University of Alberta

NONLINEAR GEOMETRIC OBSERVER DESIGN

by

Yebin Wang ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

in

Controls

Department of Electrical and Computer Engineering

Edmonton, Alberta
Fall 2008

Your file Votre référence
ISBN: 978-0-494-46451-9
Our file Notre référence
ISBN: 978-0-494-46451-9

Abstract

This thesis investigates observer design for uncontrolled multi-output nonlinear systems. The notions of special state coordinates and time scale transformations are explored to generalize an Observer Form (OF) in the geometric framework. Observer designs and stability analysis of the error dynamics are provided. Special coordinate forms are also used to derive two adaptive observer designs.

We begin by discussing a Block Triangular Form (BTF) whose structure readily allows decentralized observer design. The existence conditions of a BTF are established. Unlike with most normal forms in the literature, the existence conditions are derived in an iterative manner using a notion of extended state. A system in BTF has a relatively general dynamic structure which in some cases can make observer design nontrivial. Hence, a Block Triangular Observer Form (BTOF) is presented which is a special case of the BTF. By restricting the dynamics of the BTOF, a straightforward observer design results which is similar to exact error linearization. We consider a Time-scaled Observer Form (TOF) to generalize an OF by including output dependent time scale transformations. This work is further extended to a Time-scaled Block Triangular Observer Form (TBTOF) where time scale transformations have a more general state dependence. Existence conditions for TOF and TBOTF are given. Conditions on time scale transformations to preserve global stability of the error dynamics are presented.

Finally, this thesis discusses the adaptive observer design for two linear and nonlinear uncertainty parameterizations. The linearly parameterized case considers a more general OF with nonlinear output. Next, an adaptive observer design for nonlinearly parameterized systems transformable to an OF with linear output is considered. Local exponential convergence of the estimation error of system state and parameters is established for both cases.

Acknowledgements

I acknowledge my supervisor, Dr. Alan Lynch, for offering me the opportunity to pursue graduate studies and work with him, and for investing significant time and energy to train me. His emphasis on quality, excellent supervision, and persistence have helped me complete my Ph.D. program and this dissertation.

I would like to convey my gratitude to my examination committee members, Drs. Martin Guay, Tongwen Chen, Bob Koch, and Horacio Marquez. Their help and expertise are greatly appreciated.

I owe gratitude to all my colleagues, Cesar, Chris, Dan, Dave, Ed, Karla, Kim, Martin, Rasoul, and Tom, in the Applied Nonlinear Controls Lab for providing such an enjoyable environment to make my Ph.D. study possible. I am grateful for our time spent in discussion.

Last but not least, my acknowledgement goes to my family, especially my wife Xiao, for their love and support during the past five years.

Table of Contents

List of Figures

List of Symbols

Symbol	Description
$\mathbb{R}^+$	set of nonnegative real numbers
$\mathbb{C}^-$	set of complex numbers with negative real part
$A \backslash B$	difference of the sets A and B
$A \times B$	direct product of the sets A and B
$f \circ g$	composition of mappings
$L_f h$	Lie derivative of the function h in the direction of the vector field f
$L_f^k h$	repeated Lie derivatives $L_f^k h = L_f(L_f^{k-1} h), L_f^0 h = h$
$[f, g]$	Lie bracket of vector fields of f and g
$\mathrm{ad}_f^k g$	repeated Lie brackets, $\mathrm{ad}_f^k g = [f, \mathrm{ad}_f^{k-1} g], \mathrm{ad}_f^0 g = g$
I	identity matrix with appropriate dimension
I_k	identity matrix with dimension k
$\delta_{i,j}$	Kronecker delta, $\delta_{i,j} = 0$ for $i \neq j$, and $\delta_{i,i} = 1$
$\lVert \cdot \rVert$	vector or matrix norm

List of Acronyms

Acronym	**Description**
AOF	Adaptive Observer Form
BTF	Block Triangular Form
BTOF	Block Triangular Observer Form
EEL	Exact Error Linearization
EKF	Extended Kalman Filter
GAS	Globally Asymptotically Stable
GES	Globally Exponentially Stable
GUAS	Globally Uniformly Asymptotically Stable
IC	Initial Condition
ISS	Input-to-State Stability
LKY	Lefschetz-Kalman-Yakubovich
LMI	Linear Matrix Inequality
LP	Linearly Parameterized
LTI	Linear Time-Invariant
LTV	Linear Time-Varying
NLP	Nonlinearly Parameterized
NLTV	Nonlinear Time-Varying
NOOF	Nonlinear Output Observer Form
ODE	Ordinary Differential Equation
OF	Observer Form
PDE	Partial Differential Equation
PE	Persistently Excited
PR	Positive Real
SPR	Strictly Positive Real
TBTOF	Time-scaled Block Triangular Observer Form
TOF	Time-scaled Observer Form
TSF	Time Scale Function
UAS	Uniformly Asymptotically Stable

Chapter 1

Introduction

Many real systems are nonlinear and often modeled in an explicit state-space framework

$$\dot{\zeta} = f(\zeta, u), \tag{1.1a}$$

$$y = h(\zeta, u), \tag{1.1b}$$

where $u \in \mathbb{R}^m$ is the system input, $\zeta \in \mathbb{R}^n$ is the system state, $y \in \mathbb{R}^p$ is the measured output, the mapping $f(\cdot, u) : \mathbb{R}^n \to \mathbb{R}^n$ is a C^∞-vector field defined on $\mathbb{R}^n$ for each $u \in \mathbb{R}^m$, and h is a p-vector valued C^∞-output function defined on $\mathbb{R}^n \times \mathbb{R}^m$. The mapping f is called C^k if each of its components possesses continuous partial derivatives of all orders $\leqslant k$ on $\mathbb{R}^n$. If f is C^k for all k then f is C^∞ or smooth. In this thesis we will work in the smooth or C^∞ setting, although everything can be adapted to the C^k-setting for sufficiently large k. The problem of estimating a state from system measurements is referred to as a nonlinear observer design problem when either the system dynamics (1.1a) or measurements (1.1b) are nonlinear functions of their arguments. A nonlinear observer is a system that processes system measurements to asymptotically provide the value of the actual system state. There is practical motivation to being able to estimate a system's state without its direct measurement. This is because direct measurement is often either impossible or prohibitively expensive. Examples of applications which involve estimating a system state include bioreactor microorganism concentration estimation [46, 145], chaotic synchronization for communication systems [119, 123, 93], nonlinear map inversion in robotics [117, 112], parameter identification in uncertain models [44, 32], etc. The value of a state is also commonly required to evaluate a state feedback control law [101, 127, 128]. Although this is an important application of state estimation, the proposed work does not directly address performance of so-called estimated state feedback laws.

Most identity or full-order observers for (1.1), including those discussed in this thesis, take the form

$$\dot{\hat{\zeta}} = f(\hat{\zeta}, u) + k(\hat{\zeta}, \xi, u, y, h(\hat{\zeta}, u)), \tag{1.2a}$$

$$\dot{\xi} = \lambda(\hat{\zeta}, \xi, u, y), \tag{1.2b}$$

with

$$k(\hat{\zeta}, \xi, u, y, h(\hat{\zeta}, u)) = 0, \quad \forall \hat{\zeta} \in \mathbb{R}^n, \xi \in \mathbb{R}^q, u \in \mathbb{R}^m, y \in \mathbb{R}^p,$$

where $\hat{\zeta} \in \mathbb{R}^n$ denotes the state estimate, $\xi \in \mathbb{R}^q$, $\lambda : \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^q$ is a C^∞ mapping, $k : \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^n$ is an observer gain that is C^∞ in its arguments. The augmented state variable ξ, which for some observers is omitted, provides dynamics to the observer gain k. Such state augmentation is used for example in the Kalman Filter [68], Kalman-like observers [17], and observers for systems with unmeasurable inputs [61, 153]. The structure in (1.2a) ensures that if $\hat{\zeta}(0) = \zeta(0)$ then $\tilde{\zeta}(t) = \zeta(t) - \hat{\zeta}(t) = 0$ for all $t \geqslant 0$ [152, 166, 151]. Equivalently, $\zeta - \hat{\zeta} = 0$ is an invariant subspace for the cascade system (1.1), (1.2).

The so-called error dynamics are the differential equations determining the error $\tilde{\zeta}$. For the system (1.1) and observer (1.2) we have the error dynamics

$$\dot{\tilde{\zeta}} = \dot{\zeta} - \dot{\hat{\zeta}} = f(\zeta, u) - f(\hat{\zeta}, u) - k(\hat{\zeta}, \xi, u, y, h(\hat{\zeta}, u)). \tag{1.3}$$

Hence, the observer design problem involves a selection of the observer gain k and the dynamics vector field λ in (1.2) to ensure at least local asymptotic stability of the equilibrium $\tilde{\zeta} = 0$ of (1.3) for all system inputs u. "Local asymptotic stability" of $\tilde{\zeta} = 0$ means for all system inputs u and initial augmented observer state $\xi(0)$ there exists an open neighbourhood $U \subset \mathbb{R}^n$ of the equilibrium such that if $\tilde{\zeta}(0) \in U$ then $\|\zeta(t)\|, \|\xi(t)\|, t \geqslant 0$ are bounded and $\lim_{t\to\infty} \|\tilde{\zeta}(t)\| = 0$. For nonlinear systems, the error dynamics (1.3) is nonlinear and nonautonomous, even when the system is unforced; this makes nonlinear observer design challenging. Further, local stability is sometimes not enough in practice. Often we desire global or semi-global error stability at the origin. "Semi-global" stability is when the error dynamics is globally asymptotically stable at the origin provided the system state remains in a compact subset of $\mathbb{R}^n$.

