

$$\begin{aligned} \dot{V}_\xi &= \xi^\top (A_k^\top P_\xi + P_\xi A_k)\xi + 2\xi^\top P_\xi (ky + \phi_0) \\ &= \; -\xi^\top (I + \bar{Q}_\xi)\xi + 2\xi^\top P_\xi (ky + \phi_0) \\ &\leq \; -|\xi|^2 + \sigma_{\xi y} y^2 + \sigma_{\xi\phi_0} |\phi_0(y)|^2. \end{aligned} \tag{5.121}$$

Similarly, we take the time derivative of $V_\Xi = \mathrm{tr}(\Xi P_\Xi \Xi^\top)$ along the trajectories of (5.115), which yields

$$\begin{aligned} \dot{V}_\Xi = & \sum_{i=1}^{n} ((\Xi^\top)_i^\top (A_k^\top P_\Xi + P_\Xi A_k)(\Xi^\top)_i + 2(\Xi^\top)_i P_\Xi (\Phi^\top)_i) \\ = & \sum_{i=1}^{n} ((\Xi^\top)_i^\top (I + \bar{Q}_\Xi)(\Xi^\top)_i + 2(\Xi^\top)_i P_\Xi (\Phi^\top)_i) \\ \leq & \; |\Xi|_{\mathrm{F}}^2 + \sigma_{\Xi\Phi} |\Phi(y)|_{\mathrm{F}}^2, \end{aligned} \tag{5.122}$$

where $P_\Xi = P_\xi$, $\bar{Q}_\Xi = \bar{Q}_\xi$, and this completes the proof. □

The next step is to establish ISS for the inverse dynamics of (5.116). To derive the inverse dynamics, let

$$
\begin{aligned}
v_{m,2} &= \frac{1}{b_m}\dot{y} - \frac{1}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) - \frac{1}{\ell_{b_m}}\bar{\omega}^\top \ell_\theta \\
&= -\frac{\ell_{b_{m-1}}}{\ell_{b_m}} v_{m-1,2} - \cdots - \frac{\ell_{b_0}}{\ell_{b_m}} v_{0,2} + \frac{1}{b_m}\dot{y} \\
&\quad - ((\Xi)_2^\top + (\Phi)_1^\top)\ell_a - \frac{1}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2)
\end{aligned}
\tag{5.123}
$$

and then define the change of coordinates: $\bar{v}_{0,2} = \bar{v}_{0,2}, \ldots, \bar{v}_{m-2,2} = v_{m-2,2}, \bar{v}_{m-1,2} = v_{m-1,2} - \frac{1}{b_m} y$. The inverse dynamics of (5.116) are then described by

$$
\dot{\bar{v}} = A_{\ell_{\bar{b}}}\bar{v} + g_{\bar{v}}(y, \xi, \Xi, \bar{\varepsilon}_2, t), \tag{5.124}
$$

where $A_{\ell_{\bar{b}}} \triangleq S - e_m \ell_{\bar{b}}^\top$, $\ell_{\bar{b}} \triangleq [\frac{\ell_{b_0}}{\ell_{b_m}}, \ldots, \frac{\ell_{b_{m-1}}}{\ell_{b_m}}]^\top$, and $g_{\bar{v}}(y, \xi, \Xi, \bar{\varepsilon}_2, t) \triangleq [0, \ldots, 0, \frac{1}{b_m} y, -(\frac{\ell_{b_{m-1}}}{b_m \ell_{b_m}} + (\frac{1}{b_m})^{(1)})y - ((\Xi)_2^\top + (\Phi)_2^\top)\ell_a - \frac{1}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2)]^\top$. Exploiting the flexibility of the *congelation of variables* method we can always select ℓ_b to construct a Hurwitz $A_{\ell_{\bar{b}}}$, and therefore ISS of system (5.124) can be established as shown in the lemma that follows.

Lemma 5.2. *Consider a vector of congealed parameters $\ell_b \triangleq [\ell_{b_m}, \ldots, \ell_{b_0}]^\top$ such that the polynomial $\ell_{b_m} s^m + \ell_{b_{m-1}} s^{m-1} + \cdots + \ell_{b_0}$ is Hurwitz. Then, the system* (5.124) *is ISS with respect to the inputs y, $\phi_{0,i}(y)$, $\phi_{i,j}(y)$, ξ_2, $(\Xi)_{j2}$ and $\bar{\varepsilon}_2$, where $i = 1, \ldots, n$, $j = 1, \ldots, q$. Moreover, there is an ISS Lyapunov function $V_{\bar{v}} = |\bar{v}|^2_{P_{\bar{v}}}$, with $P_{\bar{v}} = P_{\bar{v}}^\top \succ 0$, such that the time derivative of $V_{\bar{v}}$ along the trajectories of* (5.124) *satisfies*

$$
\begin{aligned}
\dot{V}_{\bar{v}} \leq &-|\bar{v}|^2 + \sigma_{\bar{v}y} y^2 + \sigma_{\bar{v}\phi_0}|\phi_0(y)|^2 + \sigma_{\bar{v}\Phi}|\Phi(y)|^2_{\mathrm{F}} \\
&+ \sigma_{\bar{v}\xi_2}\xi_2^2 + \sigma_{\bar{v}(\Xi)_2}|(\Xi)_2|^2 + \sigma_{\bar{v}\bar{\varepsilon}_2}\bar{\varepsilon}_2^2,
\end{aligned}
\tag{5.125}
$$

where $\sigma_{(\cdot)} > 0$ are constant.

Proof. Since $\ell_{b_m} s^m + \ell_{b_{m-1}} s^{m-1} + \cdots + \ell_{b_0}$ is Hurwitz, $A_{\ell_{\bar{b}}} = S - e_m \ell_{\bar{b}}^\top$ is also Hurwitz, and therefore there exist $P_{\bar{v}} = P_{\bar{v}}^\top \succ 0$ and $Q_{\bar{v}} = Q_{\bar{v}}^\top \succ 0$ such that $A_{\ell_{\bar{b}}}^\top P_{\bar{v}} + P_{\bar{v}} A_{\ell_{\bar{b}}} + Q_{\bar{v}} = 0$. Without loss of generality we assume that $Q_{\bar{v}} = I + \bar{Q}_{\bar{v}}$, where $\bar{Q}_{\bar{v}} = \bar{Q}_{\bar{v}}^\top \succ 0$. This condition can always be satisfied by scaling $P_{\bar{v}}$. Taking the time derivative of $V_{\bar{v}}$ along the trajectories of (5.124) yields

$$\begin{aligned}\dot{V}_{\bar{v}} &= \bar{v}^{\top}(A_{\ell_{\bar{b}}}^{\top}P_{\bar{v}} + P_{\bar{v}}A_{\ell_{\bar{b}}})\bar{v} + \bar{v}^{\top}P_{\bar{v}}g_{\bar{v}} + g_{\bar{v}}P_{\bar{v}}\bar{v}^{\top} \\ &= -\bar{v}^{\top}(I + \bar{Q}_{\bar{v}})\bar{v} + 2\bar{v}^{\top}P_{\bar{v}}g_{\bar{v}}. \end{aligned} \tag{5.126}$$

Note that in $g_{\bar{v}}(y, \xi, \Xi, \bar{\varepsilon}_2, t)$ all the coefficients coupled with the "inputs" y, $\phi_{0,i}(y)$, $\phi_{i,j}(y)$, ξ_2, $(\Xi)_{j2}$, $\bar{\varepsilon}_2$ are bounded. Thus, by Lemma A.1, the condition

$$\begin{aligned}&-\bar{v}^{\top}\bar{Q}_{\bar{v}}\bar{v} + 2\bar{v}^{\top}P_{\bar{v}}g_{\bar{v}} - (\sigma_{\bar{v}y}y^2 + \sigma_{\bar{v}\phi_0}|\phi_0(y)|^2 \\ &+ \sigma_{\bar{v}\Phi}|\Phi(y)|_{\mathrm{F}}^2 + \sigma_{\bar{v}\xi_2}\xi_2^2 + \sigma_{\bar{v}(\Xi)_2}|(\Xi)_2|^2 + \sigma_{\bar{v}\bar{\varepsilon}_2}\bar{\varepsilon}_2^2) \leq 0 \end{aligned} \tag{5.127}$$

holds for some constant $\sigma_{(\cdot)} > 0$. Substituting (5.127) into (5.126) yields (5.125), which completes the proof. □

Having established the ISS properties of (5.114), (5.115) and the zero dynamics of (5.116), we proceed to the *backstepping* design on the chain of integrators (5.117). Define the error variables

$$z_1 = y, \tag{5.128}$$

$$z_i = v_{m,i} - \alpha_{i-1}, \quad i = 2, \ldots, \rho, \tag{5.129}$$

the tuning functions

$$\tau_1 = (\omega - \hat{\varrho}\bar{\alpha}_1 e_1)z_1, \tag{5.130}$$

$$\tau_i = \tau_{i-1} - \frac{\partial \alpha_{i-1}}{\partial y}\omega z_i, \quad i = 2, \ldots, \rho, \tag{5.131}$$

the virtual control laws

$$\alpha_1 = \hat{\varrho}\bar{\alpha}_1 = -\hat{\varrho}\kappa z_1, \tag{5.132}$$

$$\alpha_2 = -\hat{b}_m z_1 - (c_2 + \zeta_2) z_2 + \beta_2 + \frac{\partial \alpha_1}{\partial \hat{\theta}} \Gamma_\theta \tau_2, \tag{5.133}$$

$$\begin{aligned} \alpha_i = & -z_{i-1} - (c_i + \zeta_i) z_i + \beta_i + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma_\theta \tau_i \\ & - \sum_{j=2}^{i-1} \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}} \Gamma_\theta \frac{\partial \alpha_{i-1}}{\partial y} \omega z_j, \quad i = 3, \ldots, \rho, \end{aligned} \tag{5.134}$$

with

$$\kappa = c_1 + \frac{\epsilon_{\hat{\theta}}}{2} |\hat{\theta}|^2 + \hat{\zeta}_y + \hat{\zeta}_{\phi 0} |\bar{\phi}_0(y)|^2 + \hat{\zeta}_\Phi |\bar{\Phi}(y)|_{\mathrm{F}}^2, \tag{5.135}$$

$$\begin{aligned} \zeta_2 = & \frac{1}{2\epsilon_{\hat{\theta}}} + \frac{\rho \delta_{\Delta_{b_m}}}{2\epsilon_{\Delta_{b_m}}} + \frac{1}{2} \left(\frac{\partial \alpha_1}{\partial y} \right)^2 \times \\ & \left(\epsilon_{\Delta_{b_m}} \delta_{\Delta_{b_m}} (\hat{\varrho}^2 \kappa^2 + 1) + \epsilon_{\Delta_{\bar{\theta}}} \delta_{\Delta_{\bar{\theta}}} + \epsilon_{Y_{\varepsilon_2}} + \epsilon_{\bar{\varepsilon}_2} \right), \end{aligned} \tag{5.136}$$

$$\begin{aligned} \zeta_i = & \frac{1}{2} \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 \times \\ & \left(\epsilon_{\Delta_{b_m}} \delta_{\Delta_{b_m}} (\hat{\varrho}^2 \kappa^2 + 1) + \epsilon_{\Delta_{\bar{\theta}}} \delta_{\Delta_{\bar{\theta}}} + \epsilon_{Y_{\varepsilon_2}} + \epsilon_{\bar{\varepsilon}_2} \right), \end{aligned} \tag{5.137}$$

$$\begin{aligned} \beta_i = & \frac{\partial \alpha_{i-1}}{\partial y} (\omega_0 + \omega^\top \hat{\theta}) + \frac{\partial \alpha_{i-1}}{\partial \xi} (A_k \xi + k y + \phi_0) \\ & + \sum_{j=1}^{q} \frac{\partial \alpha_{i-1}}{\partial (\Xi^\top)_j} (A_k (\Xi^\top)_j + (\Phi^\top)_j) + k_i v_{m,1} \\ & + \sum_{j=1}^{m+i-1} \frac{\partial \alpha_{i-1}}{\partial \lambda_j} (-k_j \lambda_1 + \lambda_{j+1}) + \frac{\partial \alpha_{i-1}}{\partial \hat{\varrho}} \dot{\hat{\varrho}} \\ & + \frac{\partial \alpha_{i-1}}{\partial \hat{\zeta}_y} \dot{\hat{\zeta}}_y + \frac{\partial \alpha_{i-1}}{\partial \hat{\zeta}_{\phi_0}} \dot{\hat{\zeta}}_{\phi_0} + \frac{\partial \alpha_{i-1}}{\partial \hat{\zeta}_\Phi} \dot{\hat{\zeta}}_\Phi, \\ & i = 2, \ldots, \rho, \end{aligned} \tag{5.138}$$

the control law

$$u = \frac{1}{g(y)} (\alpha_\rho - v_{m,\rho+1}), \tag{5.139}$$

and the parameter update laws

$$\dot{\hat{\varrho}} = \gamma_{\varrho}\mathrm{sgn}(\ell_{b_m})\kappa z_1^2, \tag{5.140}$$

$$\dot{\hat{\zeta}}_y = \gamma_{\zeta_y} z_1^2, \quad \dot{\hat{\zeta}}_{\phi_0} = \gamma_{\zeta_{\phi_0}} |\phi_0|^2, \quad \dot{\hat{\zeta}}_{\Phi} = \gamma_{\zeta_\Phi} |\Phi|_{\mathrm{F}}^2, \tag{5.141}$$

$$\dot{\hat{\theta}} = \Gamma_\theta \tau_\rho, \tag{5.142}$$

where $c_i > 0$, $i = 1, \ldots, \rho$, $\epsilon_{(\cdot)} > 0$, $\gamma_{(\cdot)} > 0$, $\Gamma_\theta = \Gamma_\theta^\top \succ 0$, $\bar{\theta}(t) \triangleq \frac{b_m(t)}{\ell_{b_m}}\ell_\theta$, and $\Delta_{\bar{\theta}} \triangleq \bar{\theta}(t) - \ell_\theta$. In the definition of κ, $\bar{\phi}_0(y)$, $\bar{\Phi}(y)$ are defined such that $\phi_0(y) = \bar{\phi}_0(y)y$, $\Phi(y) = \bar{\Phi}(y)y$, which is feasible due to Remark 5.6. Moreover, the initial value of the parameter estimates are selected such that $\hat{\varrho}(0) > 0$, $\hat{\zeta}_{(\cdot)}(0) > 0$.

Remark 5.11. *Compared to the I&I scheme introduced in Section 5.1 in which "computed" damping term are adopted, in the passivity-based scheme dynamically updated "estimates" $\hat{\zeta}_{(\cdot)}$ are adopted as the coefficients of the additional damping terms, since the required damping coefficients are in general difficult to compute. Meanwhile, thanks to these adaptive damping terms, we do not need to know $\delta_{\Delta_{\bar{\theta}}}$ for a reason similar to what is explained in Remark 5.10*

Proposition 5.3. *Consider the adaptive controller described by equations* (5.128)–(5.142) *for the system described by equations* (5.114)–(5.117) *and suppose that Assumptions 1.2, 1.3, and 5.3 hold. Then, the closed-loop signals z, $\bar{x}$, $\bar{\varepsilon}$, ξ, Ξ, $\bar{v}$, $\hat{\theta}$, $\hat{\varrho}$, and $\hat{\zeta}_{(\cdot)}$ are bounded.*

Proof. We first analyze the *backstepping* error variables z_i step by step.

Step 1. Consider the dynamics of z_1, which are described by

$$\begin{aligned}
\dot{z}_1 &= \frac{b_m}{\ell_{b_m}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top \frac{b_m}{\ell_{b_m}}\ell_\theta + b_m v_{m,2} \\
&= (\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top\bar{\theta} + \ell_{b_m} v_{m,2} \\
&\quad + \Delta_{\frac{b_m}{\ell_{b_m}}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top\Delta_{\bar{\theta}} + \Delta_{b_m} v_{m,2} \\
&= (\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top\hat{\theta} + \bar{\alpha}_1 + b_m z_2 \\
&\quad \bar{\omega}^\top(\ell_\theta - \hat{\theta}) - \ell_{b_m}\left(\frac{1}{\ell_{b_m}} - \hat{\varrho}\right)\bar{\alpha}_1 \\
&\quad + \Delta_{\frac{b_m}{\ell_{b_m}}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + \bar{\omega}^\top\Delta_{\bar{\theta}} + \Delta_{b_m}\hat{\varrho}\bar{\alpha}_1,
\end{aligned} \tag{5.143}$$

where $\Delta_{\frac{b_m}{\ell_{b_m}}}(t) \triangleq \frac{b_m(t)}{\ell_{b_m}} - 1$ (recall also that $\bar{\theta}(t) = \frac{b_m(t)}{\ell_{b_m}}\ell_\theta$ and $\Delta_{\bar{\theta}} = \bar{\theta}(t) - \ell_\theta$). Note that $z_2 = v_{m,2} - \hat{\varrho}\bar{\alpha}_1$ and

$$b_m z_2 = \hat{b}_m z_2 + (\ell_{b_m} - \hat{b}_m) z_2 + \Delta_{b_m} z_2, \tag{5.144}$$

which yields

$$\bar{\omega}^\top(\ell_\theta - \hat{\theta}) + b_m z_2 = (\omega - \hat{\varrho}\bar{\alpha}_1 e_1)^\top(\ell_\theta - \hat{\theta}) + \hat{b}_m z_2 + \Delta_{b_m} z_2. \tag{5.145}$$

Considering the function $V_{z_1} \triangleq \frac{1}{2}z_1^2$ and taking the time derivative of V_{z_1} along the trajectories of (5.143) yields

$$\begin{aligned}
\dot{V}_{z_1} = {} & z_1(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) + z_1\bar{\omega}^\top\hat{\theta} + \bar{\alpha}_1 z_1 + \hat{b}_m z_1 z_2 \\
& + \Delta_{\frac{b_m}{\ell_{b_m}}}(\omega_0 + Y_{\varepsilon_2} + \bar{\varepsilon}_2) z_1 + z_1\bar{\omega}^\top\Delta_{\bar{\theta}} + \Delta_{b_m} z_1 z_2 \\
& + z_1(\omega - \hat{\varrho}\bar{\alpha}_1 e_1)^\top(\ell_\theta - \hat{\theta}) - \ell_{b_m}(\ell_{b_m}^{-1} - \hat{\varrho})\bar{\alpha}_1 z_1 \\
& + \Delta_{b_m}\hat{\varrho}\bar{\alpha}_1 z_1.
\end{aligned} \tag{5.146}$$

Invoking Lemma A.1 yields

$$\begin{aligned}
\dot{V}_{z_1} \leq {} & -\kappa z_1^2 + \left(\frac{\epsilon_{\hat{\theta}}}{2}|\hat{\theta}|^2 + \sigma_{z_1 y}\right) z_1^2 + \sigma_{z_1\phi_0}|\phi_0|^2 \\
& + \sigma_{z_1\Phi}|\Phi|_{\mathrm{F}}^2 + \sigma_{z_1\bar{\varepsilon}_2}\bar{\varepsilon}_2^2 + \sigma_{z_1\xi_2}\xi_2^2 + \sigma_{z_1(\Xi)_2}|(\Xi)_2|^2 \\
& + \sigma_{z_1\bar{v}}\bar{v}^2 + \left(\frac{1}{2\epsilon_{\hat{\theta}}} + \frac{\delta_{b_m}}{2\epsilon_{\Delta_{b_m}}}\right) z_2^2 + R_1 - \Delta_{b_m}\hat{\varrho}\bar{\kappa} z_1^2,
\end{aligned} \tag{5.147}$$

where $R_1 \triangleq z_1(\omega - \hat{\varrho}\bar{\alpha}_1 e_1)^\top(\ell_\theta - \hat{\theta}) - \ell_{b_m}(\ell_{b_m}^{-1} - \hat{\varrho})\bar{\alpha}_1 z_1$ consists of the remaining terms to be cancelled by the update law/tuning function design. Moreover, using the same argument as in Section 2.1.2 and Section 4.1, we can show that $-\Delta_{b_m}\hat{\varrho}\bar{\kappa} z_1^2 \leq 0$, and therefore this term can be dropped hereafter.

Step 2,...,ρ. Consider the sum of the functions $V_{z_i} \triangleq \frac{1}{2}z_i^2$, $i = 1, \ldots, \rho$, and take the time derivative of the sum along the trajectories of the system, which yields

$$
\begin{aligned}
\sum_{i=1}^{\rho} \dot{V}_{z_i} = \dot{V}_{z_1} - R_1 - \sum_{i=2}^{\rho} (c_i + \zeta_i) z_i^2 - \sum_{i=2}^{\rho} z_i \frac{\partial \alpha_{i-1}}{\partial y} \times \\
& (\bar{\varepsilon}_2 + Y_{\varepsilon_2} + \bar{\omega}^\top \Delta_{\bar{\theta}} + (\alpha_1 + z_2)\Delta_{b_m}) + R_\rho \\
& \leq \dot{V}_{z_1} - R_1 - \sum_{i=1}^{\rho} (c_i + \zeta_i) z_i^2 + \frac{(\rho - 1)\delta_{\Delta_{b_m}}}{2\epsilon_{\Delta_{b_m}}} z_2^2 \\
& \quad + \sum_{i=2}^{\rho} \frac{1}{2} z_i^2 \left(\frac{\partial \alpha_{i-1}}{\partial y} \right)^2 \times \\
& \quad \left(\epsilon_{\Delta_{b_m}} \delta_{\Delta_{b_m}} (\hat{\varrho}^2 \kappa^2 + 1) + \epsilon_{\Delta_{\bar{\theta}}} \delta_{\Delta_{\bar{\theta}}} + \epsilon_{Y_{\varepsilon_2}} + \epsilon_{\bar{\varepsilon}_2} \right) \\
& \quad + \sum_{i=2}^{\rho} \sigma_{z_i y} z_1^2 + \sum_{i=2}^{\rho} \sigma_{z_i \phi_0} |\phi_0|^2 + \sum_{i=2}^{\rho} \sigma_{z_i \Phi} |\Phi|_{\mathrm{F}}^2 \\
& \quad + \sum_{i=2}^{\rho} \sigma_{z_i \bar{\varepsilon}_2} \bar{\varepsilon}_2^2 + \sum_{i=2}^{\rho} \sigma_{z_i \xi_2} \xi_2^2 + \sum_{i=2}^{\rho} \sigma_{z_i (\Xi)_2} |(\Xi)_2|^2 \\
& \quad + \sum_{i=2}^{\rho} \sigma_{z_i \bar{v}} |\bar{v}|^2 + R_\rho \\
& = \; - \sum_{i=1}^{\rho} c_i z_i^2 - \hat{\zeta}_y z_1^2 - \hat{\zeta}_{\phi_0} |\phi_0|^2 - \hat{\zeta}_\Phi |\Phi|_{\mathrm{F}}^2 \\
& \quad + \sigma_{zy} z_1^2 + \sigma_{z\phi_0} |\phi_0|^2 + \sigma_{z\Phi} |\Phi|_{\mathrm{F}}^2 + \sigma_{z\bar{\varepsilon}_2} \bar{\varepsilon}_2^2 \\
& \quad + \sigma_{z\xi_2} \xi_2^2 + \sigma_{z(\Xi)_2} |(\Xi)_2|^2 + \sigma_{z\bar{v}} |\bar{v}|^2 + R_\rho, \qquad (5.148)
\end{aligned}
$$

where $\sigma_{z(\cdot)} \triangleq \sum_{i=1}^{\rho} \sigma_{z_i(\cdot)} > 0$ and $R_\rho \triangleq \tau_\rho^\top (\ell_\theta - \hat{\theta}) - \ell_{b_m} (\ell_{b_m}^{-1} - \hat{\varrho}) \bar{\alpha}_1 z_1$ consists of the remaining terms to be cancelled by the update laws. Then considering the function $V_z \triangleq \sum_{i=1}^{\rho} V_{z_i} + \frac{1}{2} |\ell_\theta - \hat{\theta}|_{\Gamma^{-1}}^2 + \frac{|\ell_{bm}|}{2\gamma_\varrho} |\ell_{b_m}^{-1} - \hat{\varrho}|^2$ and taking its time derivative along the system trajectories yields

$$
\begin{aligned}
\dot{V}_z \leq & \; - \sum_{i=1}^{\rho} c_i z_i^2 - \hat{\zeta}_y z_1^2 - \hat{\zeta}_{\phi_0} |\phi_0|^2 - \hat{\zeta}_\Phi |\Phi|_{\mathrm{F}}^2 \\
& + \sigma_{zy} z_1^2 + \sigma_{z\phi_0} |\phi_0|^2 + \sigma_{z\Phi} |\Phi|_{\mathrm{F}}^2 + \sigma_{z\bar{\varepsilon}_2} \bar{\varepsilon}_2^2 \\
& + \sigma_{z\xi_2} \xi_2^2 + \sigma_{z(\Xi)_2} |(\Xi)_2|^2 + \sigma_{z\bar{v}} |\bar{v}|^2. \qquad (5.149)
\end{aligned}
$$

Consider now the Lyapunov function candidate $V \triangleq V_z + \gamma_{V_{\bar{x}}} V_{\bar{x}} + \gamma_{V_{\bar{\varepsilon}}} V_{\bar{\varepsilon}} + \gamma_{V_\xi} V_\xi + \gamma_{V_\Xi} V_\Xi + \gamma_{V_{\bar{v}}} V_{\bar{v}} + \frac{1}{2\gamma_{\zeta_y}} (\zeta_y - \hat{\zeta}_y)^2 + \frac{1}{2\gamma_{\zeta_{\phi_0}}} (\zeta_{\phi_0} - \hat{\zeta}_{\phi_0})^2 + \frac{1}{2\gamma_{\zeta_\Phi}} (\zeta_\Phi - \hat{\zeta}_\Phi)^2$, where $\gamma_{V_{\bar{x}}} \triangleq \sigma_{\bar{\varepsilon}\bar{x}_{\rho+1}} (\sigma_{z\bar{\varepsilon}_2} + \sigma_{\bar{v}\bar{\varepsilon}_2} \sigma_{z\bar{v}})$, $\gamma_{V_{\bar{\varepsilon}}} \triangleq \sigma_{z\bar{\varepsilon}_2} + \sigma_{\bar{v}\bar{\varepsilon}_2} \sigma_{z\bar{v}}$, $\gamma_{V_\xi} \triangleq \sigma_{z\xi_2} + \sigma_{\bar{v}\xi_2} \sigma_{z\bar{v}}$, $\gamma_{V_\Xi} \triangleq \sigma_{z(\Xi)_2} + \sigma_{\bar{v}(\Xi)_2} \sigma_{z\bar{v}}$, $\gamma_{V_{\bar{v}}} \triangleq \sigma_{z\bar{v}}$ are the scaling coefficients of the corresponding partial Lyapunov function candidate, and $\zeta_y \triangleq \sigma_{zy} + \gamma_{V_{\bar{x}}}\sigma_{\bar{x}y} + \gamma_{V_{\bar{\varepsilon}}}\sigma_{\bar{\varepsilon}y} + \gamma_{V_\xi}\sigma_{\xi y} + \gamma_{V_{\bar{v}}}\sigma_{\bar{v}y}$, $\zeta_{\phi_0} \triangleq \sigma_{z\phi_0} + \gamma_{V_{\bar{x}}}\sigma_{\bar{x}\phi_0} + \gamma_{V_{\bar{\varepsilon}}}\sigma_{\bar{\varepsilon}\phi_0} + \gamma_{V_\xi}\sigma_{\xi\phi_0} + \gamma_{V_{\bar{v}}}\sigma_{\bar{v}\phi_0}$, $\zeta_\Phi \triangleq \sigma_{z\Phi} + \gamma_{V_{\bar{x}}}\sigma_{\bar{x}\Phi} + \gamma_{V_{\bar{\varepsilon}}}\sigma_{\bar{\varepsilon}\Phi} + \gamma_{V_\Xi}\sigma_{\Xi\Phi} + \gamma_{V_{\bar{v}}}\sigma_{\bar{v}\Phi}$ are the required damping coefficients to be compensated by $\hat{\zeta}_y$, $\hat{\zeta}_{\phi_0}$, and $\hat{\zeta}_\Phi$, respectively. Taking the time derivative of V along the trajectories of the system yields

$$
\begin{aligned}
\dot{V} \leq & -\sum_{i=1}^{\rho} c_i z_i^2 + (\zeta_y - \hat{\zeta}_y)(z_1^2 - \gamma_{\zeta_y}^{-1}\dot{\hat{\zeta}}_y) \\
& + (\zeta_{\phi_0} - \hat{\zeta}_{\phi_0})(|\phi_0|^2 - \gamma_{\zeta_{\phi_0}}^{-1}\dot{\hat{\zeta}}_{\phi_0}) \\
& + (\zeta_\Phi - \hat{\zeta}_\Phi)(|\Phi|_{\mathrm{F}}^2 - \gamma_{\zeta_\Phi}^{-1}\dot{\hat{\zeta}}_\Phi) = -\sum_{i=1}^{\rho} c_i z_i^2 \leq 0.
\end{aligned}
\tag{5.150}
$$

Hence z, $\bar{x}$, $\bar{\varepsilon}$, ξ, Ξ, $\bar{v}$, $\hat{\theta}$, $\hat{\varrho}$, and $\hat{\zeta}_{(\cdot)}$ are bounded, by a standard Lyapunov analysis, which completes the proof. □

We should not forget that the invariance-like proof of asymptotic output regulation requires boundedness of ε. In Proposition 5.3 we have proved the boundedness of $\bar{\varepsilon}$ after the change of coordinates described by Algorithm 5.3. However, it is not easy to directly prove the boundedness of ε since Algorithm 5.3 involves the time derivatives of y, $\phi_{0,i}(y)$, and $\phi_{i,j}(y)$, $i = 1, \ldots, n$, $j = 1, \ldots, q$, the boundedness of which is difficult to conclude. Recall that these time derivatives are present because u has to be decomposed at the design stage with the help of the inverse dynamics. Now that we have completed the design, it is more convenient to directly use the boundedness of u for concluding the boundedness of ε, provided that we can first prove the boundedness of λ, as shown in what follows.

Theorem 5.4. *Consider the system described by the equations* (5.114)–(5.117)*, with the same assumptions as in Proposition 5.3, and the adaptive controller described by the equations* (5.128)–(5.142)*. Then, all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} y(t) = 0$.

Proof. First note that boundedness of ξ, Ξ, $\hat{\theta}$, $\hat{\varrho}$, and $\hat{\zeta}_{(\cdot)}$ is guaranteed by Proposition 5.3, and therefore, the rest of the proof is devoted to establish boundedness of λ, ε, and x.

Recalling (5.89) yields

$$\begin{bmatrix} v_{0,2} \\ \vdots \\ v_{m,2} \\ \vdots \\ v_{m,\rho} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ * & * & * & \ddots & 0 \\ * & * & * & * & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}, \tag{5.151}$$

where "$*$" represents terms that do not affect the subsequent analysis. Note that $v_{0,2} = \lambda_2$ is bounded due to boundedness of $\bar{v}$ and thus λ_2 is also bounded. Note that by Vieta's formula, $-k_1$ is the sum of the roots of the characteristic polynomial of A_k, which is also the trace of A_k. Thus,

$$k_1 = -\mathrm{tr}(A_k) \tag{5.152}$$

and since A_k is Hurwitz, $\mathrm{tr}(A_k) < 0$. Hence $k_1 > 0$. Consider the dynamics of the first state variable of the input filter (5.85), that is, $\dot{\lambda}_1 = -k_1\lambda_1 + \lambda_2$ with a bounded input λ_2. Thus λ_1 is also bounded due to the bounded-input bounded-output property. Rewriting (5.151) yields

$$\begin{bmatrix} v_{0,2} \\ \vdots \\ v_{m,2} \\ \vdots \\ v_{m,\rho} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \ddots & 0 \\ * & * & \ddots & 0 \\ * & * & * & 1 \end{bmatrix} \begin{bmatrix} \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} + \begin{bmatrix} 0 \\ * \\ * \\ * \end{bmatrix} \lambda_1. \tag{5.153}$$

Since $\bar{v}$, λ_1, and z_1 (or y) are bounded, we can conclude that $\lambda_3, \ldots, \lambda_{m+1}$ are bounded by exploiting the lower-triangular structure of the matrix in (5.153). Since $\lambda_1, \ldots, \lambda_{m+1}$ are bounded, by Proposition 5.3, α_1 is bounded. Note that z_2 is bounded, thus $v_{m,2}$ is bounded, which further guarantees boundedness of λ_{m+2} due to (5.153). In the same spirit, boundedness of $\lambda_{m+3}, \ldots, \lambda_n$ can be established in a recursive way similar to the procedures in [75, Section 8.1.3], which proves that λ is bounded.

To proceed, combining boundedness of λ and Proposition 5.3 yields boundedness of α_ρ and $v_{m,\rho+1}$, and therefore boundedness of u. This further proves boundedness of ε along with boundedness of y due to (5.91). Since $x = \xi + \Omega^\top \ell_\theta + \varepsilon$, we can conclude that x is also bounded. Hence all trajectories of the closed-loop system are bounded.

Finally, recall (5.150) and note that $\dot{z}$ is bounded due to boundedness of the system parameters and all other closed-loop state variables, hence invoking Lemma A.5 yields $\lim_{t\to+\infty} z(t) = 0$, and also $\lim_{t\to+\infty} y(t) = 0$, which completes the proof. □

Remark 5.12. *Using the fact that* $\lim_{t\to+\infty} z(t) = 0$ *we can proceed to prove the convergence of* ξ, Ξ, λ, ε *and* x *to 0 by exploiting the converging-input converging-output property of the corresponding subsystems or the dependency on converging signals.*

It is beneficial to note that the constructive proof of Proposition 5.3 can be performed using the notion of *active node*. This alternative view significantly reduces the complexity of the original proof and provides more insight into the challenge of the effects of time-varying parameters, compared to the classical case, as shown below.