1.1 Literature Review

In this section we describe some of the main existing approaches to observer design which are relevant to the thesis. Particular emphasis is on geometric design methods where state coordinate transformations are used to put the system dynamics and its output equation into a special form. The discussion below will show that although observer design is a mature and sophisticated area of research, there is presently no globally applicable observer design method. We broadly categorize the design methods discussed.

The linear observer design problem, i.e., when f and h in (1.1) both depend linearly on ζ and u, for Linear Time-Invariant (LTI) and Linear Time-Varying (LTV) systems is fully solved when f and h are known exactly [122, 91]. However, many extensions to this linear observer design problem remain open, e.g. designs for systems with unknown inputs and optimal observer design. Work on observers for LTV systems originated with the Kalman Filter [70, 69]. Later work in [163] considered observer design for LTV systems in a deterministic setting. This last approach is significant since it is a basis for the Exact Error Linearization (EEL) nonlinear observer design discussed below. Although the problem of linear observer design is fully solved, nonlinear observer design continues to receive significant attention. This interest has led to a range of methods being developed over the last three decades. One of the original nonlinear designs is the Extended Kalman Filter (EKF) which is based on a (time-varying) linearization of a nonlinear system about an estimated state trajectory [143]. The EKF for a nonlinear system is in fact a standard Kalman Filter for the system's (LTV) linearization and therefore its performance can be fundamentally limited by its first order approximation. The EKF is sometimes described as heuristic as it provides limited guidance on how to choose its design parameters to ensure estimate error convergence [11, 50, 86]. On the other hand, we note recent efforts to analyze the EKF's convergence [10, 130, 144, 26].

Since the EKF was proposed, many researchers have developed methods which are mathematically well-founded. That is, they provide precise statements about performance subject to various system assumptions which are often verifiable *a priori* using modeling information. Typical measures of performance include the size of the error dynamics system's region of attraction and bounds on norms of the error's trajectory. Nevertheless, providing a rigorous design method only partially addresses the difficulties in nonlinear observer design. Other attributes are also important: applicability to a broad system class, constructive nature of the method, robustness to model error and noise, and ease of implementation. The given application will determine which criteria are most important in order for a particular method to be practically useful.

The high-gain observer design method is a popular and relatively constructive method based on a nonlinear observability assumption which is often satisfied in practice on some open set of state space [46]. Generalizations, including industrially applications, of the high-gain method are in [12, 17, 22, 38, 42, 45, 47, 48, 141, 148]. The main drawback of the high-gain method is large observer gain values [12]. Hence, the high-gain method can suffer from a lack of robustness to measurement noise. Roughly speaking, we can attribute this lack of robustness to a gain conservatively designed to overpower the worst-case nonlinearity which is measured by a Lipschitz constant. Below, we discuss EEL methods which rely on exact cancelation of system nonlinearity (as opposed to overpowering it). Exact cancelation extracts additional structure from the system which can avoid high gain. We remark that the high-gain observer theory for the multi-output case appears to be incomplete [42].

Time-invariant EEL is a geometric approach to nonlinear observer design which dualizes State Feedback Linearization [66, 62]. The method originated independently in [84, 19] shortly after the development of State Feedback Linearization [66, 62]. EEL attempts to find a change of state coordinates $z = \Phi(\zeta)$ which puts the unforced multi-output system

$$\begin{aligned} \dot{\zeta} &= f(\zeta), \\ y &= h(\zeta), \end{aligned} \tag{1.4}$$

into Observer Form (OF) without output transformation [166, 167, 109]

$$\dot{z} = Az + \gamma(y), \tag{1.5a}$$

$$y = Cz, \tag{1.5b}$$

where matrices $A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{p\times n}$ are block diagonal

$$\begin{aligned} A &= \text{blockdiag}(A^1, \dots, A^p), \\ C &= \text{blockdiag}(C^1, \dots, C^p), \end{aligned} \tag{1.6}$$

and each pair

$$A^i = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{\lambda_i \times \lambda_i}, \tag{1.7}$$

$$C^i = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{1 \times \lambda_i}$$

is in dual Brunovsky Form, and $\lambda_i, 1 \leqslant i \leqslant p$ are observability indices in Definition A.0.1. The existence conditions of (1.5) are given by Theorem A.0.2. Transformed into OF, a Luenberger-like observer

$$\dot{\hat{z}} = A\hat{z} + \gamma(y) + L(y - C\hat{z}) \tag{1.8}$$

yields a homogeneous, LTI error dynamics system

$$\dot{\tilde{z}} = (A - LC)\tilde{z},$$

where $\tilde{z} = z - \hat{z}$. EEL can be difficult to apply in practice as it requires an integrability condition to hold. Condition (iii) in Theorem A.0.2 is non-generic and restricts the range of applicability of EEL. Numerous papers have generalized and improved the original work on EEL which considered unforced, single-output systems. Significant improvements made the original work more constructive [75, 74]. A similar approach to [75, 74] is found in [124, 125] which presents a more general constructive algorithm incorporating an output transformation for multi-input, multi-output, control-affine systems. Other work in [92] also contains constructive conditions given in terms of rank conditions of matrices for the single-output, nonautonomous general system case without output transformation. The multi-output generalization of [92] is in [166], and a time-invariant version of [166] is in [167]. A substantial generalization is obtained in [90, 6, 67] by immersing an n-dimensional system into a linear N-dimensional one with $N \geqslant n$. Constructive algorithms are discussed in [7, 121, 8] for a class of systems and a specific structure of dynamic extensions. Although immersion is a generalization of state diffeomorphisms used in [84, 19], the algorithms to check immersibility of the system are difficult to describe for the most general form of dynamic extensions. Work on discrete-time EEL includes for example [89, 30] and the necessary and sufficient conditions take a similar form as the continuous-time case. However, the unit vector fields are constructed differently.

Even when an OF (1.5) exists, computing the change of coordinates is often impractical since it requires the closed-form expression for the inverse of a non-linear map. The work in [19] presents a so-called Extended Luenberger Observer which computes an approximately linearizing observer for integrable systems using straightforward operations of differentiation and integration. This approach is extended to a Generalized Observer Form with output transformation for single-output systems in [170] and for multi-output systems in [20]. Work in [113] characterizes the OF existence conditions by the exactness of one-forms. This exactness condition is weakened by approximating the nonexact one-forms by exact ones up to some order. Other work which provides approximately linear error dynamics is in [24, 99, 98]. Here optimization is used to uniformly make the system error dynamics approximately linear in some coordinates.

Important recent work in [72, 73, 82, 83] eliminates the constraint (1.5b) which forces the output to be linear in the design coordinates. Removing this constraint and applying a result on existence of solutions to systems of first-order PDEs [126], sufficient conditions are provided for the existence of a new Observer Form where the output can be in general a nonlinear function of state in the new coordinates. These conditions are met by a larger class of systems than those admitting OF (1.5). Instead of an analytic and invertible solution of PDEs, work in [2] requires the solution is only continuous and uniformly injective. However, the method can still have practical limitations since in general the solutions of the PDEs are only guaranteed to exist in a potentially small neighborhood of an operating point. On the other hand, the local nature of this result can be considered a drawback of any geometric method.

Work on time-invariant error linearization is extended in [13, 53, 55, 56, 57, 54] to consider state transformation into bilinear and state-affine forms plus input-output injection vector. In the new z-coordinates the system has the general form

$$\begin{aligned} \dot{z} &= A(u)z + \gamma(u, y), \\ y &= Cz, \end{aligned} \tag{1.9}$$

where $A \in \mathbb{R}^{n \times n}$ is C^∞ matrix-valued function of the input $u \in \mathbb{R}^m$, γ is an input-output injection vector field, and C is defined in (1.6). Extensions of the form (1.9) are considered in [15] where an output transformation is introduced for a single output system. Work in [17, 18] considers transformation to state affine forms plus triangular nonlinearity using immersion techniques. As opposed to systems admitting an OF (1.5), which are observable uniformly for all inputs, the problem of observer design for state-affine systems is complicated by the existence of inputs which makes them unobservable. Even though generically nonlinear systems have few bad inputs [146], state estimation can be difficult for inputs rendering the system sufficiently unobservable. Recent work [132, 52] considers a generalization of EEL which incorporates an output dependent time scale transformation for a single output unforced nonlinear system. Time scale transformations lead to an additional degree of freedom when transforming the system to OF. Similar to the work on time-varying linearization [17, 13, 53, 55, 56, 57, 54, 15, 18], the system's dynamics in new state coordinates has a LTV part plus output injection. The idea of time scale transformation has been applied in state feedback design in [137, 131, 51].