Alternative proof of Proposition 5.3. First, repeat the procedures of the original proof up to (5.149) and recall the dissipation inequalities (5.101), (5.107), (5.118), (5.119), and (5.125), from Assumption 5.3, Proposition 5.2, Lemma 5.1, and Lemma 5.2, respectively. Then by defining $\bar{V} = [V_z, V_{\bar{x}}, V_{\bar{\varepsilon}}, V_\xi, V_\Xi, V_{\bar{v}}]^\top$ and $\hat{\psi} = [y^2, \bar{x}^2_{\rho+1}, \bar{\varepsilon}^2_2, \xi^2_2, |(\Xi)_2|^2, |\bar{v}|^2, |\phi_0|^2, |\Phi|^2_{\mathrm{F}}]^\top$, one can write the dissipation inequalities of the subsystems in a compact form, namely,

$$\begin{aligned}\dot{\bar{V}} \leq\; & -\bar{E}\hat{\psi} - e_1 \sum_{i}^{\rho} c_i z_i^2 - e_2 \sum_{\rho+2}^{n} \bar{x}_i^2 - e_3 \left(\bar{\varepsilon}_1^2 + \sum_{3}^{n} \bar{\varepsilon}_i^2 \right) \\ & - e_4 \left(\xi_1^2 + \sum_{3}^{n} \xi_i^2 \right) - e_5 \left(|(\Xi)_1|^2 + \sum_{3}^{n} |(\Xi)_i|^2 \right), \end{aligned} \tag{5.154}$$

where e_i is the ith unit vector and

$$\bar{E} \triangleq \begin{bmatrix} (\hat{\zeta}_y - \sigma_{zy}) & 0 & -\sigma_{\bar{\varepsilon}_2} & -\sigma_{z\xi_2} & -\sigma_{z(\Xi)_2} & -\sigma_{z\bar{v}} & (\hat{\zeta}_{\phi_0} - \sigma_{z\phi_0}) & (\hat{\zeta}_\Phi - \sigma_{z\Phi}) \\ -\sigma_{\bar{x}y} & 1 & 0 & 0 & 0 & 0 & -\sigma_{\bar{x}\phi_0} & -\sigma_{\bar{x}\Phi} \\ -\sigma_{\bar{\varepsilon}y} & -\sigma_{\bar{\varepsilon}\bar{x}_{\rho+1}} & 1 & 0 & 0 & 0 & -\sigma_{\bar{\varepsilon}\phi_0} & -\sigma_{\bar{\varepsilon}\Phi} \\ -\sigma_{\xi y} & 0 & 0 & 1 & 0 & 0 & -\sigma_{\xi\phi_0} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -\sigma_{\Xi\Phi} \\ -\sigma_{\bar{v}y} & 0 & -\sigma_{\bar{v}\bar{\varepsilon}_2} & -\sigma_{\bar{v}\xi_2} & -\sigma_{\bar{v}(\Xi)_2} & 1 & -\sigma_{\bar{v}\phi_0} & -\sigma_{\bar{v}\Phi} \end{bmatrix}. \tag{5.155}$$

The matrix $\bar{E}$ can be augmented using the method in Section 3.3, described by (3.56), namely

$$\hat{E} \triangleq \begin{bmatrix} (\hat{\zeta}_y - \sigma_{zy}) & 0 & -\sigma_{\bar{\varepsilon}_2} & -\sigma_{z\xi_2} & -\sigma_{z(\Xi)_2} & -\sigma_{z\bar{v}} & 0 & 0 \\ -\sigma_{\bar{x}y} & 1 & 0 & 0 & 0 & 0 & -\sigma_{\bar{x}\phi_0} & -\sigma_{\bar{x}\Phi} \\ -\sigma_{\bar{\varepsilon}y} & -\sigma_{\bar{\varepsilon}\bar{x}_{\rho+1}} & 1 & 0 & 0 & 0 & -\sigma_{\bar{\varepsilon}\phi_0} & -\sigma_{\bar{\varepsilon}\Phi} \\ -\sigma_{\xi y} & 0 & 0 & 1 & 0 & 0 & -\sigma_{\xi\phi_0} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -\sigma_{\Xi\Phi} \\ -\sigma_{\bar{v}y} & 0 & -\sigma_{\bar{v}\bar{\varepsilon}_2} & -\sigma_{\bar{v}\xi_2} & -\sigma_{\bar{v}(\Xi)_2} & 1 & -\sigma_{\bar{v}\phi_0} & -\sigma_{\bar{v}\Phi} \\ 0 & 0 & 0 & 0 & 0 & 0 & (\hat{\zeta}_{\phi_0} - \sigma_{z\phi_0}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & (\hat{\zeta}_\Phi - \sigma_{z\Phi}) \end{bmatrix}. \tag{5.156}$$

The first 6×6 submatrix, denoted as E according to the convention in Section 3.3, is directly associated with the primary basis functions y^2, $\bar{x}_{\rho+1}^2$, $\bar{\varepsilon}_2^2$, ξ_2^2, $|(\Xi)_2|^2$, and $|\bar{v}|^2$. The underlying graph of E is depicted by Fig. 5.4, in which the *augmented* vertices associated with the basis functions $|\phi_0|^2$ and $|\Phi|^2$ are omitted for simplicity as they are not contained in any cycles. Moreover, since the damping coefficients $\hat{\zeta}_y$, $\hat{\zeta}_{\phi_0}$, $\hat{\zeta}_\Phi$ are adaptively adjusted (which makes node z an *active node*) and vertex z is contained in every directed cycles of the underlying graph, the conditions of Theorem 3.7 are satisfied. Then, if we define $V_{z\zeta} = V_z + \frac{1}{2\gamma_{\zeta_y}}(\zeta_y - \hat{\zeta}_y)^2 + \frac{1}{2\gamma_{\zeta_{\phi_0}}}(\zeta_{\phi_0} - \hat{\zeta}_{\phi_0})^2 + \frac{1}{2\gamma_{\zeta_\Phi}}(\zeta_\Phi - \hat{\zeta}_\Phi)^2$ and re-define $\bar{V} = [V_{z\zeta}, V_{\bar{x}}, V_{\bar{\varepsilon}}, V_\xi, V_\Xi, V_{\bar{v}}]^\top$, an overall storage function can be constructed via linear scaling, namely, $V \triangleq \varsigma^\top \bar{V}$. By Proposition 3.3, for all $\hat{\varpi} > 0$, there exists $\varsigma > 0$, depending on $\hat{\varpi}$, such that

$$\dot{V} \leq -\hat{\varpi}^\top \hat{\psi} - \varsigma^\top \tilde{\psi} \leq 0, \tag{5.157}$$

where $\tilde{\psi} \triangleq [\sum_i^\rho c_i z_i^2,\ \sum_{\rho+2}^n \bar{x}_i^2,\ \bar{\varepsilon}_1^2 + \sum_3^n \bar{\varepsilon}_i^2,\ \xi_1^2 + \sum_3^n \xi_i^2,\ |(\Xi)_1|^2 + \sum_3^n |(\Xi)_i|^2]^\top$. The rest of the proof is identical to the original proof of Proposition 5.3 and the claims in the proposition hold true. □

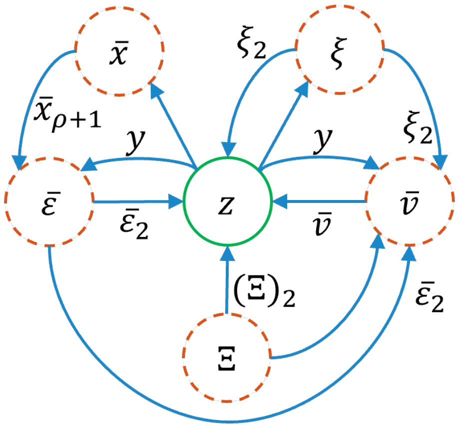

Figure 5.4: The underlying direct graph described by E (the first 6×6 submatrix of $\bar{E}$ in (5.155), with the augmented vertices omitted).

It is worth comparing the time-varying case with the classical time-invariant case, say, the scheme in [75, Chapter 8], by using the node-schematic interpretations in Fig. 5.5. Though the structure of the schematics is more complex than the ones presented in previous sections and chapters, the core idea is essentially the same: in the time-invariant case (Fig. 5.5(b)), the interconnections of the subsystems are acyclic, for which boundedness and convergence properties can be established separately; whereas in the time-varying case (Fig. 5.5(a)), cyclic interconnections are created by the time-varying perturbations, hence a *small-gain-like* analysis is needed to establish boundedness and convergence properties for the overall system. Furthermore, the proposed scheme uses adaptive damping terms parametrized by $\hat{\zeta}_y$, $\hat{\zeta}_{\phi_0}$, and $\hat{\zeta}_\Phi$, which exploits Proposition 3.3 as a synthesis tool to dominate the cyclic interconnection. This does not require the bounds of unknown parameters generated by the change of coordinates, which is favourable in practice as the computation of such bounds is typically difficult.

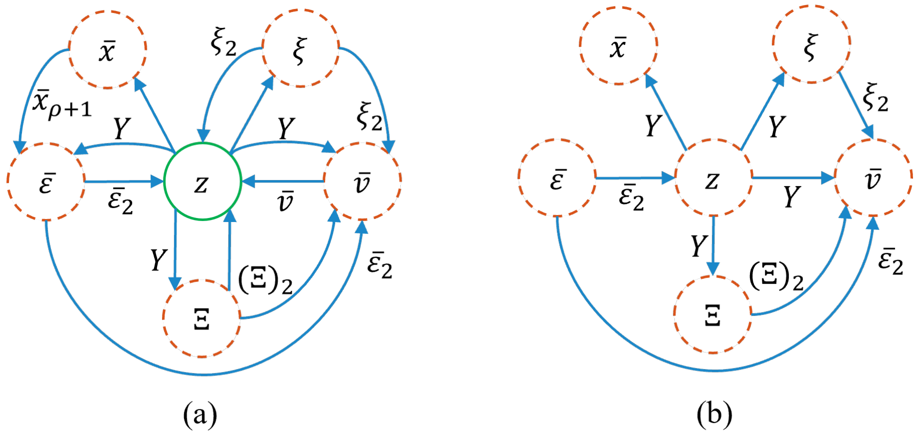

Figure 5.5: Schematic interpretation of the interconnected z, $\bar{x}$, $\bar{\varepsilon}$, ξ, Ξ, and $\bar{v}$ subsystems. (a) shows the interconnection of the case in which the system parameters are time-varying and (b) shows the interconnection of the classical time-invariant case. The signal Y is a collection of the y-related signals y, $\phi_0(y)$, and $\Phi(y)$.

5.2.5 Simulations

To compare the proposed controller with the classical adaptive controller, consider the nonlinear system described by the equations

$$\begin{aligned} \dot{x}_1 &= a_1(t)x_1^2 + x_2, \\ \dot{x}_2 &= a_2(t)x_1^2 + x_3 + b_1(t)u, \\ \dot{x}_3 &= a_3(t)x_1^2 + b_0(t)u, \\ y &= x_1, \end{aligned} \tag{5.158}$$

where the time-varying parameters are defined by

$$b_1(t) = 1 + 0.2\sin(5t), \quad b_0(t) = 6 + \sin(10t), \tag{5.159}$$

$$a(t) = [1, -2, -2]^\top - 10\,\mathrm{sgn}\left(\frac{\partial \alpha_1}{\partial y} z_2\right) \frac{(\bar{\omega})_{3:5}}{|(\bar{\omega})_{3:5}|}, \tag{5.160}$$

with $(\bar{\omega})_{3:5} \triangleq [(\bar{\omega})_3, (\bar{\omega})_4, (\bar{\omega})_5]^\top$. Each of these parameters comprises a constant component and a time-varying component. For parameter a the time-varying component is also state-dependent, and it is designed to destabilize the system. It is not difficult to verify that Assumption 5.3 is satisfied since $\frac{b_0(t)}{b_1(t)} \geq \frac{5}{1.2} > 0$, for all $t \geq 0$.

Consider now three scenarios similar to the ones in Section 4.4 and Section 5.1.5.

The "Baseline" scenario is the case in which the system parameters are constant (evaluated by their constant components in (5.159)) and the controller used is the classical adaptive backstepping controller. The "Controller 1" scenario considers the time-varying system parameters and uses a modified backstepping controller with projection for $\hat{\theta}$ (confined in a ball centred at the origin and with radius 20) and $\hat{\varrho}$ (confined in a ball centred at the origin and with radius 6). The "Controller 2" scenario considers the time-varying parameters (the same as the ones in the "Controller 1" scenario) and adopts the proposed controller. To compare the three scenarios fairly, set the common controller parameters as $c_1 = c_2 = 1$, $\Gamma_\theta = I$, $\gamma_\varrho = 1$ and the initial conditions $\hat{\theta}(0) = 0, \hat{\varrho}(0) = 1$. Each scenario uses an identical set of *K-filters* given by (5.83)–(5.85). The filter gains are obtained by solving the algebraic Riccati equation (5.104) with $Q_{\bar{\varepsilon}} = 10$ and $\gamma_{\bar{\varepsilon}} = 100$, and the filter states are initialized to 0. The initial condition for the system state is set to $x(0) = [1, 0, 0]^\top$. For the parameters solely used in the "Controller 2" scenario, set $\gamma_{(\cdot)} = 1$, $\epsilon_{(\cdot)} = 1$, $\delta_{\Delta_{b_m}} = 0.2$, $\epsilon_{\Delta_{\bar{\theta}}}\delta_{\Delta_{\bar{\theta}}} = 1$ (note that the knowledge of $\delta_{\Delta_{\bar{\theta}}}$ is not required as mentioned in Remark 5.11), with the dynamic damping parameters initialized as $\hat{\zeta}_y(0) = 1$, $\hat{\zeta}_\Phi(0) = 1$ (non-zero initial conditions provide additional damping from the beginning to counteract the time-varying effects).

To keep the comparison fair when using state-dependent parameters (similarly to Section 4.4 and Section 5.1.5), the scenarios "Controller 1" and "Controller 2" are simulated in parallel and simulated twice, generating two sets of simulation data. For the first set, the state-dependent time-varying parameters of both scenarios are generated by the "Controller 1" scenario; and for the second simulation set the parameters are generated by the "Controller 2" scenario. Also note that the "Baseline" scenario does not contain state-dependent parameters and is simulated only once.

The responses of the closed-loop signals in each scenario and in each simulation set are plotted in Fig. 5.8 and Fig. 5.9, respectively, and the parameters used in each simulation set are shown in Fig. 5.6 and Fig. 5.7, respectively. In Simulation set 1, the state-dependent parameters are designed to destabilize the closed-loop system of the "Controller 1" scenario. One can observe from Fig. 5.6 and Fig. 5.8 that the parameter variations between 3.5 second and 4 second excite oscillations in the "Controller 1" scenario, even though the output seems to have been "regulated" to 0 before 3.5 second. Whereas in Simulation set 2, in which the parameters are designed to destabilize the system controlled by the proposed Controller 2, the proposed controller restores the performance of the "Baseline"scenario, with no additional oscillation caused by parameter variations being observed. These results show that the proposed controller (Controller 2) outperforms the classical controller (Controller 1) in the presence of time-varying parameters and effectively prevents the oscillations caused by parameter variations.

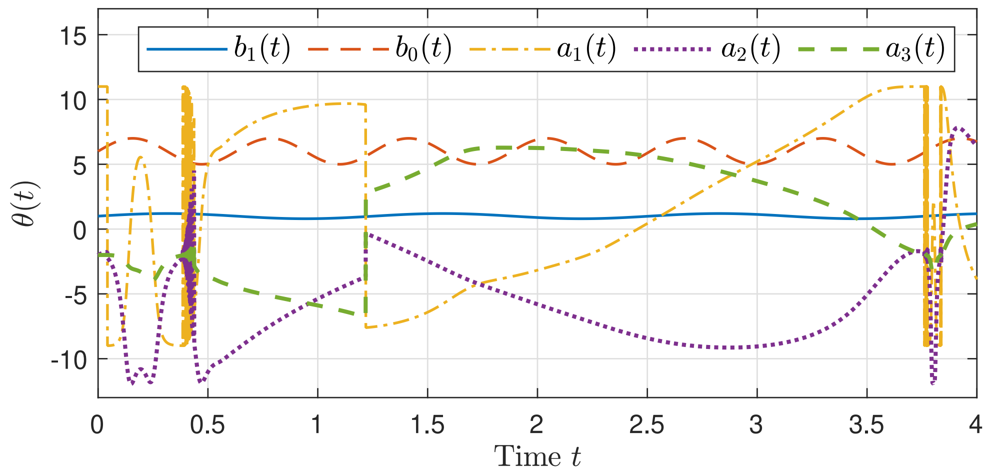

Figure 5.6: Simulation set 1: time-varying parameters generated by the closed-loop system controlled by Controller 1.

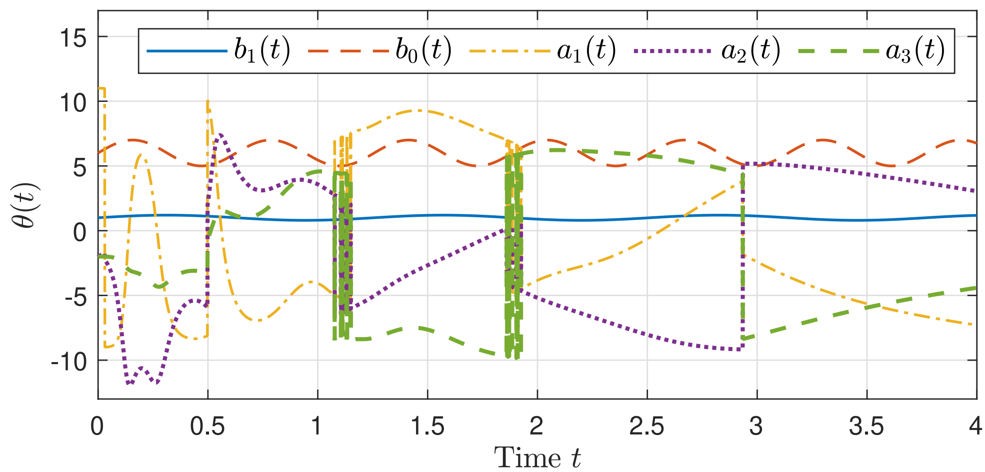

Figure 5.7: Simulation set 2: time-varying parameters generated by the closed-loop system controlled by Controller 2.

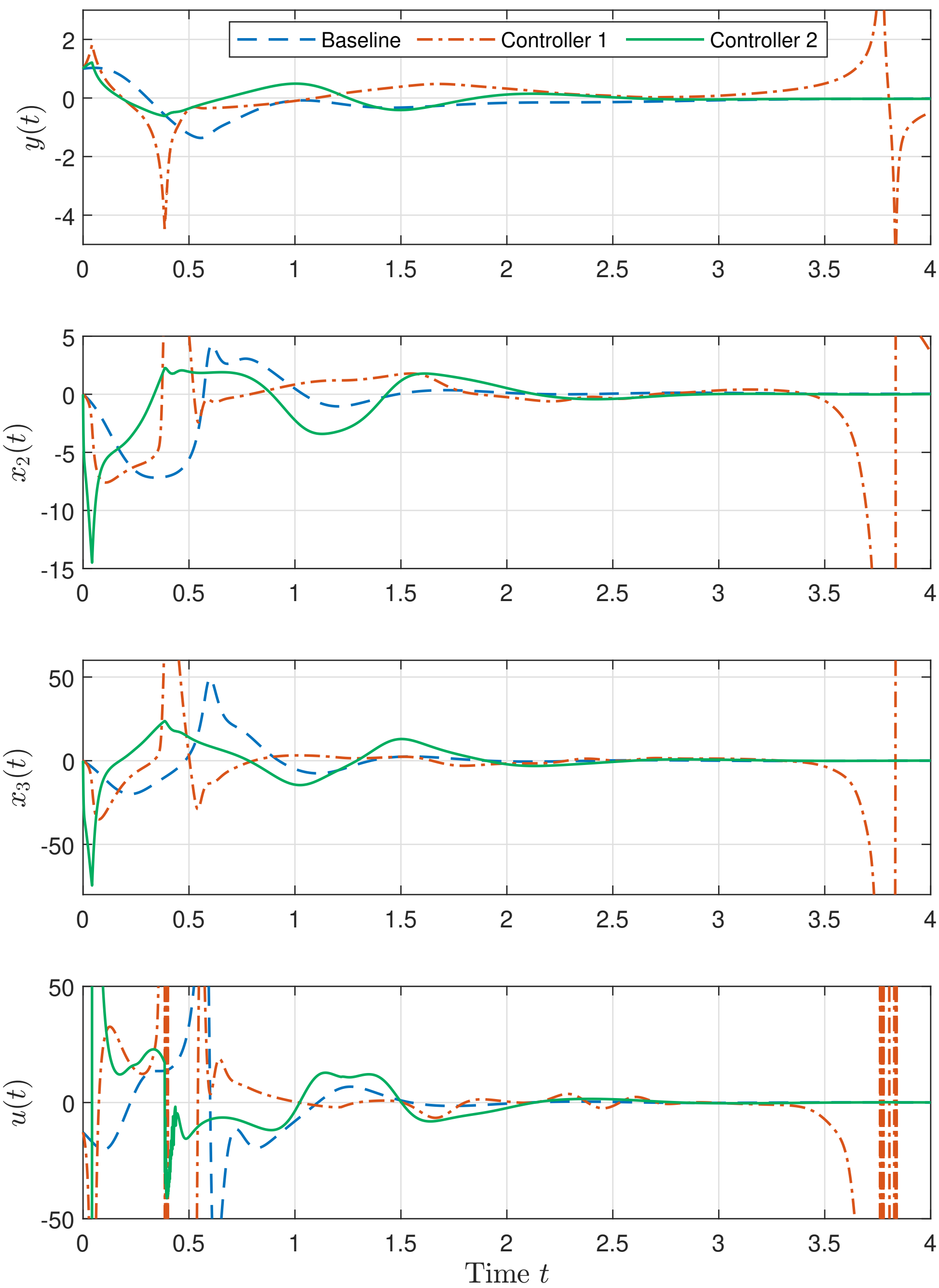

Figure 5.8: Simulation set 1: time histories of the system state and control effort driven by different controllers and the parameters shown in Fig. 5.6.

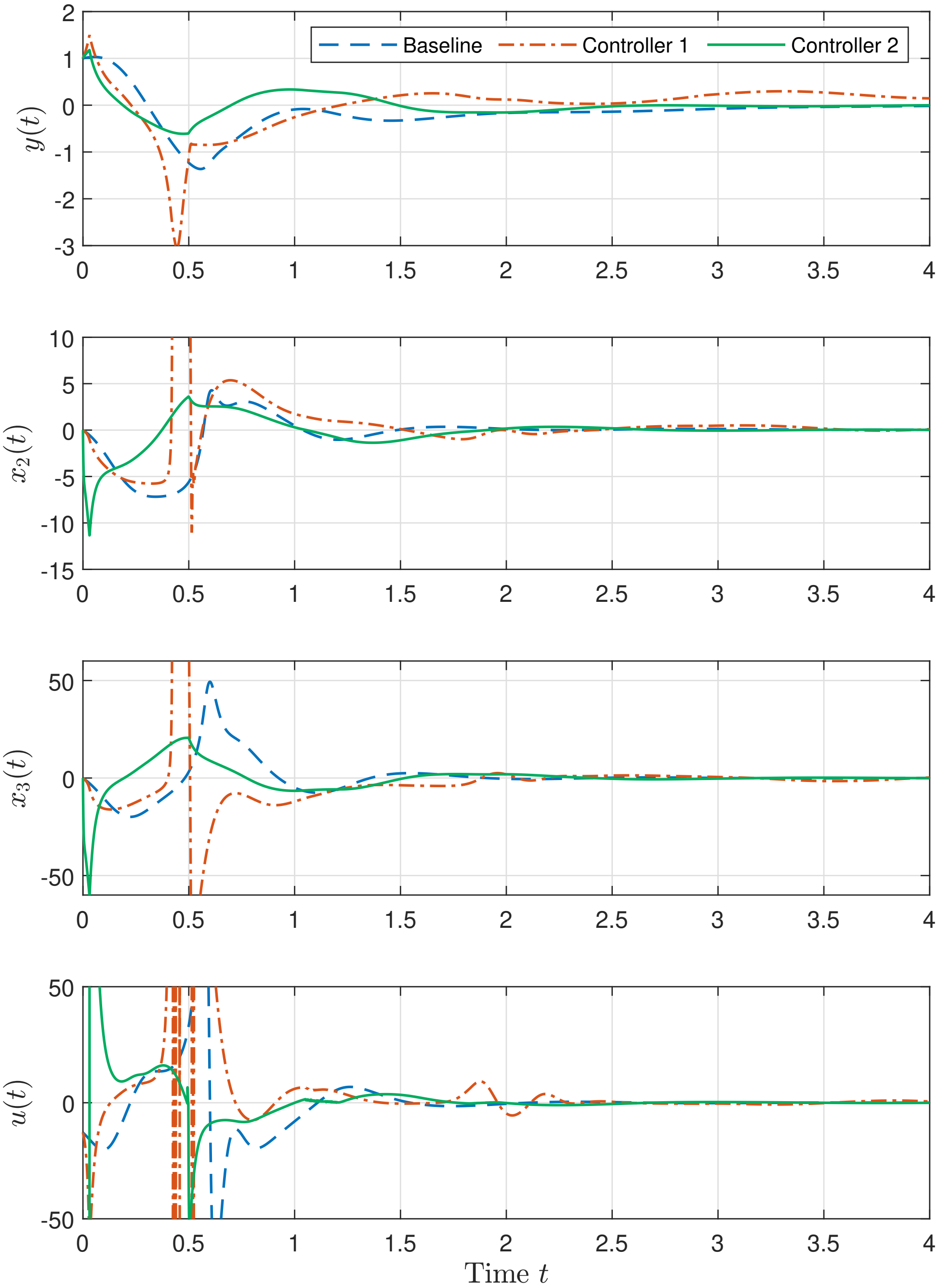

Figure 5.9: Simulation set 2: time histories of the system state and control effort driven by different controllers and the parameters shown in Fig. 5.7.

Chapter 6

Applications

In the preceding chapters the closed-loop performance of the proposed adaptive control schemes have been tested with numerical examples and have been compared with those resulting from the use of classical control schemes designed for time-invariant systems. This chapter aims to shed some light on potential applications of the *congelation of variables* method and the notion of *active node* proposed in the thesis. In the first application, control problem for an actuator servo is solved by treating state-dependent nonlinearities as time-varying parameters and by applying the modified I&I scheme. The second application is concerned with the solution of a disease control problem for interconnected settlements exploiting the notion of *active nodes*, which allows introducing quarantine measures to fewer settlements.

6.1 Series Elastic Actuators

In this section the adaptive I&I scheme introduced in Section 4.2 is exploited to design a controller for the so-called series elastic actuators (SEAs) [104]. SEAs are widely used in robotics: they turn a force control problem into a position control problem using the elastic characteristic of the link, exploiting the well-known Hooke's law.

Control problems arise in SEA due to the extra dynamics caused by the elastic linkage compared to traditional servo problems. A variety of control methods have been applied to SEAs, including PID control [104], PD control with a disturbance observer [71], adaptive control [12], and sliding mode control [8]. In most works, the elastic linkage is modelled as a linear spring with known stiffness and the force exerted on the load is determined by the relative position between the load and the actuator. However, in general, the elastic linkage has nonlinear elastic characteristic. Such nonlinearity is either designed on purpose [138], [108], or unavoidable due to the properties of elastic material [107].

A nonlinear spring can be described by the equation

$$F_s = K_s(d)d, \tag{6.1}$$

where F_s is the elastic force of the spring, K_s is the apparent stiffness function, and d is the deflection of the spring. In the case of a linear spring, K_s is a constant. Since $d(t)$ is time-varying, the value of $K_s\big(d(t)\big)$ is also time-varying and with a slight abuse of notation one can also write $K_s(t)$. This allows viewing K_s as a time-varying parameter and applying an adaptive control scheme for time-varying systems to the SEA position/force control problem, which circumvents the need for incorporating the detailed description of the nonlinearity of the spring into the control design step.

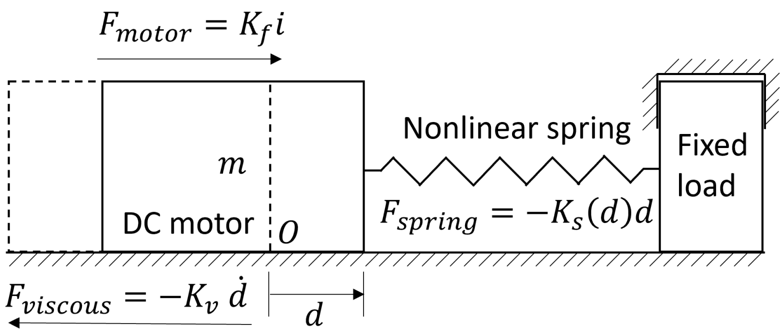

Figure 6.1: Schematic of the SEA with a fixed load.

We now consider the SEA connected with a fixed load as shown in Fig. 6.1. This is the scenario in which the end-effector is in contact with the object to be manipulated and gradually exerts a force on the object (for instance, the egg-grasping task). The goal of the control is to let the DC motor drive the moving end of the spring to the desired deflection d_* such that the force exerted on the load is the desired value $K_s(d_*)$. The transient stage should behave in an over-damped manner so that the force on the end-effector does not cause damage. The physical model of the SEA with fixed load driven by a translational DC motor (a compound of DC motor, gearbox, and linkages that turn the rotary motion into translational motion) is given by the differential equations

$$
\begin{aligned}
m\ddot{d} = & -K_s(d)d - K_v\dot{d} + K_f i, \\
L\dot{i} = & -Ri - K_b\dot{d} + V_{in},
\end{aligned} \tag{6.2}
$$

where m is the apparent mass of the moving parts (the total inertia of the rotor of the motor, the gearbox and other linkages); K_v is the viscous friction constant; K_f is the current-to-force constant; K_b is the back-electromotive-force constant; L is the inductance of the armature; R is the resistance of the armature; i is the current across the armature; and V_{in} is the voltage on the armature. Assume in addition that the constants of the DC motor are known and the only unknown "parameter" is $K_s\big(d(t)\big)$.

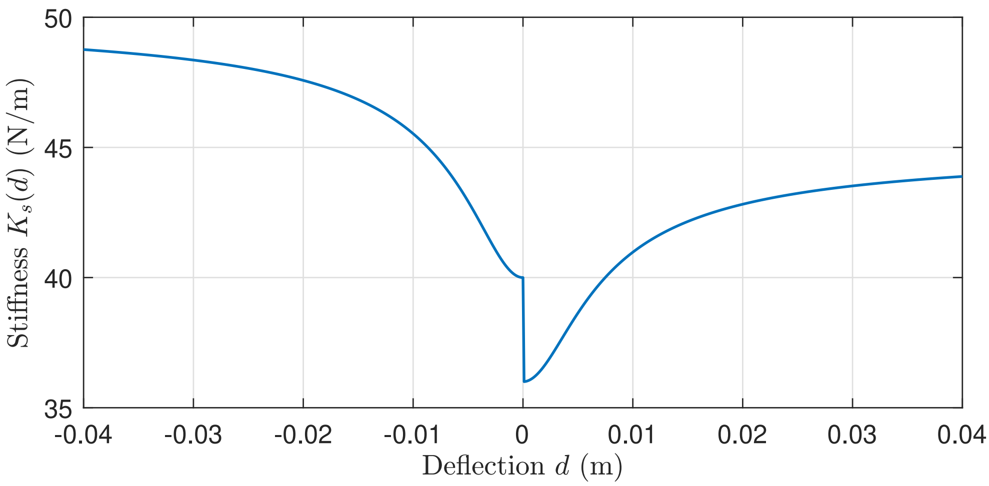

Figure 6.2: Plot of the stiffness function K_s of the nonlinear spring.

Consider an asymmetric nonlinear spring[1], the stiffness of which (plotted in Fig. 6.2) is given by

$$
K_s(d) = \begin{cases} 2K_{s1}\left(1 - \dfrac{l_{01}}{\sqrt{d^2 + l_1^2}}\right), & d \geq 0 \\ 2K_{s2}\left(1 - \dfrac{l_{02}}{\sqrt{d^2 + l_2^2}}\right), & d < 0, \end{cases} \tag{6.3}
$$

[1]A realization of the nonlinearity by means of linear springs is discussed in [138, Fig. 8 (b)] and a realization of the asymmetry is presented in [108, Fig. 4].

where K_{s1} and K_{s2} are the stiffness constants of the linear springs used to realize the nonlinear spring device, l_{01}, l_{02}, l_1, l_2 are parameters related to geometric configurations such that $l_{01} \leq l_1$, $l_{02} \leq l_2$. The system (6.2) has an equilibrium at $d = d_*$, $\dot{d} = 0$, $i = i_*$, with input $V_{in} = V_{in*}$. In a regulator problem we want to shift the origin of the state variables to the desired set point. To this end, define the shifted elastic characteristic $K_{s*}(d)$ such that $K_{s*}(d - d_*) = K_s(d)$. This allows writing (6.2) into the 3-dimensional state space model

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \phi(x_1)\theta(x_1) - ax_2 + x_3, \\
\dot{x}_3 &= u,
\end{aligned} \tag{6.4}$$

where $x_1 \triangleq d - d_*$, $x_2 \triangleq \dot{d}$, $x_3(t) \triangleq \frac{K_f}{m}(i - i_*)$, $u(t) \triangleq \frac{K_f}{L}\big((V_{in} - Ri - K_b\dot{d}) - (V_{in*} - Ri_*)\big)$, $a \triangleq \frac{K_v}{m}$, $\phi(x_1) \triangleq -x_1$, and

$$\theta(x_1) \triangleq \frac{1}{m}\left(K_{s*}(x_1) + \frac{K_{s*}(x_1) - K_{s*}(0)}{x_1}d_*\right). \tag{6.5}$$

Due to boundedness of $K_{s*}(x_1)$ and the Lipschitz continuity of $K_{s*}(x_1)$ at $x_1 = 0$, $\theta\big(x_1(t)\big)$ can be treated as a bounded time-varying parameter and denoted by $\theta(t)$, with a slight abuse of notation.