Most of the design procedures discussed above assume knowledge of an exact system model. Evidently, this assumption is rarely valid in practice and a design's robustness to model error is desirable. Significant work has been placed on designing adaptive observers to achieve robustness by simultaneously estimating the unknown parameters and system states. Compared to observer design assuming a known model, the design of adaptive observers is more challenging and error convergence can only be proven under restrictive conditions. Adaptive observer design for linear systems has been largely solved [31, 58, 80, 81, 88, 97, 96, 115, 133, 171]. On the other hand, the nonlinear case remains only partially solved. The existing work is applicable to a small class of nonlinear systems state equivalent to a certain form with a specific uncertainty parameterization. For example, uncertainty is assumed to be linearly parameterized (LP). Results on nonlinear adaptive observer design are usually developed by combining linear adaptive theory and a nonlinear geometric framework [11, 105, 102, 103, 104, 16, 37, 169, 14, 168]. For example, work in [11, 105, 102, 103, 104] transforms the system to an Adaptive Observer Form (AOF) and a LTI observer design in [96] is applied. The resulting adaptive observer includes a state estimator, a parameter identifier, and an auxiliary filter which is excited by the coefficient of unknown parameters. A different approach to that taken in [11, 105, 102, 103, 104] is in [37] which considers a different class of system whose dynamics are Lipschitz and have a LP uncertainty. Based on results in [129], [37] solves an adaptive observer problem using LMIs. The work in [105, 103, 104, 37] was unified under the same framework in [16]. Here Lyapunov's direct method is applied to design the parameter update law. This work treats the AOF in [103] and [37] as special cases. However, the results are given as non-constructive existence conditions which could be difficult to apply in practice.

Many practical systems such as biochemical processes and machines with friction often contain unknown parameters that enter the systems' dynamics nonlinearly, e.g. friction dynamics in [4] and bacterial growth systems in [43, 172]. From a theoretical point of view, adaptive observer and control design of nonlinearly parameterized (NLP) systems is interesting and challenging. The major difficulty in extending approaches developed for LP systems to the NLP case is due to the non-convexity of the underlying cost function which prevents the application of a gradient algorithm [95, 25]. There is little known general theory on adaptive observer design for NLP systems. Some work on specific systems includes a microbial growth process application [172]. For the control problem, backstepping and high-gain techniques are applied in [108] to a class of system admitting an OF. Recent work [9, 3, 95, 142, 78] proposes a control design using a novel parameter update law based on a min-max optimization strategy instead of a gradient algorithm. Other approaches include a linear approximation design [71], a control Lyapunov function method [49], a parameter separation technique combined with feedback domination design [94], and a backstepping design [150]. The work above is not generally applicable and is restricted by factors including the type of parameterization and nonlinear system structure. For example, work [3] requires convex/concave parameterizations, [95] considers general nonlinear parameterizations where closed-form solutions of a min-max optimization problem are nontrivial, [79, 78, 150] discuss special classes of triangular systems, and [94] investigates a class of feedback linearizable systems. Evidently, further research is required to solve the control problem for NLP systems.

1.2 Contributions of the Thesis

Chapter 2 presents a Block Triangular Form (BTF). The primary reason for transforming a system into BTF is that it permits lower dimensional observer designs to be performed subsystem-at-a-time while effectively treating "upper" subsystem states as known measurements. Reducing the dimension of the observer is important as it simplifies the design. Existence conditions for the BTF are weaker than those of the OF [85, 165, 167] and can therefore be applied to a broader class of systems. The main contribution of this chapter is to establish new BTF existence conditions with the ordering of observability indices. The ordering constraint is finally removed and BTF existence conditions are derived which are less constructive. This chapter is based on [154, 158].

Chapter 3 discusses the existence conditions of a Block Triangular Observer Form (BTOF), BTOF-based observer design, and the stability of the error dynamics. As a special case of BTF, the existence conditions of the BTOF are weaker than the OF but stronger than the BTF. A main contribution of this chapter is to provide generalized BTOF existence conditions. Our work shows that different BTOF existence conditions can be formulated depending on which system variables are treated as parameters. Treating variables as parameters leads to more restrictive existence conditions. Another contribution is to provide a rigorous proof of global and semi-global error dynamics stability. The resulting stability proofs lead to a simple condition on the observer gain. This chapter is based on [155, 159].

Chapter 4 investigates a Time-scaled Observer Form (TOF). Output dependent time scale transformations are introduced to enlarge the class of systems admitting OF. A main contribution of this chapter is to extend the single output time scaling work [52, 132] to multi-output systems. TOFs with single and multiple time scale transformations are discussed and existence conditions are provided. The necessary and sufficient conditions on time scale transformations to preserve the global exponential stability of the error dynamics are presented. This chapter is based on [160, 161].

Chapter 5 considers a Time-scaled BTOF (TBTOF) which generalizes a BTOF to include time scale transformations. Due to a block triangular structure, the time scale transformation introduced in this chapter allows more general dependence than the TOF case. A main contribution of this chapter is to derive the existence conditions for a TBTOF. The error dynamics stability for TBTOF-based observer design is studied. This chapter is based on [39, 156, 157].

Chapter 6 discusses adaptive observer design for uncertain unforced nonlinear multi-output systems. A relatively large class of systems admitting the Nonlinear Output Observer Form (NOOF) in [73, 82] is considered. Hence, the proposed adaptive observer should also be widely applicable. Due to the system's nonlinear dependence of the output in the new coordinates, only local exponential convergence of the error dynamics is achieved with the proposed adaptive observer. Next, we consider the adaptive observer design for NLP systems in OF. Local exponential convergence of the state estimation error of system state and parameters is established.

Chapter 2

A Block Triangular Form (BTF)

In this chapter and in Chapters 3–5, we consider special coordinates which enable observer design for the uncontrolled multi-output system

$$\begin{aligned} \dot{\zeta} &= f(\zeta), \\ y &= h(\zeta) \end{aligned} \tag{2.1}$$

with C^{∞} vector fields $f : \mathbb{R}^n \to \mathbb{R}^n$, and C^{∞} output functions $h : \mathbb{R}^n \to \mathbb{R}^p$. Throughout this thesis we consider the design of observer which provides convergent estimate error dynamics when $t \to \infty$. Hence, we assume that (2.1) has a unique solution on $[t_0, \infty)$. The Observer Form (OF) and related special forms, such as the Block Triangular Form (BTF) considered in this chapter, define coordinates in which the system has a special structure which facilitates observer designs. This chapter considers a BTF which permits lower dimensional observer designs to be performed subsystem-at-a-time while effectively treating "upper" subsystem states as known measurements. The idea of performing a design in subsystems came from work in [134]. We consider existence conditions for the BTF. These conditions are weaker than those of the (Block Triangular) OFs used in EEL [134, 165, 167] and can therefore be applied to a broader class of systems.

This chapter is organized as follows. In Section 2.1 we introduce some fundamental concepts and system forms. Two different existence conditions for a BTF are established in Sections 2.2 and 2.3. Examples of the construction of BTF coordinates and a BTF-based observer design are presented in Section 2.4.

2.1 Introduction

Before introducing a number of special state coordinates, we define some notation. Given a C^∞ vector field $f : \mathbb{R}^n \to \mathbb{R}^n$, and a C^∞ function $\alpha : \mathbb{R}^n \to \mathbb{R}$, the function

$$L_f\alpha = \frac{\partial\alpha}{\partial x}f$$

is the *Lie derivative* of α along f. The *differential* or *gradient* of a C^∞ function $\alpha : \mathbb{R}^n \to \mathbb{R}$ is denoted $\mathrm{d}\alpha$ and has local coordinate description

$$\mathrm{d}\alpha = \frac{\partial\alpha}{\partial x} = (\frac{\partial\alpha}{\partial x_1}, \ldots, \frac{\partial\alpha}{\partial x_n}).$$

Given a C^∞ one-form $\omega : \mathbb{R}^n \to \mathbb{R}^n$ and a vector field $f : \mathbb{R}^n \to \mathbb{R}^n$, the *inner product* of ω and f is the function

$$\langle\omega(x), f(x)\rangle = \sum_{i=1}^{n} \omega_i(x) f_i(x),$$

where ω_i, f_i are the components of ω, f in local coordinates, respectively. The *Lie bracket* of two C^∞ vector fields $f, g : \mathbb{R}^n \to \mathbb{R}^n$ is defined as

$$[f, g] = \frac{\partial g}{\partial x}f - \frac{\partial f}{\partial x}g.$$

See [64, 120] for further details.

The approach taken for observer design in this thesis relies on special coordinates which exist if we assume system (2.1) is observable in some sense. A number of notions of observability exist [60, 118, 85]. For instance, Definition A.0.1 uniquely defines observability indices to ensure a single normal form. However, this uniqueness is actually unnecessarily restrictive for transforming a system into a special coordinate form. Hence, we introduce an alternate definition of observability in a similar manner as in [109, 85].

Definition 2.1.1. System (2.1) is locally observable in U_0 with indices $\lambda_i, 1 \leqslant i \leqslant p$, if after suitable reordering of the h_i's,

$$\dim(\mathrm{span}\{L_f^{j-1}\mathrm{d}h_i(\zeta),\ 1 \leqslant j \leqslant \lambda_i;\ 1 \leqslant i \leqslant p\}) = n, \tag{2.2}$$

for every $\zeta \in U_0$. A system is globally observable if $L_f^{j-1}h_i, 1 \leqslant j \leqslant \lambda_i, 1 \leqslant i \leqslant p$ are globally defined coordinates on $\mathbb{R}^n$. □

Note that in [120, 60] the notion of indistinguishability of states is introduced and used to derive a rank condition for observability. For the mapping $x = T(\zeta) = (h_1, \cdots, L_f^{\lambda_1 - 1} h_1, \cdots, L_f^{\lambda_p - 1} h_p)^T$ to be a global diffeomorphism, in addition to satisfying (2.2) for all $\zeta \in \mathbb{R}^n$, the following condition must hold [164, 77]

$$\lim_{\|\zeta\| \to \infty} \|T(\zeta)\| = \infty.$$

The indices λ_i in Definition 2.1.1 are not uniquely defined [134].

Example 2.1. The non-uniqueness of observability indices in Definition 2.1.1 is demonstrated in the two-output system

$$\dot{\zeta} = f(\zeta) = \begin{pmatrix} \zeta_2^1 \\ \zeta_3^1 \\ \zeta_1^1 + \zeta_2^1 \zeta_3^1 \\ \zeta_1^2 + \zeta_3^1 \end{pmatrix},$$

$$y = \begin{pmatrix} h_1(\zeta) \\ h_2(\zeta) \end{pmatrix} = \begin{pmatrix} \zeta_1^1 \\ \zeta_1^2 \end{pmatrix},$$

where $\zeta = (\zeta_1^1, \zeta_2^1, \zeta_3^1, \zeta_1^2)^T$. The system is globally observable with indices $(3, 1)$ since the observability matrix with indices $(3, 1)$ has the representation in ζ-coordinates

$$Q = \begin{pmatrix} \mathrm{d}h_1 \\ \mathrm{d}L_f h_1 \\ \mathrm{d}L_f^2 h_1 \\ \mathrm{d}h_2 \end{pmatrix} = I_4, \quad \forall \zeta \in \mathbb{R}^4,$$

and the mapping

$$T(\zeta) = \begin{pmatrix} h_1 \\ L_f h_1 \\ L_f^2 h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \zeta_1^1 \\ \zeta_2^1 \\ \zeta_3^1 \\ \zeta_1^2 \end{pmatrix}$$

is globally defined. On the other hand, the observability matrix with indices $(2, 2)$ has the representation in ζ-coordinates

$$Q = \begin{pmatrix} \mathrm{d}h_1 \\ \mathrm{d}L_f h_1 \\ \mathrm{d}h_2 \\ \mathrm{d}L_f h_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

which is nonsingular for all $\zeta \in \mathbb{R}^4$. Verifying that the mapping

$$T(\zeta) = (h_1, L_f h_1, h_2, L_f h_2)^T$$

globally defines new coordinates, we confirm that the system is also globally observable with indices $(2,2)$. On the other hand, according to Definition A.0.1, the observability indices make use of the lowest order independent Lie derivatives and are uniquely defined as $(2,2)$. $\square$

When system (2.1) is locally observable according to Definition 2.1.1, we can define new coordinates $x_j^i = L_f^{j-1}h_i, 1 \leqslant j \leqslant \lambda_i, 1 \leqslant i \leqslant p$ to locally transform system (2.1) into an Observable Form whose ith subsystem is

$$
\begin{aligned}
\dot{x}^i = f^i(x) &= \begin{pmatrix} x_2^i \\ x_3^i \\ \vdots \\ \varphi^i(x) \end{pmatrix}, \\
y_i &= x_1^i,
\end{aligned}
\tag{2.3}
$$

where $\varphi^i(x) = L_f^{\lambda_i}h_i$, and $x^i = (x_1^i, \ldots, x_{\lambda_i}^i)^T$. The vector field $f^i(x)$ can be written using two notations: in matrix format (2.3) or alternatively

$$
f^i(x) = \sum_{k=1}^{\lambda_i - 1} x_{k+1}^i \frac{\partial}{\partial x_k^i} + \varphi^i(x) \frac{\partial}{\partial x_{\lambda_i}^i}.
$$

Treating the lower subsystem outputs as inputs, we denote $\varphi^i(x)$ as $\varphi^i(x^{<i>}, y_{[i+1,p]}^{<\alpha^i>})$, where

$$
\begin{aligned}
x^{<i>} &= ((x^1)^T, \ldots, (x^i)^T)^T \in \mathbb{R}^{\nu_i}, \quad \nu_i = \sum_{k=1}^{i} \lambda_k, \\
y_{[i+1,p]}^{<\alpha^i>} &= ((y_{i+1}^{<\alpha_{i+1}^i>})^T, \ldots, (y_p^{<\alpha_p^i>})^T)^T, \\
y_j^{<\alpha_j^i>} &= (y_j, \dot{y}_j, \ldots, y_j^{(\alpha_j^i)})^T,
\end{aligned}
$$

and the non-negative indices $\alpha_j^i, 1 \leqslant i < j \leqslant p$ denote the highest order of time derivative of y_j appearing in φ^i. Since $x_k^j = y_j^{(k-1)}$, there can be time derivatives of y_j of order at most $\lambda_j - 1$ appearing in φ^i, i.e.,

$$
\alpha_j^i \leqslant \lambda_j - 1, \quad 1 \leqslant i < j \leqslant p.
$$

Putting system (2.1) into Observable Form allows us to derive the existence conditions in the framework of [41].

Assuming that system (2.1) is observable according to Definition 2.1.1 and can be interconnected in a state cascade output feedback form, system (2.1) is said to be in BTF [134, 139, 140, 154]. That is, in BTF the ith subsystem has the form

$$\begin{aligned} \dot{\bar{x}}^i &= \bar{f}^i(\bar{x}^{<i>}, y_{[i+1,p]}), \\ y_i &= \bar{h}_i(\bar{x}), \end{aligned} \tag{2.4}$$

where $\bar{x}^i = (\bar{x}^i_1, \dots, \bar{x}^i_{\lambda_i})^T$ denotes the state vector of the ith subsystem, $\bar{x}^{<i>} = ((\bar{x}^1)^T, \dots, (\bar{x}^i)^T)^T$ denotes the state vector of the upper i subsystems, $y_{[i+1,p]} = (y_{i+1}, \dots, y_p)^T$, and $\bar{h}_i$ is the output function expressed in the $\bar{x}$-coordinates. We denote $\bar{f}^i(\bar{x}^{<i>}, y_{[i+1,p]})$ in matrix format $(\bar{f}^i_1(\bar{x}^{<i>}, y_{[i+1,p]}), \dots, \bar{f}^i_{\lambda_i}(\bar{x}^{<i>}, y_{[i+1,p]}))^T$ or the alternate form

$$\bar{f}^i(\bar{x}^{<i>}, y_{[i+1,p]}) = \sum_{k=1}^{\lambda_i} \bar{f}^i_k(\bar{x}^{<i>}, y_{[i+1,p]}) \frac{\partial}{\partial \bar{x}^i_k},$$

where $f^i_k(\bar{x}^{<i>}, y_{[i+1,p]})$ is the kth component of $\bar{f}^i(\bar{x}^{<i>}, y_{[i+1,p]})$. In order to simplify presentation we make use of both notations in the sequel. We remark that a BTF considered in [139] is a special case of (2.4) in that $y_i = \bar{x}^i_1$. This special case can be useful for observer design and is discussed below.

As shown in Section 2.4, non-uniqueness of λ_i can add a degree of freedom when checking BTF existence conditions. That is, for a particular choice of indices λ_i the system may be transformable to a BTF and for another choice it is not.

2.2 BTF Existence Conditions : State Approach

Two approaches have been proposed to solve the existence conditions of a BTF. For the ith subsystem, the first approach, which is taken in [134, 139] and Section 2.3, treats only part of variables in the ith subsystem dynamics as states, and the rest of variables as *parameters*. This parameter assumption reduces the complexity of the existence conditions, but it also limits the range of applicability for the resulting conditions. It is worth noting that in [134, 139], where only the current subsystem state x^i and output derivatives $y_j^{<\alpha^i_j>}, i+1 \leqslant j \leqslant p$ are considered as states, incomplete existence conditions for the BTF result due to the implicit assumption that the upper subsystem states are parameters. In Section 2.3, only $y_{[1,i-1]}$ is treated as a parameter. Another approach which is considered in this section treats all variables in the ith subsystem dynamics as *states*. However, this approach in general leads to very complex existence conditions. We therefore impose the ordering constraint on observability indices in this section to derive relatively straightforward existence conditions.

2.2.1 Main Result

Theorem 2.2.1. *Given an observable system (2.1) with $\lambda_i \geqslant \max_{i+1\leqslant j\leqslant p} \lambda_j - 1$, with its first $i-1$ subsystems in BTF, and its last $p-i+1$ subsystems in Observable Form, define the prolonged vector field F_e^i and the extended state x_e^i as*

$$F_e^i(x_e^i) = \sum_{j=1}^{i-1} \bar{f}^j(\bar{x}^{<j>}, y_{[j+1,p]}) + f^i(x_e^i) + \sum_{j=i+1}^{p} \sum_{k=0}^{\alpha_j^i} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}, \tag{2.5a}$$

$$x_e^i = ((\bar{x}^{<i-1>})^T, (x^i)^T, (y_{[i+1,p]}^{<\alpha_{i+1}^i>})^T)^T \in \mathbb{R}^N, \tag{2.5b}$$

where $N = \nu_i + L, L = \sum_{j=i+1}^{p}(\alpha_j^i + 1)$. The ith subsystem in (2.3) can be locally transformed into BTF in (2.4) by the extended state transformation $\bar{x}^i = \Psi^i(x_e^i)$ and the first $i-1$ subsystems remain identical if and only if

$$\left[\mathrm{ad}_{F_e^i}^q \frac{\partial}{\partial y_j^{(\alpha_j^i)}}, \mathrm{ad}_{F_e^i}^r \frac{\partial}{\partial y_l^{(\alpha_l^i)}}\right] = 0, \quad \begin{cases} 1 \leqslant i < j, l \leqslant p; \\ 0 \leqslant q \leqslant \alpha_j^i; \\ 0 \leqslant r \leqslant \alpha_l^i. \end{cases} \tag{2.6}$$