Since system (6.4) contains *unmatched* uncertainty, the controller design requires the use of the backstepping procedures as follows. Note that in general the over-parametrized scheme used in Section 4.2.3 is required for *lower-triangular systems*, because the construction of the β function (4.41) in Section 4.2.1 does not possess the *lower-triangular* dependency. However, for system (6.4), the selection (4.41) is valid since

$$\beta(x, \hat{x}) = \gamma_\theta \Phi(\hat{x})x = \gamma_\theta \begin{bmatrix} 0 & -\hat{x}_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -\gamma_\theta \hat{x}_1 x_2 \tag{6.6}$$

is brought to the control synthesis stage in the second step of backstepping, indicating that (6.6) has a *lower-triangular* dependency (β only depends on variables with a subscript less or equal to 2). Therefore, overparametrization is not needed for the considered SEA system and we can directly use the dynamic scaling estimator discussed in Section 4.2.1. Proceed now with the backstepping procedure.

Step 1. Define $z_1 = x_1$ and $z_2 = x_2 - \alpha_1$, the first two backstepping error variables, and let the first virtual control law be

$$\alpha_1 = -\sigma_1, \tag{6.7}$$

where σ_1 is a damping function to be defined. This yields

$$\dot{z}_1 = z_2 + \alpha_1 = -\sigma_1 + z_2. \tag{6.8}$$

Step 2. Define $z_3 = x_3 - \alpha_2$ and

$$\alpha_2 = -\sigma_2 - z_1 - \phi_2(\hat{\theta} + \beta) + ax_2 + \frac{\partial \alpha_1}{\partial x_1} x_2, \tag{6.9}$$

with σ_2 a damping function to be defined. Then the dynamics of z_2 becomes

$$\dot{z}_2 = z_3 + \alpha_2 = -\sigma_2 - z_1 + z_3 + \phi_2(rz_\theta - \Delta_\theta), \tag{6.10}$$

with $z_\theta \triangleq \dfrac{\hat{\theta} - \ell_\theta + \beta}{r}$.

Step 3. Let the actual control law

$$\begin{aligned} u = & - \sigma_3 - z_2 + \frac{\partial \alpha_2}{\partial x_1} x_2 + \frac{\partial \alpha_2}{\partial x_2} \big((\hat{\theta} + \beta) - ax_2 + x_3 \big) \\ & + \frac{\partial \alpha_2}{\partial r} \dot{r} + \frac{\partial \alpha_2}{\partial \hat{x}} \dot{\hat{x}} + \frac{\partial \alpha_2}{\partial \hat{\theta}} \dot{\hat{\theta}}, \end{aligned} \tag{6.11}$$

with σ_3 to be defined. This yields the dynamics of the third error variable

$$\dot{z}_3 = -\sigma_3 - z_2 + \frac{\partial \alpha_2}{\partial x_2} \phi_2(rz_\theta - \Delta_\theta). \tag{6.12}$$

Proposition 6.1. *Consider system* (6.4) *with the dynamic scaling estimator given by* (4.41)*,* (4.42)*,* (4.43)*,* (4.48)*,* (4.47)*, and the controller* (6.11)*. Select the damping terms* *as*

$$\sigma_1 = \left(c_1 + \frac{13}{2}\delta_{\Delta_\theta}^2\right) z_1, \tag{6.13}$$

$$\sigma_2 = (c_2 + r^2 + 1) z_2, \tag{6.14}$$

$$\sigma_3 = \left(c_3 + \left(\frac{\partial \alpha_2}{\partial x_2}\right)^2 (r^2+1)\right) z_3, \tag{6.15}$$

with $c_{(\cdot)} > 0$. *Then, all trajectories of the closed-loop system are bounded and* $\lim_{t\to+\infty} x(t) = 0$. *In particular,* $d \to d_*$ *as* $t \to +\infty$.

Proof. Consider the Lyapunov function candidate $V_z = \frac{1}{2}z_1^2 + \frac{1}{2}z_2^2 + \frac{1}{2}z_3^2$. The time derivative of V_z along the trajectories of the system is such that

$$\begin{aligned}
\dot{V}_z =& -\sigma_1 z_1 + z_1 z_2 - \sigma_2 z_2 - z_1 z_2 + z_2 z_3 + z_2 \phi_2 (r z_\theta - \Delta_\theta) \\
& - \sigma_3 z_3 - z_2 z_3 + z_3 \frac{\partial \alpha_2}{\partial x_2} \phi_2 (r z_\theta - \Delta_\theta) \\
\le & - \sigma_1 z_1 - \sigma_2 z_2 - \sigma_3 z_3 + z_2^2 (r^2 + 1) \\
& + z_3^2 \left(\frac{\partial \alpha_2}{\partial x_2}\right)^2 (r^2+1) + \frac{1}{2}(\phi_2 z_\theta)^2 + \frac{1}{2}\delta_{\Delta_\theta}^2 \phi_2^2.
\end{aligned} \tag{6.16}$$

Using the damping terms (6.13)-(6.15) yields

$$\dot{V}_z \le -c_1 z_1^2 - c_2 z_2^2 - c_3 z_3^2 - 6\phi_2^2 + \frac{1}{2}(\phi_2 z_\theta)^2. \tag{6.17}$$

It can be concluded from Proposition 4.3 that $\lim_{t\to+\infty} z(t) = 0$ and all trajectories of the closed-loop system are bounded. Using a standard argument for stability in the backstepping scheme, $\lim_{t\to+\infty} z(t) = 0$ implies $\lim_{t\to+\infty} x_1(t) = 0$ and $\lim_{t\to+\infty} \alpha_1(t) = 0$, which gives $\lim_{t\to+\infty} x_2(t) = 0$, since $\lim_{t\to+\infty} z_2(t) = 0$. In the same way we can prove that $\lim_{t\to+\infty} x_3(t) = 0$ and this completes the proof. □

Remark 6.1. *In practice it is not necessary to use exactly* δ_{Δ_θ} *when designing the controller. It leads to a conservative design that can guarantee boundedness and convergence properties even in the worst case, however, the resulting controller can have high gains and large-amplitude control signals, which may cause robustness issues. Typically a discounted version of the "radius"* $\delta_d < \delta_{\Delta_\theta}$ *can be implemented to make a "softer" version of the* *proposed controller.*

Consider now the SEA with the parameters: $K_{s1} = 45$ N/m, $K_{s2} = 50$ N/m, $l_{01} = l_{02} = 1 \times 10^{-3}$ m, $l_1 = l_2 = 5 \times 10^{-3}$ m, $m = 1$ kg, $R = 3\ \Omega$, $L = 1.5 \times 10^{-4}$ H, $K_f = 1.5$ N/A, $K_b = 2.5$ V·s/m, $K_v = 0.01$ N·s/m, and the estimator-controller setting: $\gamma_\theta = 1 \times 10^3$, $\lambda = \epsilon = 5$, $\kappa = 100$, $c_1 = 2$, $c_2 = 2$, $c_3 = 2$, $\delta_d^2 = 2$. Let $d_* = 2 \times 10^{-2}$ m, and let the initial condition be $x(0) = [-0.04,\ 0,\ -0.85]^\top$ (the third element enforces zero initial armature current).

The closed-loop system responses are shown in Fig. 6.4. The closed-loop is tuned to behave in an "overdamped" manner so that the force is smoothly exerted on the load and thus prevents damage to the manipulated load. Fig. 6.3 shows the time history of the state-dependent parameter θ during the transient stage. Note that θ is bounded as is stated in the discussion on (6.5).

The SEA example shows that the *congelation of variables* method can be exploited to extend the use of adaptive control, which has mainly been used for coping with constant parametric uncertainties in the classical scenarios, to deal with state-dependent nonlinearities (treated as time-varying parameters). It is therefore natural to apply this idea to other systems with state-dependent nonlinearities of similar characteristics.

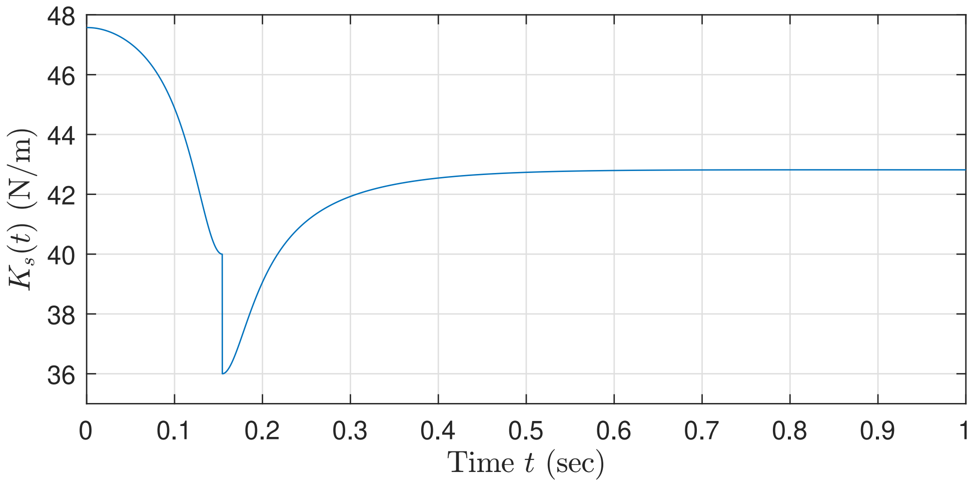

Figure 6.3: Time history of θ.

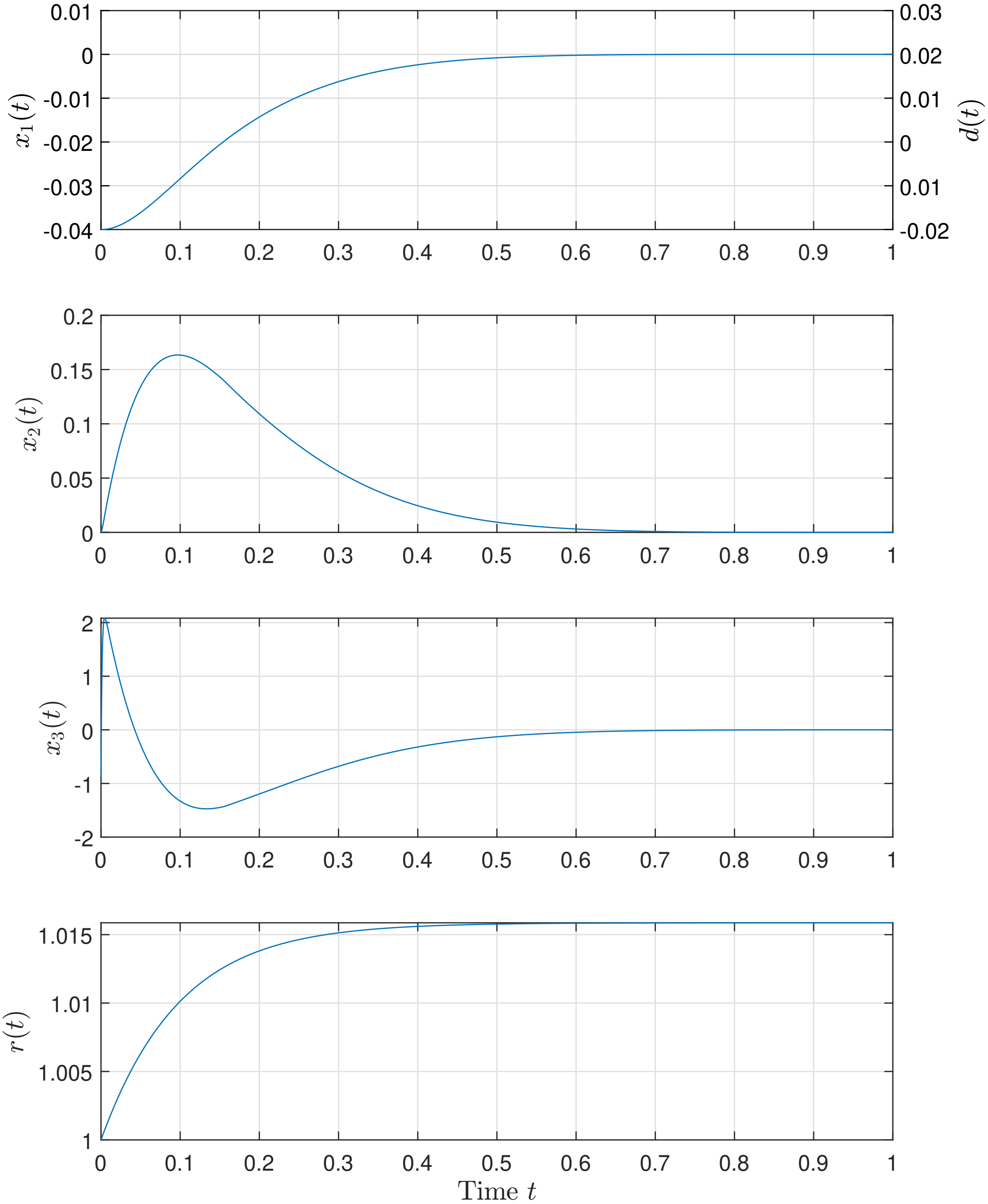

Figure 6.4: Time histories of the states of the closed-loop system.

6.2 Disease Control of Interconnected Settlements

In this section we discuss how to use the notion of *active nodes* to solve a disease control problem. In particular, consider the deterministic *Susceptible-Infectious-Quarantined-Susceptible* (SIQS) model [35, 46], a modification from the *Susceptible-Infectious-Susceptible* (SIS) model (see *e.g.* [91, Section 10.3]). The isolated version the SIQS model is described by the equations

$$\begin{aligned}\dot{S} &= -\frac{\beta SI}{N} + \gamma(I+Q),\\ \dot{I} &= \frac{\beta SI}{N} - (\gamma+\delta)I,\\ \dot{Q} &= \delta I - \gamma Q,\end{aligned} \tag{6.18}$$

where $S(t) \geq 0$ is the susceptible population, $I(t) \geq 0$ is the infectious population, $\beta > 0$ is the infection parameter, $\gamma > 0$ is the recovery parameter, and $\delta > 0$ is the quarantine parameter. Since $\dot{S} + \dot{I} + \dot{Q} = 0$, the total population N remains constant, that is,

$$S(t) + I(t) = N, \tag{6.19}$$

for all $t \geq 0$. This model describes infectious diseases that are in general non-lethal and for which immunity is not acquired after recovery, like influenza. Therefore the infectious and quarantined population returns to the susceptible population once recovered, instead of being removed. It is possible to extend this model to a scenario in which n interconnected settlements and the population migrating among them are considered. For simplification, in this example we assume that (6.19) holds for every settlement, that is,

$$S_i(t) + I_i(t) + Q_i(t) = N_i, \tag{6.20}$$

for $i = 1 \dots n$, as the population that the facilities of each settlement can support is approximately constant in the considered time period. Consider now the interconnected version of the infectious population dynamics

$$\dot{I}_i = \frac{\beta_i S_i I_i}{N_i} - \left(\gamma_i + \delta_i + \sum_{j \in \mathcal{S}_i} \frac{\mu_{ji}}{N_i}\right) I_i + \sum_{j \in \mathcal{P}_i} \frac{\mu_{ij}}{N_j} I_j, \tag{6.21}$$

where $\mu_{ij} \geq 0$ denotes the rate of migration from settlement j to settlement i. Clearly, equation (6.20) holds if and only if

$$\sum_{j \in \mathcal{S}_i} \mu_{ji} = \sum_{j \in \mathcal{P}_i} \mu_{ij}, \tag{6.22}$$

which indicates the balance of migration. Substituting (6.20) into (6.21) yields

$$\dot{I}_i = -\left(\gamma_i + \delta_i - \beta_i + \frac{1}{N_i} \sum_{j \in \mathcal{S}_i} \mu_{ji}\right) I_i + \sum_{j \in \mathcal{P}_i} \frac{\mu_{ij}}{N_j} I_j - P_i, \tag{6.23}$$

where $P_i \triangleq \frac{\beta_i}{N_i}(I_i^2 + Q_i I_i)$. It is not difficult to verify that $\mathbb{R}_+$ is forward invariant for S_i, I_i, Q_i and therefore $P_i(t) \geq 0$ for all $t \geq 0$. Thus one can select $V_i(I_i) = I_i$ as the storage function for each node and the associated dissipation inequality is

$$\dot{V}_i \leq \; - a_i I_i + \sum_{j \in \mathcal{P}_i} b_{ij} I_j, \tag{6.24}$$

where $a_i \triangleq \gamma_i + \delta_i - \beta_i + \frac{1}{N_i} \sum_{j \in \mathcal{S}_i} \mu_{ji}$, $b_{ij} \triangleq \frac{\mu_{ij}}{N_j}$. Note that so far we have not specified whether the parameters γ_i, β_i, μ_{ij} are constant or time-varying. In the classical SIQS model these parameters are assumed to be constant. It turns out that if these parameters are time-varying, a_i and b_{ij} are also time-varying scalar parameters, which requires rewriting (6.24) into the dissipation inequality

$$\dot{V}_i \leq \; - \ell_{a_i} I_i + \sum_{j \in \mathcal{P}_i} \ell_{b_{ij}} I_j, \tag{6.25}$$

with constant parameters $\ell_{a_i} \triangleq \inf_{t \geq 0} a_i(t)$ and $\ell_{b_{ij}} \triangleq \sup_{t \geq 0} b_{ij}(t)$. This, however, does not alter the structure of the underlying directed graph.