The extended state transformation Ψ^i is a solution of the $\lambda_i \cdot L$ PDEs

$$\frac{\partial \Psi^i(x_e^i)}{\partial x_e^i} \mathrm{ad}_{F_e^i}^q \frac{\partial}{\partial y_j^{(\alpha_j^i)}} = 0, \quad 0 \leqslant q \leqslant \alpha_j^i; i+1 \leqslant j \leqslant p. \tag{2.7}$$

Remark 2.2.2. The notion of a prolonged vector field comes from [41]. The main differences between Theorem 2.2.1 and previous work [139, Thm. 1] involve the definitions of the extended state x_e^i and the prolonged vector field F_e^i. Previously, x_e^i and F_e^i did not include the upper subsystem states. That is,

$$F_e^i(x_e^i) = f^i(x^{<i>}, y_{[i+1,p]}^{<\alpha^i>}) + \sum_{j=i+1}^{p} \sum_{k=0}^{\alpha_j^i} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}, \tag{2.8a}$$

$$x_e^i = ((x^i)^T, (y_{[i+1,p]}^{<\alpha^i>})^T)^T. \tag{2.8b}$$

The necessary and sufficient Lie bracket conditions in [139, Thm. 1] are the same as (2.6). The difference in the definitions of x_e^i and F_e^i leads to different vector fields appearing in conditions (2.6). This difference will be demonstrated by two counterexamples in Section 2.2.2. However, the Lie bracket conditions using either (2.5) or (2.8) are equivalent in two simple cases. The first case is when $\alpha_j^i \leqslant 1, i+1 \leqslant j \leqslant p, 1 \leqslant i \leqslant p-1$. In this case we have $\bar{x}_1^j = \bar{x}_1^j(y), 1 \leqslant j \leqslant i-1$ and the Lie bracket conditions using (2.5) or (2.8) are equivalent. That is, the Lie bracket conditions reduces to φ^i not having terms involving $\dot{y}_j\dot{y}_l, i+1 \leqslant j, l \leqslant p$. The other case is when the first subsystem is transformed. No upper subsystem for the first subsystem implies the same prolonged vector fields and extended state for both definitions. Hence, for the two subsystem case, which is considered for the examples in [139], the Lie bracket conditions are equivalent. $\square$

Remark 2.2.3. To simplify notation we have specified that the extended state transformation Ψ^i depends on x_e^i, however Ψ^i can only be a function of

$$\bar{x}^{<i-1>}, x^i, y_{i+1}^{<\alpha_{i+1}^i-1>}, \ldots, y_p^{<\alpha_p^i-1>}.$$

This is because allowing $y_j^{(\alpha_j^i)}$ dependence in Ψ^i is not required to remove $y_j^{(\alpha_j^i)}$ from the ith subsystem. As shown in the proof of Theorem 2.2.1, provided that conditions (2.6) hold, such a Ψ^i can always be constructed. $\square$

Remark 2.2.4. Theorem 2.2.1 removes all time derivatives of lower subsystem outputs from the ith subsystem. In order to reduce the order of time derivatives of y_j by $\alpha_j^i - \beta_j^i, i < j \leqslant p$, we change the range of indices in (2.6) to $0 \leqslant q \leqslant \alpha_j^i - \beta_j^i$; $0 \leqslant r \leqslant \alpha_l^i - \beta_l^i, i < j, l \leqslant p$. $\square$

Remark 2.2.5. Given an observable system (2.1) with $\lambda_i \geqslant \max_{i+1 \leqslant j \leqslant p} \lambda_j - 1$, the first $i-1$ subsystems have a globally defined BTF, and the last $p-i+1$ subsystems have a globally defined Observable Form, The ith subsystem in (2.3) can be globally transformed into BTF in (2.4), and the first $i-1$ subsystems remain unchanged if and only if (2.6) are satisfied on $\mathbb{R}^N$ and the vector fields $\mathrm{ad}^k_{F_e^i}\frac{\partial}{\partial y_j^{(\alpha_j^i)}}, 0 \leqslant k \leqslant \alpha_j^i, i+1 \leqslant j \leqslant p$ are complete. A vector field f is said to be complete if the domain of all its integral curves can be extended to all of $\mathbb{R}$ [162, 21]. Thus if through every point there is an integral curve that exists for all time, we show that a vector field is complete. A vector field is incomplete if we find one integral curve that cannot be extended for all time. $\square$

Next, two examples are given to illustrate Theorem 2.2.1 and its differences from the previous work in [139, Thm. 1].

2.2.2 Counterexamples

The first example shows the existence conditions based on (2.8) are not necessary, and the second example shows the conditions are not sufficient. We consider a three-output system with indices $\lambda_1 = 2, \lambda_2 = \lambda_3 = 4$

$$
\begin{aligned}
\dot{x}^1 &= f^1(x) = \begin{pmatrix} x_2^1 \\ x_1^3 \end{pmatrix}, \\
\dot{x}^2 &= f^2(x) = \begin{pmatrix} x_2^2 \\ x_3^2 \\ x_4^2 \\ x_2^1 y_3^{(3)} + \kappa(x) \end{pmatrix}, \\
\dot{x}^3 &= f^3(x) = \begin{pmatrix} x_2^3 \\ x_3^3 \\ x_4^3 \\ \varphi^3(x) \end{pmatrix}, \\
y &= \begin{pmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \end{pmatrix},
\end{aligned}
\tag{2.9}
$$

where initially we take $\kappa(x) = x_2^2 x_4^2/(x_2^1)^2$. The first subsystem is already in BTF coordinates (i.e., $\bar{x}^1 = x^1$). Hence, we only consider the transformation of the second subsystem into BTF. Since $\alpha_3^2 = 3$, we define the extended state

$$x_e^2 = ((\bar{x}^1)^T, (x^2)^T, (y_3^{<3>})^T)^T,$$

and the prolonged vector field

$$F_e^2 = ((\bar{f}^1)^T, (f^2)^T, \dot{y}_3, \ddot{y}_3, y_3^{(3)}, y_3^{(4)})^T.$$

We can verify that the Lie bracket conditions (2.6) of Theorem 2.2.1 hold. Using PDEs (2.7) to solve for the state transformation $\Psi^2(x_e^2)$, we can transform the second subsystem into a BTF

$$
\begin{aligned}
\dot{\bar{x}}_1^2 &= \bar{x}_2^1 y_3 - \bar{x}_2^2 \bar{x}_2^1, \\
\dot{\bar{x}}_2^2 &= \bar{x}_3^2 + \frac{3}{2\bar{x}_2^1}((y_3)^2 - (\bar{x}_2^2)^2), \\
\dot{\bar{x}}_3^2 &= \bar{x}_4^2 - \frac{1}{(\bar{x}_2^1)^2}(5(\bar{x}_2^2)^3 - 5\bar{x}_2^2(y_3)^2 - 2(\bar{x}_2^2)^2 y_3 - 3\bar{x}_3^2\bar{x}_2^1\bar{x}_2^2 + 3y_3\bar{x}_3^2\bar{x}_2^1 + 2(y_3)^3), \\
\dot{\bar{x}}_4^2 &= \frac{\bar{x}_2^2 - y_3}{(\bar{x}_2^1)^3}(6\bar{x}_3^2\bar{x}_2^1\bar{x}_2^2 - 4\bar{x}_4^2(\bar{x}_2^1)^2 - 7(\bar{x}_2^2)^3 - 5\bar{x}_2^2(y_3)^2 \\
&\quad - 27(\bar{x}_2^2)^2 y_3 + 18 y_3 \bar{x}_3^2 \bar{x}_2^1 + 15(y_3)^3), \\
y_2 &= \bar{x}_1^2.
\end{aligned}
$$

Following the results in [139, Thm. 1], we define

$$x_e^2 = ((x^2)^T, (y_3^{<3>})^T)^T,$$

and

$$F_e^2 = ((f^2)^T, \dot{y}_3, \ddot{y}_3, y_3^{(3)}, y_3^{(4)})^T.$$

It is trivial to verify the Lie bracket conditions (2.6) are not satisfied, which implies the second subsystem is not transformable to BTF. This leads to a contradiction as a BTF was computed above. Hence, the Lie bracket conditions (2.6) with definition (2.8) are not necessary.

In order to show the conditions of [139, Thm. 1] are not sufficient, we consider system (2.9) with $\kappa(x) = 0$. As above, we only consider the transformation of the second subsystem. Defining F_e^2 and x_e^2 to include the first subsystem as in Theorem 2.2.1, we observe the Lie bracket conditions (2.6) are satisfied nowhere

$$\left[\mathrm{ad}^2_{F_e^2}\frac{\partial}{\partial y_3^{(3)}}, \mathrm{ad}^3_{F_e^2}\frac{\partial}{\partial y_3^{(3)}}\right] = 2\frac{\partial}{\partial x_4^2}.$$

Alternately as in [139, Thm. 1], defining F_e^2 and x_e^2 to exclude the first subsystem, we observe the Lie bracket conditions (2.6) for $\alpha_3^2 = 3, j = l = 3$ are satisfied, and the second subsystem appears to be transformable to BTF for some Ψ^2. Consequently, there should also exist changes of coordinates ψ^{21}, ψ^{22} which successively remove $y_3^{(3)}$ and $(\ddot{y}_3, \dot{y}_3)$ from Σ^2. However, we show that attempting to remove $y_3^{(3)}$ and then $(\ddot{y}_3, \dot{y}_3)$ from the second subsystem using [139, Thm. 1] leads to a contradiction and the conditions in this result are therefore not sufficient.

Remark 2.2.4 can be applied to [139, Thm. 1] and in order to remove only $y_3^{(3)}$ from the second subsystem, we verify

$$\left[\mathrm{ad}_{F_e^2}\frac{\partial}{\partial y_3^{(3)}}, \frac{\partial}{\partial y_3^{(3)}}\right] = 0,$$

and solve PDEs (2.7) to get a transformation ψ^{21} which removes $y_3^{(3)}$. Without any loss of generality (shown below), we choose the solution

$$\psi^{21}(x_e^2) = \begin{pmatrix} x_1^2 \\ x_2^2 - x_2^1 y_3 \\ x_3^2 - x_2^1 \dot{y}_3 \\ x_4^2 - x_2^1 \ddot{y}_3 \end{pmatrix},$$

which leads to

$$\dot{\bar{x}}^{21} = \begin{pmatrix} \bar{x}_2^{22} + x_2^1 y_3 \\ \bar{x}_3^2 x_2^1 - (y_3)^2 \\ \bar{x}_4^2 - y\dot{y}_3 \\ -y_3\ddot{y}_3 \end{pmatrix},$$

where the state of the second subsystem is re-notated as $\bar{x}^{21}$. Next, we define

$$x_e^{21} = ((\bar{x}^{21})^T, (y_3^{<2>})^T)^T,$$
$$F_e^{21} = ((\bar{f}^{21})^T, \dot{y}_3, \ddot{y}_3, y_3^{(3)})^T,$$

where $\dot{\bar{x}}^{21} = \bar{f}^{21}(\bar{x}_e^{21})$, and verify the Lie bracket conditions for removing $(\ddot{y}_3, \dot{y}_3)$ are not satisfied

$$\left[\mathrm{ad}_{F_e^{21}} \frac{\partial}{\partial \ddot{y}_3}, \mathrm{ad}^2_{F_e^{21}} \frac{\partial}{\partial \ddot{y}_3}\right] = -2\frac{\partial}{\partial \bar{x}_4^{21}}.$$

By [139, Thm. 1] this means there exists no change of coordinates ψ^{22} to transform the second subsystem into a BTF, but because Ψ^2 was previously shown to exist, there must exist ψ^{22} given by $\psi^{22}(x_e^{21}) = \Psi^2(x_e^2) \circ (\psi^{21}(x_e^2))^{-1}$. This leads to a contradiction. Above, we made a specific choice for ψ^{21}. However, this choice will not affect the existence of ψ^{22}. For instance, suppose we consider an alternate transformation $\bar{\psi}^{21}$ in place of ψ^{21}, and that after transformation by $\bar{\psi}^{21}$ there exists $\bar{\psi}^{22}$ to put the second subsystem in a BTF. If we define $\pi = \bar{\psi}^{21} \circ (\psi^{21})^{-1}$ then ψ^{22} can be constructed as $\psi^{22}(\bar{x}_e^{21}) = \bar{\psi}^{22} \circ \pi(\bar{x}_e^{21})$. Therefore we can pick any ψ^{21} without loss of generality.

2.2.3 Proof of Main Result: Theorem 2.2.1

We assume the ith subsystem of (2.1) to be in Observable Form (2.3) and that its first $i-1$ subsystems are in BTF. We determine the existence conditions to transform the ith ($i \leqslant p-1$) subsystem (2.3) into BTF (2.4) while leaving the first $i-1$ subsystems unchanged. As originally suggested in [134], we treat the y_j, $i+1 \leqslant j \leqslant p$ which appear in (2.3) as inputs so that the problem of transforming the ith subsystem to BTF is converted into the problem of removing derivatives of inputs as considered in [41]. In particular, we wish to *completely* eliminate, as opposed to merely lowering input derivative order. The important point is that we consider *the first i subsystems*, and not the ith subsystem in isolation, as the original system. The problem is to transform the prolonged vector field (2.5a) into

$$\bar{F}_e^i(\bar{x}^{<i>}, y_{[i+1,p]}^{<\alpha^i>}) = \sum_{j=1}^{i} \bar{f}^j(\bar{x}^{<j>}, y_{[j+1,p]}) + \sum_{j=i+1}^{p} \sum_{k=0}^{\alpha_j^i} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}, \tag{2.10}$$

where we have removed the derivatives of $y_{[i+1,p]}$ from the ith subsystem dynamics.
Necessity: We assume there exists a transformation

$$\Psi = ((\bar{x}^{<i-1>})^T, (\Psi^i)^T, (y_{[i+1,p]}^{<\alpha^i>})^T)^T \tag{2.11}$$

putting the ith subsystem into BTF and leaving the other subsystems unchanged. According to Remark 2.2.3, the components of the inverse transformation $x^i_{k+1} = (\Psi^{-1})^i_{k+1}$ have the dependence

$$x^i_{k+1} = (\Psi^{-1})^i_{k+1}(\bar{x}^{<i>}, y_{i+1}^{<\alpha^i_{i+1}-\lambda_i+k>}, \dots, y_p^{<\alpha^i_p-\lambda_i+k>}),$$

for $0 \leqslant k \leqslant \lambda_i - 1$ and $\lambda_i \geqslant \max_{i+1 \leqslant j \leqslant p} \lambda_j - 1 \geqslant \max_{i+1 \leqslant j \leqslant p} \alpha^i_j$, and particularly we have

$$y_i = (\Psi^{-1})^i_1(\bar{x}^{<i>}, y_{[i+1,p]}).$$

Hence, the vector field (2.10) can be rewritten as

$$\bar{F}^i_e(\bar{x}^{<i>}, y^{<\alpha^i>}_{[i+1,p]}) = \sum_{j=1}^{i} \bar{f}^j(\bar{x}^{<i>}, y_{[i+1,p]}) + \sum_{j=i+1}^{p} \sum_{k=0}^{\alpha^i_j} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}.$$

By induction we calculate

$$\mathrm{ad}^k_{-\bar{F}^i_e} \frac{\partial}{\partial y_j^{(\alpha^i_j)}} = \frac{\partial}{\partial y_j^{(\alpha^i_j - k)}}, \quad 0 \leqslant k \leqslant \alpha^i_j; i+1 \leqslant j \leqslant p.$$

These unit vector fields satisfy the Lie bracket conditions (2.6), and since these conditions hold independent of coordinates, necessity is shown.

Sufficiency: We use a constructive method to prove sufficiency. The extended state transformation $\Psi(x^i_e)$ is the inverse of a composition of flows of vector fields. These vector fields are defined on the extended N-dimensional state space $\mathcal{S}$ which is constructed by extending the ν_i-dimensional state space of the first i subsystems with $y^{<\alpha^i>}_{[i+1,p]}$, the L output time derivatives in φ^i. We begin by introducing N vector fields defined on $\mathcal{S}$

$$\begin{aligned} X^j_k &= \frac{\partial}{\partial \bar{x}^j_k}, && 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \\ Y_k &= \frac{\partial}{\partial x^i_k}, && 1 \leqslant k \leqslant \lambda_i, \\ Z^j_k &= \mathrm{ad}^{\alpha^i_j - k}_{-F^i_e} \frac{\partial}{\partial y_j^{(\alpha^i_j)}}, && 0 \leqslant k \leqslant \alpha^i_j; i+1 \leqslant j \leqslant p. \end{aligned} \tag{2.12}$$

We denote the vector fields (2.12) as $W_1, \dots, W_N$ and order them as follows: the first ν_{i-1} vector fields are $X^1_1, \dots, X^1_{\lambda_1}, \dots, X^{i-1}_1, \dots, X^{i-1}_{\lambda_{i-1}}$, followed by the next λ_i vector fields $Y_1, \dots, Y_{\lambda_i}$, and finally followed by the L vector fields $Z^{i+1}_0, \dots, Z^{i+1}_{\alpha^i_{i+1}}, \dots,$

$Z_0^p, \ldots, Z_{\alpha_p^i}^p$. Given an initial condition $s_0 \in \mathbb{R}^N$, Ψ is given by $\Psi(x_e^i) = \Phi^{-1}(x_e^i) + s_0$ where

$$\Phi(t_1, \ldots, t_{N-1}, t_N) = \phi_{W_N}^{t_N}(\cdots(\phi_{W_2}^{t_2}(\phi_{W_1}^{t_1}(s_0)))\cdots) \tag{2.13}$$

with ϕ_V^t denoting the flow of a vector field V defined on $\mathcal{S}$, i.e., the solution of the differential equation

$$\frac{d}{dt}\phi_V^t(s) = V(\phi_V^t(s))$$

passing through s at $t = 0$.

We show $\Psi(x_e^i)$ has the following properties

(i) it is a local diffeomorphism;

(ii) it rectifies the prolonged vector fields $\mathrm{ad}_{-F_e^i}^k \frac{\partial}{\partial y_j^{(\alpha_j^i)}}$ into $\frac{\partial}{\partial y_j^{(\alpha_j^i)}}$;

(iii) it preserves $\bar{x}^{<i-1>}, y_{[i+1,p]}^{<\alpha^i>}$, i.e., Ψ takes the form of (2.11);

(iv) it transforms the ith subsystem into a BTF.