It can be observed that in the case in which neither migration nor quarantine is considered (that is, all μ-terms and δ-terms are zero), the infectious population of each settlement converges to zero if the *basic reproduction number* denoted by $R_{0i} \triangleq \frac{\beta_i}{\gamma_i} < 1$, or equivalently, $\beta_i > \gamma_i$. If quarantine measures are adopted, there exists a $\delta_i > 0$ such that $\delta_i > \beta_i - \gamma_i$, which guarantees that the local infectious population asymptotically converges to zero. For convenience of the subsequent discussion, we assume that the quarantine force at each settlement is at least locally "sufficient" in the sense that the condition

$$\delta_i > \beta_i - \gamma_i - \frac{1}{N_i} \sum_{j \in \mathcal{S}_i} \mu_{ji} \tag{6.26}$$

holds, for $i = 1, \ldots, n$. The counterpart of this condition in the interconnected case is, however, more complicated due to the non-zero b_{ij}-terms. The infection in the network system can be amplified through the interconnections even if the damping effect at each node is sufficient for the isolated scenario. This is part of the reasons why cutting off transport is considered as a public health control measure. In what follows we show, using the notion of *active nodes*, that it is possible to bring the infectious population to zero at each settlement without the need for cutting off inter-settlement transport, provided the quarantine forces at the settlements that serve as the "transport hubs" are adjustable.

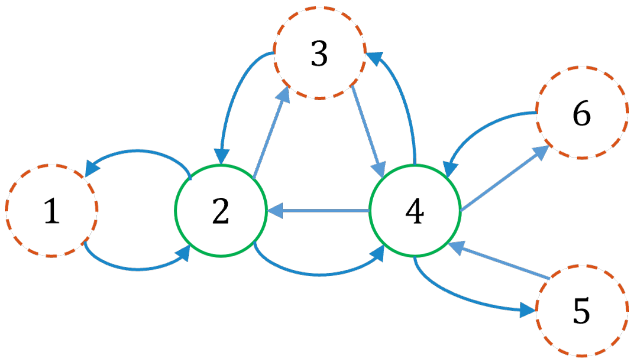

Figure 6.5: The directed graph describing the migration among the six settlements.

Consider now the six interconnected settlements described by the graph in Fig. 6.5, with each settlement containing population of $N \triangleq [N_1, \ldots, N_6]^\top = [12, 50, 30, 70, 24, 18]^\top$ (unit: thousand people) migration parameters defined by $\mu_{ij} \triangleq (\mathrm{M})_{ij}$, where $\mathrm{M}(t) \triangleq$

$\mathrm{M}_{const} + \mathrm{M}_{TV}(t)$,

$$\mathrm{M}_{const} \triangleq \begin{bmatrix} 0 & 6 & 0 & 0 & 0 & 0 \\ 6 & 0 & 9 & 7 & 0 & 0 \\ 0 & 7 & 0 & 13 & 0 & 0 \\ 0 & 9 & 11 & 0 & 12 & 8 \\ 0 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 \end{bmatrix}, \tag{6.27}$$

and

$$\mathrm{M}_{TV}(t) \triangleq \begin{bmatrix} 0 & \sin\frac{2\pi}{7}t & 0 & 0 & 0 & 0 \\ \sin\frac{2\pi}{7}t & 0 & -1.4\sin\frac{2\pi}{7}t & 0.8\cos\frac{2\pi}{7}t & 0 & 0 \\ 0 & -1.4\sin\frac{2\pi}{7}t & 0 & -1.8\cos\frac{2\pi}{7}t & 0 & 0 \\ 0 & 0.8\cos\frac{2\pi}{7}t & -1.8\cos\frac{2\pi}{7}t & 0 & 2\sin\frac{2\pi}{7}t & 1.1\cos\frac{2\pi}{7}t \\ 0 & 0 & 0 & 2\sin\frac{2\pi}{7}t & 0 & 0 \\ 0 & 0 & 0 & 1.1\cos\frac{2\pi}{7}t & 0 & 0 \end{bmatrix}, \tag{6.28}$$

with unit: thousand people/day. The sinusoidal signals in $\mathrm{M}_{TV}(t)$ is included to model the weekly fluctuations of migration rates. One can verify that the given $\mathrm{M}(t)$ matrix satisfies conditions (6.22) for $i = 1, \ldots, 6$ and for all $t \geq 0$. Only one infectious disease prevailing among the settlements is considered. The associated infection and recovery parameters are $\beta_i = 2.0$ and $\gamma_i = 1.4$ (unit: day^{-1}), for $i = 1, \ldots, 6$, respectively[2]. The initial conditions are $I(0) = [0.03,\ 0.045,\ 0.015,\ 0.012,\ 0.015,\ 0.022]^\top$ and $Q(0) = \mathit{0}$.

In this example the quarantine forces δ_i is the control variable to be designed. Consider now two quarantine policies. In Scenario 1, a regional quarantine policy is applied to all six settlements, with identical quarantine force at each settlement, that is, $\delta_i = 0.5$, for $i = 1, \ldots, 6$. This verifies the conditions (6.26). This, however, does not guarantee that the infectious population converges to zero, as shown in Fig. 6.6. The infectious population and quarantine population reach a nonzero equilibrium in the 3-month period considered, which is also called the *endemic* steady state, a common phenomenon of SIS or SIQS

[2]The values of β and γ are adopted from the simulation model of seasonal influenza considered in [99].

models. Note that the node dissipation inequalities (6.24) are in a similar form as (3.2). Furthermore, conditions (6.26) guarantee that $a_i > 0$, for $n = 1, \ldots, 6$. Therefore, instead of strengthening the quarantine measures in the whole region, we can exploit the notion of *active nodes* and the adaptive technique discussed in Section 3.4.3. This case, denoted by Scenario 2, only requires applying more strict quarantine measures at settlements 2 and 4 (denoted by green solid circles in Fig. 6.5). More specifically, we replace the constant quarantine forces δ_2 and δ_4 with adaptive estimates $\hat{\delta}_2$ and $\hat{\delta}_4$, updated by the equations

$$\dot{\hat{\delta}}_2 = 1.2 I_2, \qquad \hat{\delta}_2(0) = 0.5, \tag{6.29}$$

$$\dot{\hat{\delta}}_4 = 3 I_4, \qquad \hat{\delta}_4(0) = 0.5. \tag{6.30}$$

Note that due to the monotonicity and the given initial conditions of $\hat{\delta}_2$ and $\hat{\delta}_4$, the projection operation is not needed in this scenario. Invoking Proposition 3.3, we conclude that the supply rate in the network dissipation inequality is negative definite in I and applying standard invariance analysis yields convergence of I to zero, which is confirmed by the simulation results shown in Figs. 6.7 and 6.8.

The example discussed is simple but reveals that it is possible to exploit the notion of *active nodes* and the transport topology for a more efficient public health policy, which does not require an overall strict quarantine or transport lockdown over all nodes and may improve the trade-off between safety and disruption of normal life. It should, however, be emphasized that the SIQS model considered is only an elementary epidemiological model and does not consider the side effects brought by excessive enforcement of quarantine. A settlement that serves as a “transport hub” typically also provides essential services to the neighbouring settlements. An excessively strengthened quarantine policy at the “transport hubs” can effectively prevent the spread of diseases, whereas at the same time hinder the delivery of necessary services and supplies to the whole region. In the light of this, public health policy-making requires comprehensive consideration of all aspects of the society to achieve comprehensive and sustainable welfare.

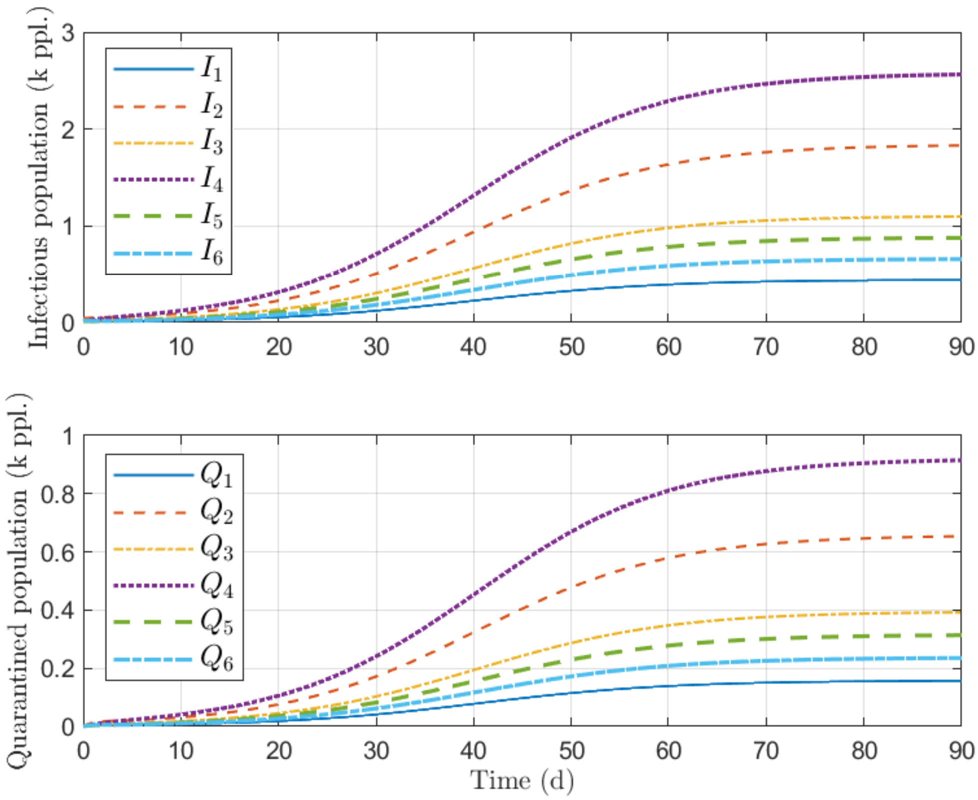

Figure 6.6: Time histories of the infectious and the quarantined populations in Scenario 1.

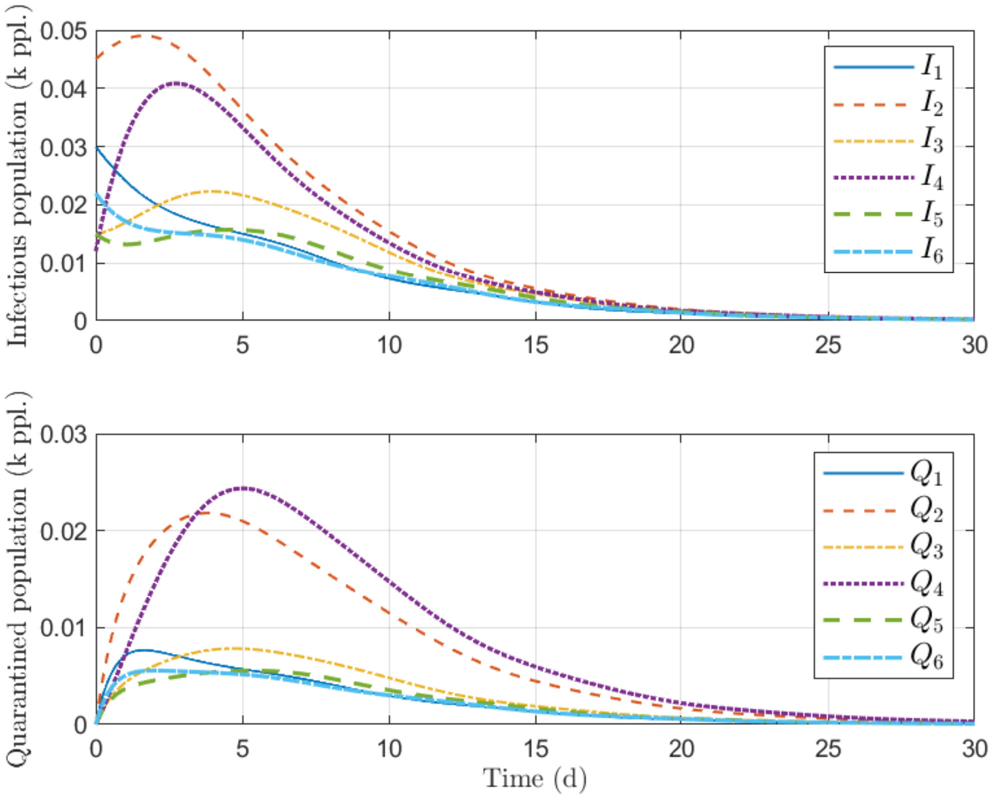

Figure 6.7: Time histories of the infectious and the quarantined populations in Scenario 2.

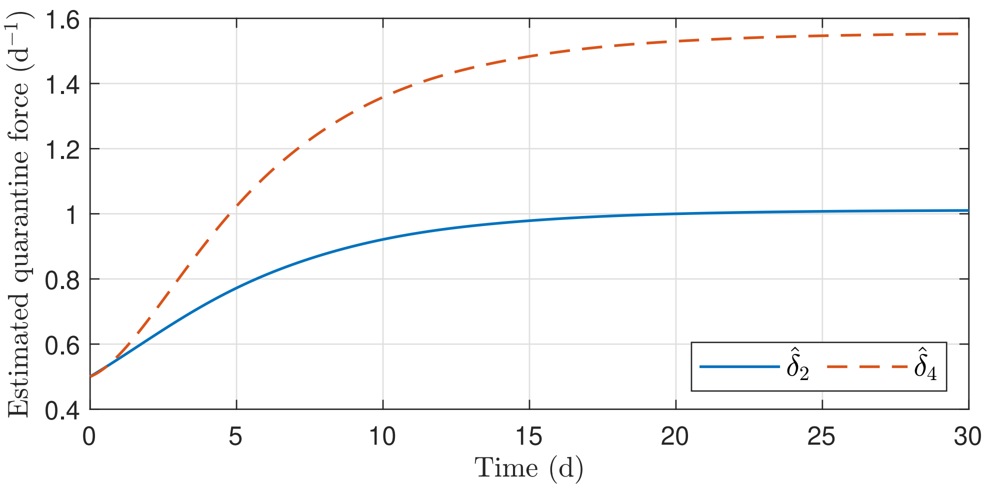

Figure 6.8: Time histories of the estimated quarantine forces in Scenario 2.

Chapter 7

Conclusion

7.1 Summary of the Results

The thesis has proposed and discussed adaptive control schemes for time-varying systems, along with the two key theoretical tools:

- the *congelation of variables* method for re-formulating the original problem; and
- the notion of *active nodes* for dominance design over cyclic interconnection.

The thesis has reviewed the challenges brought by parameter variations in the classical control, that is, the "corruption" of passivity or $\mathcal{L}_2$-stability due to parameter variations and has introduced the idea of the *congelation of variables*. Using such a method, the time-varying parameters can be divided into a constant unknown component and a time-varying perturbation. The constant component yields an underlying adaptive control problem that can be solved by classical parameter update schemes, and the time-varying perturbation creates cyclic interconnections among the subsystems to be dominated via a small-gain-like framework.

To conduct a dominance design over the network of subsystems with cyclic structure, the notion of *active nodes* has been introduced, which suggests that as long as the node systems the damping coefficients of which are adjustable, *i.e.* the *active nodes*, form an FVS of the underlying graph of the network system, the overall dissipation inequality can be made negative by means of linear or nonlinear scaling of the node dissipation inequalities. Within this framework analysis and control synthesis conditions have been proposed for systems with quadratic supply rates, sum-type nonlinear supply rates, and linearly parametrized nonlinear supply rates. The methods for effectively placing the *active nodes*, and the schemes for obtaining damping coefficients with and without explicit computation, respectively, have been provided.

Using the two tools proposed, the thesis has then investigated the state-feedback adaptive control problem for a class of nonlinear systems with *lower-triangular* structure. The combination of the *congelation of variables* method with the passivity-based scheme, the I&I scheme, and the identification-based scheme has been elaborated. Simulation results show that the proposed scheme has superior performance in the presence of time-varying parameters, compared to classical schemes.

The thesis has also studied the output-feedback adaptive control problems. The I&I scheme for SISO linear time-varying systems, and the passivity-based scheme for a class of SISO nonlinear systems have been proposed. The control synthesis and stability analysis have been proven both using the constructive method and the small-gain-like framework based on *active nodes*. The notion of *active nodes* significantly reduces the complexity of the analysis and increases the flexibility of the results. The simulation results show that the performance of the proposed output-feedback schemes is also superior to that of their classical counterparts.