The property (i) can be shown by verifying that the Jacobian matrix of Φ at s_0 is nonsingular. Computing the ith column vector of the Jacobian matrix $\frac{\partial \Phi(t)}{\partial t}$

$$\begin{aligned}\frac{\partial \Phi(t)}{\partial t_i} =& \frac{\partial \Phi}{\partial \Upsilon_{N-1}} \cdots \frac{\partial \Upsilon_{N-i+2}}{\partial \Upsilon_{N-i+1}} \frac{\partial}{\partial t_{N-i+1}} \left(\phi_{W_{N-i+1}}^{t_{N-i+1}}(\ldots(\phi_{W_1}^{t_1}(s_0))\ldots)\right) \\ =& \frac{\partial \Phi}{\partial \Upsilon_{N-1}} \cdots \frac{\partial \Upsilon_{N-i+2}}{\partial w_{N-i+1}} W_{N-i+1}(\Upsilon_{N-i+1}(s_0)),\end{aligned}$$

where $t = (t_1, \ldots, t_N)^T$, and $\Upsilon_k = \phi_{W_k}^{t_k}(\ldots(\phi_{W_1}^{t_1}(s_0))\ldots), 1 \leqslant k \leqslant N-1$. Taking $t_k = 0$ for $1 \leqslant k \leqslant N$, we have

$$\begin{aligned}W_{N-i+1}(\Upsilon_{N-i+1}(s_0)) &= W_{N-i+1}(s_0), \\ \frac{\partial \Upsilon_{k+1}}{\partial \Upsilon_k} &= I_n.\end{aligned}$$

Since $\lambda_i \geqslant \max_{i+1 \leqslant j \leqslant p} \lambda_j - 1$, by direct calculation all the vector fields in (2.6) have the form

$$Z_k^j = \mathrm{ad}_{-F_e^i}^{\alpha_j^i - k} \frac{\partial}{\partial y_j^{(\alpha_j^i)}} = \sum_{s=1}^{\lambda_i} \sigma_s^{k,j}(x_e^i) \frac{\partial}{\partial x_s^i} + \frac{\partial}{\partial y_j^{(k)}}, \tag{2.14}$$

for $0 \leqslant k \leqslant \alpha_j^i, i+1 \leqslant j \leqslant p$, and where $\sigma_s^{k,j}$ is a smooth function. We therefore conculde that the Jacobian matrix

$$\frac{\partial \Phi}{\partial t}(0) = (W_1(s_0), \ldots, W_N(s_0)) = \begin{pmatrix} I_{\nu_i} & * \\ 0 & I_L \end{pmatrix}$$

is nonsingular and Ψ is a local diffeomorphism.

The property (ii) can be shown by verifying

$$\frac{\partial \Psi}{\partial x_e^i}\left(W_{N-L+1},\ldots,W_N\right)=\begin{pmatrix}0\\ I_L\end{pmatrix}. \tag{2.15}$$

From the Lie bracket conditions (2.6), i.e., $W_k, N-L+1 \leqslant k \leqslant N$ commute, we know their flows commute [162]. This implies

$$\begin{aligned}\Phi(t) &= \phi_{W_N}^{t_N}(\cdots(\phi_{W_2}^{t_2}(\phi_{W_1}^{t_1}(s_0)))\cdots)\\ &= \phi_{W_k}^{t_k}(\phi_{W_N}^{t_N}(\ldots(\phi_{W_{k+1}}^{t_{k+1}}(\phi_{W_{k-1}}^{t_{k-1}}\cdots(\phi_{W_1}^{t_1}(s_0))\ldots))\ldots)),\end{aligned}$$

for $N-L+1 \leqslant k \leqslant N$. Similar to the proof of property (i), we can derive

$$\frac{\partial \Phi}{\partial t_k}(t)=W_k(\Phi(t)), \quad N-L+1 \leqslant k \leqslant N,$$

and

$$\frac{\partial \Phi}{\partial t}=\left(*,W_{N-L+1},\ldots,W_N\right).$$

Considering the formula of the extended state transformation $\Psi(x_e^i)=\Phi^{-1}(x_e^i)+s_0$, we have

$$\frac{\partial \Psi}{\partial x_e^i}\frac{\partial \Phi}{\partial t}=\frac{\partial \Psi}{\partial x_e^i}\left(*,W_{N-L+1},\ldots,W_N\right)=I_N,$$

and complete the proof of (2.15).

To show the property (iii), we compute the effect of Φ on a point

$$(\bar{x}^{<i-1>},x^i,y_{[i+1,p]}^{<\alpha^i>})\in\mathcal{S}.$$

Clearly, for $Y_k, 1 \leqslant k \leqslant \lambda_i$

$$\phi_{Y_k}^{t_l}(x_e^i)=((\bar{x}^{<i-1>})^T,x_1^i,\ldots,\underbrace{x_k^i+t_l}_{\substack{k+\nu_{i-1}\\ \text{entry}}},\ldots,x_{\lambda_i}^i,(y_{[i+1,p]}^{<\alpha^i>})^T)^T. \tag{2.16}$$

We have a similar result for the flow of $X_k^j, 1 \leqslant k \leqslant \lambda_j, 1 \leqslant j \leqslant i-1$

$$\begin{aligned}\phi_{X_k^j}^{t_l}(x_e^i)=((\bar{x}^{<j-1>})^T,\ldots,\underbrace{\bar{x}_k^j+t_l}_{\substack{\nu_{j-1}+k\\ \text{entry}}},\ldots,(\bar{x}^{j+1})^T,\\ \ldots,(\bar{x}^{i-1})^T,(x^i)^T,(y_{[i+1,p]}^{<\alpha^i>})^T)^T.\end{aligned} \tag{2.17}$$

Considering (2.14), we have

$$\begin{aligned}\phi^{t_l}_{Z^j_k}(x^i_e) =&((\bar{x}^{<i-1>})^T, (\tilde{x}^i)^T, (y^{<\alpha^i>}_{[i+1,j-1]})^T, (y_j^{<k-1>})^T, \\ &y_j^{(k)} + t_l, (y_j^{[(k+1),(\alpha^i_j)]})^T, (y^{<\alpha^i>}_{[j+1,p]})^T)^T, \end{aligned} \tag{2.18}$$

where $\tilde{x}^i(x^i_e)$ is some smooth function, $l = \nu_i + \sum_{r=i+1}^{j-1}(\alpha^i_r + 1) + k + 1$, and $y_j^{[(k+1),(\alpha^i_j)]} = (y_j^{(k+1)}, \dots, y_j^{(\alpha^i_j)})^T$. From (2.16)–(2.18) we conclude that Φ defined in (2.13) transforms a point

$$s_0 = (\bar{x}_0^{<i-1>}, x^i_0, (y^{<\alpha^i>}_{[i+1,p]})_0) \in \mathcal{S}$$

into

$$(\bar{\bar{x}}^{<i-1>}, \tilde{x}^i(x^i_e), (\tilde{y}^{<\alpha^i>}_{[i+1,p]})) \in \mathcal{S},$$

where

$$\begin{aligned}\bar{\bar{x}}^j_k &= t_{\nu_{j-1}+k} + (\bar{x}^j_k)_0, \quad 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \\ \tilde{y}^{(k)}_j &= t_{\nu_i + \sum_{r=i+1}^{j-1}(\alpha^i_r+1)+k+1} + (y_j^{(k)})_0, \quad 0 \leqslant k \leqslant \alpha^i_j; i+1 \leqslant j \leqslant p\end{aligned}$$

with $(\cdot)_0$ denoting the initial point, and $\tilde{x}^i(x^i_e) \in \mathbb{R}^{\lambda_i}$ is a smooth function. Thus Ψ, which differs from Φ^{-1} by a shift of the initial condition s_0, preserves $\bar{x}^j_k, 1 \leqslant k \leqslant \lambda_j, 1 \leqslant j \leqslant i-1$ and $y_j^{(k)}, 0 \leqslant k \leqslant \alpha^i_j, i+1 \leqslant j \leqslant p$. Therefore, only x^i is transformed or Ψ has the special structure given by (2.11).

Finally we show Ψ transforms the ith subsystem into BTF, i.e., the property (iv). Since $\mathrm{ad}^k_{-F^i_e}\frac{\partial}{\partial y_j^{(\alpha^i_j)}}, 0 \leqslant k \leqslant \alpha^i_j, i+1 \leqslant j \leqslant p$ are rectified into $\frac{\partial}{\partial y_j^{(\alpha^i_j-k)}}$ by Ψ, i.e., in new coordinates,

$$\mathrm{ad}^k_{-\bar{F}^i_e}\frac{\partial}{\partial y_j^{(\alpha^i_j)}} = \frac{\partial}{\partial y_j^{(\alpha^i_j-k)}}, \tag{2.19}$$

where $\bar{F}^i_e$ denotes F^i_e in the new coordinates. Since $\lambda_k \geqslant \max_{k+1 \leqslant j \leqslant p}\lambda_j - 1 \geqslant \max_{k+1 \leqslant j \leqslant p}\alpha^k_j$, we know y_k only depends on $(\bar{x}^{<k>}, y_{[k+1,p]})$ for $1 \leqslant k \leqslant i$. After applying the extended state transformation Ψ, $\bar{F}^i_e$ is written as

$$\bar{F}^i_e = \sum_{k=1}^{i-1}\bar{f}^k(\bar{x}^{<i>}, y_{[i+1,p]}) + \bar{f}^i + \sum_{j=i+1}^{p}\sum_{k=0}^{\alpha^i_j} y_j^{(k+1)}\frac{\partial}{\partial y_j^{(k)}}.$$