Potential applications, including the servo control problem of SEA, in which the nonlinearity are viewed as a time-varying parameter to simplify the design, and a disease control problem among interconnected settlements, in which the settlements serving as "traffic hubs" are set as *active nodes*, have been discussed. These provide some examples on how the proposed theoretical results can be exploited in practice.

7.2 Future Research Directions

Though the thesis has provided tools and a framework to systematically solve the adaptive control problem for a certain class of systems, the current results are far from comprehensive. This leaves some interesting directions for future study, listed in what follows.

- The current results rely on the condition that the effects of time-varying parameters vanish at the reference set-point or trajectory (in the thesis such a condition is enforced by Assumption 1.4). The relaxation of such a condition yields a disturbance rejection problem. This requires either the knowledge of the *exosystem* that generates the time-varying parameters, or a non-smooth control law that resembles (higher-order) sliding-mode control. It is interesting to see if the same result can be achieved without using the knowledge of parameters model or non-smooth control laws.

- The dynamic nonlinear damping terms for dominance design, in the current results, can only be updated by passivity-based schemes, whereas in the I&I and identification-based schemes, one has to use static nonlinear damping terms. This is due to the fact that these damping terms are introduced by the designer, not come inherently from the system. Thus, while these terms can be "compensated", they cannot be "identified" or "observed". It would be interesting to see if one can use I&I or identification-based to achieve such dynamic dominance design.

- The *active nodes* defined in the thesis possess damping coefficients that can be adjusted up to $+\infty$, that is, the damping coefficients, viewed as a whole, have a conic admissible region. It is interesting to see how the results based on FVS may change if one considers a polyhedral admissible region for the damping coefficients, resulting, for example, in putting upper limits for these coefficients.

- The analysis for the output-feedback schemes explicitly considers the inverse dynamics of the system. Though the systems are assumed to be minimum-phase, the spirit of the analysis may provide a better framework to investigate adaptive control problem for nonminimum-phase systems, which remains an open topic.

- The thesis focuses on output-feedback problems for SISO systems. It would be interesting to understand what modifications are needed to extend the current results to the multiple-input multiple-output case.

Appendix A

Useful Lemmas

Lemma A.1 (Young's inequality)**.** *Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}^n$, and $\epsilon > 0$. The inequality*

$$a^\top b \leq \frac{1}{2\epsilon}|a|^2 + \frac{1}{2}\epsilon|b|^2 \tag{A.1}$$

holds.

Lemma A.2 (Young's inequality for functions [139])**.** *Consider a continuous strictly increasing function f and positive real numbers, $a \in [0, c]$, $b \in [0, f(c)]$, $c > 0$, Then the inequality*

$$ab \leq \int_0^a f(s)ds + \int_0^b f^{-1}(s)ds \tag{A.2}$$

holds, and in particular, the equality holds if and only if $b = f(a)$.

Lemma A.3 (Hadamard's lemma)**.** ([45], see also, *e.g.*, [95, Lemma 2.8] for modern interpretations) *A smooth mapping $\phi : \mathbb{R}^n \to \mathbb{R}^q$ can be written as*

$$\phi(x) = \phi(\bar{x}) + \bar{\Phi}(x)(x - \bar{x}), \tag{A.3}$$

where $\bar{x} \in \mathbb{R}^n$ and $\bar{\Phi} : \mathbb{R}^n \to \mathbb{R}^{q\times n}$ is a smooth mapping. A possible selection of $\bar{\Phi}$ is

$$\bar{\Phi}(x) = \int_0^1 \nabla^\top \phi\big(\bar{x} + s(x - \bar{x})\big)\mathrm{ds}, \tag{A.4}$$

where $\nabla^\top \phi$ denotes the Jacobian of ϕ.

The following two lemmas are useful to establish convergence.

Lemma A.4 (Barbalat's lemma [9]). *Let $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a uniformly continuous function. If $\left| \lim_{t \to +\infty} \int_0^{+\infty} \phi(s) ds \right| = M < +\infty$, then*

$$\lim_{t \to +\infty} \phi(t) = 0. \tag{A.5}$$

Lemma A.5. *Consider a storage function $V \geq 0$ and the associated dissipation inequality $\dot{V} \leq -W(y)$, where $W(\cdot)$ is a differentiable positive definite function and $y : \mathbb{R} \to \mathbb{R}^{n_y}$ is differentiable. If $\frac{\partial W}{\partial y}$ and $\dot{y}$ are bounded, then*

$$\lim_{t \to +\infty} y(t) = 0. \tag{A.6}$$

Proof. Since W is positive definite and $\dot{V} \leq -W(y)$, one has

$$0 \leq \int_0^{+\infty} W\big(y(t)\big) \mathrm{d}t \leq \Big(V\big(y(0)\big) - V\big(y(+\infty)\big)\Big) \leq V\big(y(0)\big). \tag{A.7}$$

Note that $\dot{W} = \frac{\partial W}{\partial y}\dot{y}$ is bounded due to boundedness of $\frac{\partial W}{\partial y}$ and $\dot{y}$, and $W \circ y$ is uniformly continuous in t. Invoking Barbalat's lemma yields

$$\lim_{t \to +\infty} W\big(y(t)\big) = 0. \tag{A.8}$$

Hence $\lim_{t \to +\infty} y(t) = 0$ due to the positive definiteness of W. $\square$

It should be noted that Lemma A.5 is a commonly used argument in the proofs for adaptive control systems (see, *e.g.*, [125]) and it is formulated in this form for the convenience of the expressions in this thesis. Boundedness analysis and convergence analysis for adaptive control systems, especially for those using auxiliary filters, are typically done separately. For this reason, Lemma A.5 and its counterparts commonly found in the literature, are more popular than LaSalle-Yoshizawa Theorem (see, *e.g.*, [75, Theorem A.8]), which proves boundedness and convergence together, but requires incorporating all state variables into the proof at once. Nevertheless, both the spirits and the proofs of these results are essentially similar as they are all based on Barbalat's lemma.

The next lemma provides a useful property of adjustable class-$\mathcal{K}_\infty$ functions.

Lemma A.6. *Consider class-$\mathcal{K}_\infty$ functions $\underline{\alpha}$, β. There exists $\underline{\gamma} \in \mathcal{K}_\infty$ such that for all class-$\mathcal{K}_\infty$ function $\gamma \geq \underline{\gamma}$, there exists a class-$\mathcal{K}_\infty$ function $\alpha \geq \underline{\alpha}$, depending on γ, which satisfies*

$$\alpha - \beta \geq \gamma. \tag{A.9}$$

Proof. Letting $\underline{\gamma} = \underline{\alpha}$ and $\alpha = \beta + \gamma + \delta$, with $\delta \in \mathcal{K}_\infty$. Since β, γ,and δ are class-$\mathcal{K}_\infty$ functions, α is also a class-$\mathcal{K}_\infty$ function. In addition,

$$\alpha \geq \gamma + \beta \geq \underline{\alpha} + \beta \geq \underline{\alpha}, \tag{A.10}$$

which satisfies (A.9). The proof is now complete. □

In short, Lemma A.6 reveals that an adjustable $\alpha \in \mathcal{K}_\infty$, lower-bounded by $\underline{\alpha} \in \mathcal{K}_\infty$, guarantees the existence of an adjustable $\gamma \in \mathcal{K}_\infty$, lower-bounded by $\underline{\gamma} \in \mathcal{K}_\infty$ and satisfying (A.9).

References

[1] A. V. Aho and N. J. Sloane, "Some doubly exponential sequences," *Fibonacci Quart.*, vol. 11, no. 4, pp. 429–437, 1973.

[2] D. Angeli and A. Astolfi, "A tight small-gain theorem for not necessarily ISS systems," *Syst. & Control Lett.*, vol. 56, no. 1, pp. 87–91, 2007.

[3] A. M. Annaswamy and K. S. Narendra, "Adaptive control of simple time-varying systems," in *Proc. 28th IEEE Conf. Decis. and Control.* IEEE, 1989, pp. 1014–1018.

[4] S. Aranovskiy, A. Bobtsov, R. Ortega, and A. Pyrkin, "Parameters estimation via dynamic regressor extension and mixing," in *Proc. Amer. Control Conf.* IEEE, 2016, pp. 6971–6976.

[5] A. Astolfi, D. Karagiannis, and R. Ortega, *Nonlinear and Adaptive Control with Applications.* Springer Science & Business Media, 2007.

[6] A. Astolfi and R. Ortega, "Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 590–606, 2003.

[7] K. J. Astrom and B. Wittenmark, *Adaptive Control.* Courier Corporation, 2013.

[8] J. Bae, K. Kong, and M. Tomizuka, "Gait phase-based smoothed sliding mode control for a rotary series elastic actuator installed on the knee joint," in *Proc. Amer. Control Conf.* IEEE, 2010, pp. 6030–6035.

[9] I. Barbalat, "Systems of differential equations of nonlinear oscillations," *Rev. Math. Pures Appl. (in French)*, vol. 4, no. 2, pp. 267–270, 1959.

[10] P. Bernard, M. Bin, and L. Marconi, "Adaptive output regulation via nonlinear luenberger observer-based internal models and continuous-time identifiers," *Automatica*, vol. 122, p. 109261, 2020.

[11] M. Bin and L. Marconi, "'Class-type' identification-based internal models in multivariable nonlinear output regulation," *IEEE Trans. on Autom. Control*, vol. 65, no. 10, pp. 4369–4376, 2019.

[12] J. A. Blaya and H. M. Herr, "Adaptive control of a variable-impedance ankle-foot orthosis to assist drop-foot gait," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 12, no. 1, pp. 24–31, 2004.

[13] U. Borison, "Self-tuning regulators for a class of multivariable systems," *Automatica*, vol. 15, no. 2, pp. 209–215, 1979.

[14] J. Chen, Y. Liu, S. Lu, B. O'Sullivan, and I. Razgon, "A fixed-parameter algorithm for the directed feedback vertex set problem," in *Proc. 40th annu. ACM symp. Theory Comput.*, 2008, pp. 177–186.

[15] K. Chen, "Adaptive control with varying parameters," Master's thesis, Imperial College London, 2017.

[16] K. Chen and A. Astolfi, "Adaptive control of linear systems with time-varying parameters," in *Proc. Amer. Control Conf.* IEEE, 2018, pp. 80–85.

[17] ——, "I&I adaptive control for systems with varying parameters," in *Proc. 57th IEEE Conf. Decis. and Control.* IEEE, 2018, pp. 2205–2210.

[18] ——, "Output-feedback adaptive control for systems with time-varying parameters," in *Proc. 11th IFAC Symp. Nonlinear Control Syst.* IFAC, 2019, pp. 1061–1066.

[19] ——, "Output-feedback I&I adaptive control for linear systems with time-varying parameters," in *Proc. 58th IEEE Conf. Decis. and Control.* IEEE, 2019, pp. 1965–1970.

[20] ——, "Adaptive control for nonlinear systems with time-varying parameters and control coefficient," in *Proc. of IFAC World Congr.* IFAC, 2020, pp. 3895–3900.

[21] ——, "On the active nodes of network systems," in *Proc. 59th IEEE Conf. Decis. and Control*. IEEE, 2020, pp. 5561–5566.

[22] ——, "Adaptive control for systems with time-varying parameters," *IEEE Trans. Autom. Control*, vol. 66, no. 5, pp. 1986–2001, 2021.

[23] ——, "Identification-based adaptive control for systems with time-varying parameters," in *Proc. 60th IEEE Conf. Decis. and Control*. IEEE, 2021, pp. 1083–1088.

[24] ——, "Adaptive control for systems with time-varying parameters—a survey," *Trends in Nonlinear and Adaptive Control*, pp. 217–247, 2022.

[25] ——, "Decentralized adaptive control for interconnected cyber-physical systems under coordinated attacks," in *Proc. 61st IEEE Conf. Decis. and Control*. IEEE, 2022, to be published.

[26] ——, "Active nodes of network systems with sum-type dissipation inequalities," *IEEE Trans. Autom. Control*, submitted for publication.

[27] ——, "On the adaptive tracking problem of time-varying systems using the congelation for variables method," in *12th IFAC Symp. Nonlinear Control Syst. (NOLCOS 2022)*. IFAC, submitted for publication.

[28] Y.-Y. Chen, K. Chen, and A. Astolfi, "Adaptive formation tracking control of directed networked vehicles in a time-varying flowfield," *J. Guid. Control and Dyn.*, vol. 44, no. 10, pp. 1883–1891, 2021.

[29] ——, "Adaptive formation tracking control for first-order agents with a time-varying flow parameter," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2481–2488, 2022.

[30] S. Dashkovskiy, H. Ito, and F. Wirth, "On a small gain theorem for ISS networks in dissipative Lyapunov form," *Eur. J. Control*, vol. 17, no. 4, pp. 357–365, 2011.

[31] S. Dashkovskiy and S. Pavlichkov, "Stability conditions for infinite networks of nonlinear systems and their application for stabilization," *Automatica*, vol. 112, p. 108643, 2020.

[32] S. Dashkovskiy, B. S. Ruffer, and F. R. Wirth, "An ISS small gain theorem for general networks," *Math. Control, Signals and Syst.*, vol. 19, no. 2, pp. 93–122, 2007.

[33] S. N. Dashkovskiy, B. S. Ruffer, and F. R. Wirth, "Small gain theorems for large scale systems and construction of ISS Lyapunov functions," *SIAM J. Control and Optim.*, vol. 48, no. 6, pp. 4089–4118, 2010.

[34] H. Elliott, R. Cristi, and M. Das, "Global stability of adaptive pole placement algorithms," *IEEE Trans. Autom. Control*, vol. 30, no. 4, pp. 348–356, 1985.

[35] Z. Feng and H. R. Thieme, "Recurrent outbreaks of childhood diseases revisited: the impact of isolation," *Math. Biosci.*, vol. 128, no. 1-2, pp. 93–130, 1995.

[36] A. Feuer and A. Morse, "Adaptive control of single-input, single-output linear systems," *IEEE Trans. Autom. Control*, vol. 23, no. 4, pp. 557–569, 1978.

[37] R. Freeman and P. V. Kokotovic, *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques.* Springer Science & Business Media, 2008.

[38] R. A. Freeman and P. V. Kokotovic, "Tracking controllers for systems linear in the unmeasured states," *Automatica*, vol. 32, no. 5, pp. 735–746, 1996.

[39] G. Frobenius, "On matrices of non-negative elements," *Proc. Roy. Prussian Acad. Sci. (in German)*, pp. 456–477, 1912.

[40] F. Giri, M. M'saad, J.-M. Dion, and L. Dugard, "Pole placement direct adaptive control for time-varying ill-modeled plants," *IEEE Trans. Autom. Control*, vol. 35, no. 6, pp. 723–726, 1990.

[41] G. C. Goodwin and E. K. Teoh, "Adaptive control of a class of linear time varying systems," *IFAC Proc. Volumes*, vol. 16, no. 9, pp. 1–6, 1983.

[42] G. Goodwin, P. Ramadge, and P. Caines, "Discrete-time multivariable adaptive control," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 449–456, 1980.

[43] G. C. Goodwin and D. Q. Mayne, "A parameter estimation perspective of continuous time model reference adaptive control," *Automatica*, vol. 23, no. 1, pp. 57–70, 1987.

[44] G. C. Goodwin and K. S. Sin, *Adaptive filtering Prediction and Control.* Courier Corporation, 2014.

[45] J. Hadamard, *Lessons on Wave Propagation and the Equations of Hydrodynamics.* A. Hermann (in French), 1903.

[46] H. Hethcote, M. Zhien, and L. Shengbing, "Effects of quarantine in six endemic models for infectious diseases," *Math. Biosci.*, vol. 180, no. 1-2, pp. 141–160, 2002.

[47] N. Hovakimyan and C. Cao, $\mathcal{L}_1$ *Adaptive Control Theory: Guaranteed Robustness with Fast Adaptation.* SIAM, 2010.

[48] J. Huang, W. Wang, C. Wen, and J. Zhou, "Adaptive control of a class of strict-feedback time-varying nonlinear systems with unknown control coefficients," *Automatica*, vol. 93, pp. 98–105, 2018.

[49] P. Ioannou and K. Tsakalis, "A robust direct adaptive controller," *IEEE Trans. Autom. Control*, vol. 31, no. 11, pp. 1033–1043, 1986.

[50] P. A. Ioannou and P. V. Kokotovic, "Instability analysis and improvement of robustness of adaptive control," *Automatica*, vol. 20, no. 5, pp. 583–594, 1984.

[51] P. A. Ioannou and J. Sun, *Robust Adaptive Control.* PTR Prentice-Hall Upper Saddle River, NJ, 1996.

[52] H. Ito, "State-dependent scaling problems and stability of interconnected iISS and ISS systems," *IEEE Trans. Autom. Control*, vol. 51, no. 10, pp. 1626–1643, 2006.

[53] H. Ito, Z.-P. Jiang, S. N. Dashkovskiy, and B. S. Ruffer, "Robust stability of networks of iISS systems: construction of sum-type Lyapunov functions," *IEEE Trans. Autom. Control*, vol. 58, no. 5, pp. 1192–1207, 2012.

[54] H. Ito and C. M. Kellett, "iISS and ISS dissipation inequalities: Preservation and interconnection by scaling," *Math. Control, Signals. and Syst.*, vol. 28, no. 3, pp. 1–36, 2016.

[55] M. Jerrum, A. Sinclair, and E. Vigoda, "A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries," *J. ACM*, vol. 51, no. 4, pp. 671–697, 2004.

[56] Z.-P. Jiang, A. R. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Math. Control, Signals and Syst.*, vol. 7, no. 2, pp. 95–120, 1994.

[57] Z.-P. Jiang and T. Liu, "Small-gain theory for stability and control of dynamical networks: A survey," *Ann. Rev. Control*, vol. 46, pp. 58–79, 2018.

[58] Z.-P. Jiang, I. M. Mareels, and Y. Wang, "A Lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems," *Automatica*, vol. 32, no. 8, pp. 1211–1215, 1996.

[59] Z.-P. Jiang and Y. Wang, "A generalization of the nonlinear small-gain theorem for large-scale complex systems," in *Proc. 7th World Congr. Intell. Control and Automat.* IEEE, 2008, pp. 1188–1193.

[60] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," in *Proc. Amer. Control Conf.* IEEE, 1991, pp. 649–654.

[61] D. Karagiannis and A. Astolfi, "Nonlinear adaptive control of systems in feedback form: an alternative to adaptive backstepping," *IFAC Proc. Vol.*, vol. 37, no. 11, pp. 67–72, 2004.

[62] ——, "Nonlinear adaptive control of systems in feedback form: An alternative to adaptive backstepping," *Syst. & Control Lett.*, vol. 57, no. 9, pp. 733–739, 2008.

[63] D. Karagiannis, A. Astolfi, and R. Ortega, "Two results for adaptive output feedback stabilization of nonlinear systems," *Automatica*, vol. 39, no. 5, pp. 857–866, 2003.

[64] D. Karagiannis, Z.-P. Jiang, R. Ortega, and A. Astolfi, "Output-feedback stabilization of a class of uncertain non-minimum-phase nonlinear systems," *Automatica*, vol. 41, no. 9, pp. 1609–1615, 2005.

[65] D. Karagiannis, M. Sassano, and A. Astolfi, "Dynamic scaling and observer design with application to adaptive control," *Automatica*, vol. 45, no. 12, pp. 2883–2889, 2009.

[66] R. M. Karp, "Reducibility among combinatorial problems," in *Proc. Symp. Complexity Comput. Computations.* Springer, 1972, pp. 85–103.

[67] C. Kawan, A. Mironchenko, A. Swikir, N. Noroozi, and M. Zamani, “A Lyapunov-based small-gain theorem for infinite networks,” *IEEE Trans. Autom. Control*, 2020.

[68] C. M. Kellett and P. M. Dower, “Input-to-state stability, integral input-to-state stability, and $\mathcal{L}_2$-gain properties: Qualitative equivalences and interconnected systems,” *IEEE Trans. Autom. Control*, vol. 61, no. 1, pp. 3–17, 2015.

[69] C. M. Kellett and F. R. Wirth, “Nonlinear scaling of iISS-Lyapunov functions,” *IEEE Trans. Autom. Control*, vol. 61, no. 4, pp. 1087–1092, 2015.

[70] H. K. Khalil, *Nonlinear Systems.* Prentice Hall, 2002.

[71] K. Kong, J. Bae, and M. Tomizuka, “Control of rotary series elastic actuator for ideal force-mode actuation in human–robot interaction applications,” *IEEE/ASME Trans. Mechatron.*, vol. 14, no. 1, pp. 105–118, 2009.

[72] G. Kreisselmeier, “Adaptive observers with exponential rate of convergence,” *IEEE Trans. Autom. Control*, vol. 22, no. 1, pp. 2–8, 1977.

[73] ——, “Adaptive control of a class of slowly time-varying plants,” *Syst. & Control Lett.*, vol. 8, no. 2, pp. 97–103, 1986.

[74] M. Krstic and P. V. Kokotovic, “Adaptive nonlinear design with controller-identifier separation and swapping,” *IEEE Trans. Autom. Control*, vol. 40, no. 3, pp. 426–440, 1995.

[75] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and Adaptive Control Design*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc, 1995.

[76] E. Lavretsky and K. A. Wise, *Robust and Adaptive Control With Aerospace Applications.* Springer, 2013.

[77] A. Lempel and I. Cederbaum, “Minimum feedback arc and vertex sets of a directed graph,” *IEEE Trans. Circuit Theory*, vol. 13, no. 4, pp. 399–403, 1966.

[78] Z. Li and M. Krstic, “Optimal-design of adaptive tracking controllers for nonlinear-systems,” *Automatica*, vol. 33, no. 8, pp. 1459–1473, 1997.

[79] W. Lin and C. Qian, "Adaptive control of nonlinearly parameterized systems: a nonsmooth feedback framework," *IEEE Trans. Autom. Control*, vol. 47, no. 5, pp. 757–774, 2002.

[80] ——, "Adaptive control of nonlinearly parameterized systems: the smooth feedback case," *IEEE Trans. Autom. Control*, vol. 47, no. 8, pp. 1249–1266, 2002.

[81] T. Liu, D. J. Hill, and Z.-P. Jiang, "Lyapunov formulation of ISS cyclic-small-gain in continuous-time dynamical networks," *Automatica*, vol. 47, no. 9, pp. 2088–2093, 2011.

[82] T. Liu, P. Zhang, and Z.-P. Jiang, "Basic stability and small-gain tools for system synthesis," in *Robust Event-Triggered Control of Nonlinear Systems*. Springer, 2020, pp. 19–40.

[83] D. Lokshtanov, M. Ramanujan, and S. Saurabh, "A linear time parameterized algorithm for directed feedback vertex set," *arXiv:1609.04347*, 2016.

[84] R. Marino and P. Tomei, "An adaptive output feedback control for a class of nonlinear systems with time-varying parameters," *IEEE Trans. Autom. Control*, vol. 44, no. 11, pp. 2190–2194, 1999.

[85] ——, "Robust adaptive regulation of linear time-varying systems," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1301–1311, 2000.

[86] ——, "Adaptive control of linear time-varying systems," *Automatica*, vol. 39, no. 4, pp. 651–659, 2003.

[87] Merriam-Webster Staff, *Merriam-Webster's Collegiate Dictionary*. Merriam-Webster, 2004.

[88] R. H. Middleton and G. C. Goodwin, "Adaptive control of time-varying linear systems," *IEEE Trans. Autom. Control*, vol. 33, no. 2, pp. 150–155, 1988.

[89] R. Monopoli and C. Hsing, "Parameter adaptive control of multivariable systems," *Int. J. Control*, vol. 22, no. 3, pp. 313–327, 1975.

[90] A. Morse, "Global stability of parameter-adaptive control systems," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 433–439, 1980.

[91] J. D. Murray, *Mathematical Biology: I. An introduction*, 3rd ed. Springer, 2002.

[92] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems.* Prentice Hall, 1989.

[93] K. Narendra, Y.-H. Lin, and L. Valavani, "Stable adaptive controller design, Part II: Proof of stability," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 440–448, 1980.

[94] K. Nassiri-Toussi and W. Ren, "Indirect adaptive pole-placement control of MIMO stochastic systems: self-tuning results," *IEEE Trans. Autom. Control*, vol. 42, no. 1, pp. 38–52, 1997.

[95] J. Nestruev, *Smooth Manifolds and Observables*, 2nd ed. Springer Science & Business Media, 2020.

[96] A. Nijenhuis and S. Herbert, *Combinatorial Algorithms: for Computers and Calculators.* Academic Press, 1978.

[97] R. Ortega, L. Praly, S. Aranovskiy, B. Yi, and W. Zhang, "On dynamic regressor extension and mixing parameter estimators: Two luenberger observers interpretations," *Automatica*, vol. 95, pp. 548–551, 2018.

[98] P. Osbourne, "New developments in the design of model reference adaptive control systems," in *29th Annual Meeting.* Inst. of Aeronautical Sci., 1961.

[99] D. Osthus, K. S. Hickmann, P. C. Caragea, D. Higdon, and S. Y. Del Valle, "Forecasting seasonal influenza with a state-space SIR model," *Ann. Appl. Statist.*, vol. 11, no. 1, pp. 202–224, 2017.

[100] O. S. Patil, R. Sun, S. Bhasin, and W. E. Dixon, "Adaptive control of time-varying parameter systems with asymptotic tracking," *IEEE Trans. Autom. Control*, 2022.

[101] R. J. Plemmons, "M-matrix characterizations.I–Nonsingular M-matrices," *Linear Algebr. and its Appl.*, vol. 18, no. 2, pp. 175–188, 1977.

[102] J.-B. Pomet and L. Praly, "Adaptive nonlinear regulation: Estimation from the Lyapunov equation," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 729–740, 1992.

[103] L. Praly, “Asymptotic stabilization via output feedback for lower triangular systems with output dependent incremental rate,” *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 1103–1108, 2003.

[104] G. A. Pratt and M. M. Williamson, “Series elastic actuators,” in *Proc. 1995 IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, vol. 1. IEEE, 1995, pp. 399–406.

[105] I. Razgon, “Computing minimum directed feedback vertex set in $O^*(1.9977^n)$,” in *Proc. 10th Italian Conf. Theor. Comput. Sci.* World Scientific, 2007, pp. 70–81.

[106] H. Rios, D. Efimov, J. A. Moreno, W. Perruquetti, and J. G. Rueda-Escobedo, “Time-varying parameter identification algorithms: Finite and fixed-time convergence,” *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3671–3678, 2017.

[107] D. Rollinson, S. Ford, B. Brown, and H. Choset, “Design and modeling of a series elastic element for snake robots,” in *Proc. ASME 2013 Dynamic Syst. Control Conf.* ASME, 2013, pp. V001T08A002–V001T08A002.

[108] E. J. Rouse, L. M. Mooney, and H. M. Herr, “Clutchable series-elastic actuator: Implications for prosthetic knee design,” *The Int. J. of Robotics Research*, vol. 33, no. 13, pp. 1611–1625, 2014.

[109] J. G. Rueda-Escobedo and J. A. Moreno, “Discontinuous gradient algorithm for finite-time estimation of time-varying parameters,” *Int. J. of Control*, vol. 89, no. 9, pp. 1838–1848, 2016.

[110] H. J. Ryser, *Combinatorial Mathematics.* Amer. Math. Soc., 1963, vol. 14.

[111] G. Scarciotti, L. Praly, and A. Astolfi, “A small-gain-like theorem for large-scale systems,” in *52nd IEEE Conf. Decis. and Control.* IEEE, 2013, pp. 6593–6596.

[112] ——, “Invariance-like theorems and “lim inf” convergence properties,” *IEEE Trans. Autom. Control*, vol. 61, no. 3, pp. 648–661, 2015.

[113] O. Schuck, “Honeywell’s history and philosophy in the adaptive control field,” in *Proc. Self Adaptive Flight Control Symp.* Wright Patterson AFB, Ohio, 1959.

[114] R. A. Servedio and A. Wan, “Computing sparse permanents faster,” *Inf. Process. Lett.*, vol. 96, no. 3, pp. 89–92, 2005.

[115] D. D. Siljak, *Large-Scale Dynamic Systems: Stability and Structure.* North Holland, 1978.

[116] H. A. Simon, "Dynamic programming under uncertainty with a quadratic criterion function," *Econometrica*, pp. 74–81, 1956.

[117] G. Smith and R. Walford, "The identification of a minimal feedback vertex set of a directed graph," *IEEE Trans. Circuits and Syst.*, vol. 22, no. 1, pp. 9–15, 1975.

[118] E. D. Sontag, "Comments on integral variants of ISS," *Syst. & Control Lett.*, vol. 34, no. 1-2, pp. 93–100, 1998.

[119] E. D. Sontag and Y. Wang, "On characterizations of the input-to-state stability property," *Syst. & Control Lett.*, vol. 24, no. 5, pp. 351–359, 1995.

[120] ——, "Notions of input to output stability," *Syst. & Control Lett.*, vol. 38, no. 4-5, pp. 235–248, 1999.

[121] J. Sun, "A modified model reference adaptive control scheme for improved transient performance," *IEEE Trans. Autom. Control*, vol. 38, no. 8, pp. 1255–1259, 1993.

[122] D. Swaroop and J. K. Hedrick, "String stability of interconnected systems," *IEEE Tran. Autom. Control*, vol. 41, no. 3, pp. 349–357, 1996.

[123] J. J. Sylvester, "On a point in the theory of vulgar fractions," *Amer. J. Math.*, vol. 3, no. 4, pp. 332–335, 1880.

[124] Y. Tan, I. Kanellakopoulos, and Z.-P. Jiang, "Nonlinear observer/controller design for a class of nonlinear systems," in *Proc. 37th Conf. Decis. and Control*, vol. 3. IEEE, 1998, pp. 2503–2508.

[125] G. Tao, "A simple alternative to the Barbalat lemma," *IEEE Trans. Autom. Control*, vol. 42, no. 5, p. 698, 1997.

[126] ——, *Adaptive Control Design and Analysis.* John Wiley & Sons, 2003.

[127] ——, "Multivariable adaptive control: A survey," *Automatica*, vol. 50, no. 11, pp. 2737–2764, 2014.

[128] ——, "Lecture notes on adaptive control," pp. 46–51, 2019. [Online]. Available: http://www.ece.virginia.edu/~gt9s/wiley-book/2003book-notes.pdf

[129] K. Tsakalis and P. A. Ioannou, "Adaptive control of linear time-varying plants," *Automatica*, vol. 23, no. 4, pp. 459–468, 1987.

[130] M. Vidyasagar, *Input-Output Analysis of Large-Scale Interconnected Systems: Decomposition, Well-posedness and Stability.* Springer Berlin Heidelberg, 1981.

[131] C. Wang and L. Guo, "Adaptive cooperative tracking control for a class of nonlinear time-varying multi-agent systems," *J. of the Franklin Inst.*, vol. 354, no. 15, pp. 6766–6782, 2017.

[132] J. Wang, N. Hovakimyan, and C. Cao, "$\mathcal{L}_1$ adaptive augmentation of gain-scheduled controller for racetrack maneuver in aerial refueling," in *Guid. Navigation and Control Conf.* AIAA, 2009, p. 5739.

[133] L. Wang and C. M. Kellett, "Adaptive tracking control via immersion and invariance: An (i) ISS perspective," in *Proc. 58th Conf. Decis. and Control.* IEEE, 2019, pp. 7019–7024.

[134] ——, "Robust I&I adaptive tracking control of systems with nonlinear parameterization: An ISS perspective," *arXiv preprint arXiv:2102.13411*, 2021.

[135] C. Wen, "An indirect robust continuous-time adaptive controller with minimal modifications," *Automatica*, vol. 31, no. 2, pp. 293–296, 1995.

[136] H. Whitaker, J. Yamron, and A. Kezer, "Design of model reference control systems for aircraft," Instrumentation Laboratory, Massachusetts Inst. of Technol., Tech. Rep. R-164, 1958.

[137] B. Xian, D. M. Dawson, M. S. de Queiroz, and J. Chen, "A continuous asymptotic tracking control strategy for uncertain nonlinear systems," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1206–1211, 2004.

[138] J. Yamaguchi and A. Takanishi, "Development of a biped walking robot having antagonistic driven joints using nonlinear spring mechanism," in *Proc. 1997 IEEE Int. Conf. Robotics and Automat.*, vol. 1. IEEE, 1997, pp. 185–192.

[139] W. H. Young, "On classes of summable functions and their Fourier series," *Proc. Roy. Soc. Ser. A*, vol. 87, no. 594, pp. 225–229, 1912.

[140] J. F. Zhang, "Robust stabilization of LTI systems via indirect adaptive controllers," *IEEE Trans. Autom. control*, vol. 40, no. 5, pp. 900–906, 1995.

[141] Y. Zhang, B. Fidan, and P. A. Ioannou, "Backstepping control of linear time-varying systems with known and unknown parameters," *IEEE Trans. Autom. Control*, vol. 48, no. 11, pp. 1908–1925, 2003.

[142] Y. Zhang and P. A. Ioannou, "Adaptive control of linear time varying systems," in *Proc. 35th IEEE Conf. Decis. and Control*, vol. 1. IEEE, 1996, pp. 837–842.

[143] Z. Zhang, X.-J. Xie, and S. S. Ge, "Adaptive tracking for uncertain MIMO nonlinear systems with time-varying parameters and bounded disturbance," *IEEE Trans. Syst., Man, Cybern.*, 2019.

[144] J. Zhou and C. Wen, *Adaptive Backstepping Control of Uncertain Systems: Nonsmooth Nonlinearities, Interactions or Time-Variations.* Springer, 2008.

[145] J. Zhou, C. Wen, and Y. Zhang, "Adaptive output control of a class of time-varying uncertain nonlinear systems," *Nonlinear Dyn. Syst. Theory*, vol. 5, pp. 285–298, 2005.