From (2.19), we readily verify by induction that

$$\frac{\partial \bar{f}^i}{\partial y_j^{(\alpha^i_j-k)}} = 0, \quad 0 \leqslant k \leqslant \alpha^i_j - 1; i+1 \leqslant j \leqslant p,$$

which implies that $\bar{f}^i$ is independent of $y_j^{(k)}, 1 \leqslant k \leqslant \alpha_j^i, i+1 \leqslant j \leqslant p$, and the ith subsystem is in BTF. ■

Remark 2.2.6. According to the structure of the vector fields in (2.14), we have

$$\mathcal{Z} = \left(Z_0^{i+1}, \dots, Z_{\alpha_{i+1}^i}^{i+1}, \dots, Z_0^p, \dots, Z_{\alpha_p^i}^p \right) = \begin{pmatrix} 0_{\nu_{i-1} \times L} \\ \Omega_{(\lambda_i + L) \times L} \end{pmatrix}.$$

Thus Ψ^i can be solved from

$$\frac{\partial \Psi^i}{\partial x_e^i} \Omega = 0_{\lambda_i \times L}, \tag{2.20}$$

where x_e^i is defined in (2.8b). Written in the form (2.20) we remark the similarity with the PDEs provided in [139, Thm. 1]. However, the first λ_i rows of Ω can differ. This accounts for PDEs in [139, Thm. 1] leading to the incorrect transformations for some systems. For the same reason, the Lie bracket conditions (2.6) using (2.8) can lead to the wrong conclusion about the existence of a BTF. □

Remark 2.2.7. In a more general setting, for $\mathcal{Z}$ to be rectified into the form of $(0, 0, I)^T$ by an extended state transformation Ψ (2.11), the first ν_{i-1} rows of $\mathcal{Z}$ must be zero. In Theorem 2.2.1, this necessary condition is inferred from the ordering constraint on λ_i. □

Example 2.2. To illustrate the construction of the BTF coordinates by the composition of flows of vector fields, we consider a system

$$\begin{aligned} \dot{x}^1 &= f^1 = \begin{pmatrix} x_2^1 \\ c \end{pmatrix}, \\ \dot{x}^2 &= f^2 = \begin{pmatrix} x_2^2 \\ x_2^1 \ddot{y}_3 \end{pmatrix}, \\ \dot{x}^3 &= f^3 = \begin{pmatrix} x_2^3 \\ x_3^3 \\ \varphi^3(x) \end{pmatrix}, \\ y &= \begin{pmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \end{pmatrix}, \end{aligned}$$

where c is a constant and φ^3 a C^∞ function. The system is in Observable Form with indices $(2, 2, 3)$. The first and the third subsystems are already in BTF and we only consider the transformation of the second subsystem. We have $\alpha_3^2 = 2$ since $\ddot{y}_3$ appears in f^2. Defining the extended state x_e^2 and the prolonged vector field F_e^2

$$\begin{aligned} x_e^2 &= \left((x^1)^T, (x^2)^T, (y_3^{<2>})^T \right)^T, \\ F_e^2 &= ((f^1)^T, (f^2)^T, \dot{y}_3, \ddot{y}_3, y_3^{(3)})^T, \end{aligned}$$

we verify the conditions of Theorem 2.2.1 with $j = l = 3$ and $i = 2$. Taking the vector fields

$$\begin{aligned}
&W_1 = \frac{\partial}{\partial x_1^1}, \quad W_2 = \frac{\partial}{\partial x_2^1}, \quad W_3 = \frac{\partial}{\partial x_1^2}, \quad W_4 = \frac{\partial}{\partial x_2^2},\\
&W_5 = \mathrm{ad}^2_{-F_e^2}\frac{\partial}{\partial \ddot{y}_3} = x_2^1\frac{\partial}{\partial x_2^1} - c\frac{\partial}{\partial x_2^2} + \frac{\partial}{\partial y_3},\\
&W_6 = \mathrm{ad}_{-F_e^2}\frac{\partial}{\partial \ddot{y}_3} = x_2^1\frac{\partial}{\partial x_2^2} + \frac{\partial}{\partial \dot{y}_3},\\
&W_7 = \frac{\partial}{\partial \ddot{y}_3},
\end{aligned}$$

we compute the composition of flows of $W_k, 1 \leqslant k \leqslant 7$ with the initial condition $s_0 = 0$.

$$\phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1\\0\\0\\0\\0\\0\\0 \end{pmatrix},$$

$$\phi_{W_2}^{t_2}\circ\phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1\\t_2\\0\\0\\0\\0\\0 \end{pmatrix},$$

$$\phi_{W_3}^{t_3}\circ\phi_{W_2}^{t_2}\circ\phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1\\t_2\\t_3\\0\\0\\0\\0 \end{pmatrix},$$

$$\phi_{W_4}^{t_4}\circ\phi_{W_3}^{t_3}\circ\phi_{W_2}^{t_2}\circ\phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1\\t_2\\t_3\\t_4\\0\\0\\0 \end{pmatrix},$$

$$\phi_{W_5}^{t_5} \circ \phi_{W_4}^{t_4} \circ \phi_{W_3}^{t_3} \circ \phi_{W_2}^{t_2} \circ \phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1 \\ t_2 \\ t_3 + t_2 t_5 \\ t_4 - c t_5 \\ t_5 \\ 0 \\ 0 \end{pmatrix},$$

$$\phi_{W_6}^{t_6} \circ \phi_{W_5}^{t_5} \circ \phi_{W_4}^{t_4} \circ \phi_{W_3}^{t_3} \circ \phi_{W_2}^{t_2} \circ \phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1 \\ t_2 \\ t_3 + t_2 t_5 \\ t_4 - c t_5 + t_2 t_6 \\ t_5 \\ t_6 \\ 0 \end{pmatrix},$$

$$x = \phi_{W_7}^{t_7} \circ \phi_{W_6}^{t_6} \circ \phi_{W_5}^{t_5} \circ \phi_{W_4}^{t_4} \circ \phi_{W_3}^{t_3} \circ \phi_{W_2}^{t_2} \circ \phi_{W_1}^{t_1}(0) = \begin{pmatrix} t_1 \\ t_2 \\ t_3 + t_2 t_5 \\ t_4 - c t_5 + t_2 t_6 \\ t_5 \\ t_6 \\ t_7 \end{pmatrix}.$$

The transformation Ψ is obtained by solving the inverse of the composition of flows

$$\Psi = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \end{pmatrix} = \begin{pmatrix} x_1^1 \\ x_2^1 \\ x_1^2 - x_2^1 x_1^3 \\ x_2^2 - x_2^1 x_2^3 + c x_1^3 \\ x_1^3 \\ x_2^3 \\ x_3^3 \end{pmatrix},$$

which define the BTF coordinates

$$\bar{x}^2 = \begin{pmatrix} x_1^2 - x_2^1 x_1^3 \\ x_2^2 - x_2^1 x_2^3 + c x_1^3 \end{pmatrix}.$$

□

2.2.4 BTF with Linear Output Function

As we will discuss in the next chapter, a BTF is an intermediate step between the original ζ-coordinates and Block Triangular Observer Form (BTOF) coordinates. For the ith subsystem in BTF to be transformable to BTOF, the ith output is required to be independent of lower subsystem states and outputs. Furthermore, a BTF system with a linear output is more likely to admit a BTOF. Hence, the existence conditions of a BTF with linear output function are significant. Using the same procedure as in the sufficiency side of the proof of Theorem 2.2.1, we obtain the following theorem which ensures $y_j = \bar{x}_1^j, 1 \leqslant j \leqslant p$.

Theorem 2.2.8. *Given an observable system* (2.1) *whose first* $i-1$ *subsystems are in BTF, the ith subsystem* $1 \leqslant i \leqslant p-1$ *in* (2.3) *can be locally transformed into BTF* (2.4) *by* $\Psi^i(x_e^i) = (x_1^i, \Psi_2^i, \dots, \Psi_{\lambda_i}^i)^T$ *leaving the remaining subsystems unchanged if* $\alpha_j^i + 1 \leqslant \lambda_i, i+1 \leqslant j \leqslant p$, *and the Lie bracket conditions* (2.6) *hold.*

Proof: Since $\alpha_j^i + 1 \leqslant \lambda_i$, the vector fields Z_k^j can be written as

$$Z_k^j = \begin{cases} \frac{\partial}{\partial y_j^{(k)}} + \sum_{r=0}^{\alpha_j^i - k} \sigma_r^{k,j}(x_e^i) \frac{\partial}{\partial x_{\lambda_i - r}^i}, & 0 \leqslant k \leqslant \alpha_j^i - 1, \\ \frac{\partial}{\partial y_j^{(\alpha_j^i)}}, \quad k = \alpha_j^i, \end{cases}$$

for smooth functions $\sigma_r^{k,j}(x_e^i)$, i.e., Z_k^j has no $\frac{\partial}{\partial x_1^i}$ component. Therefore, the composition of flows defined in (2.13) preserves x_1^i and we have $\bar{x}_1^i = x_1^i$. Finally, $\Psi^i = (x_1^i, \Psi_2^i, \dots, \Psi_{\lambda_i}^i)^T$ transforms the ith subsystem into BTF. □

Remark 2.2.9. If system (2.1) has indices $\lambda_1 \geqslant \ldots \geqslant \lambda_p$ and is transformable to BTF, we have $y_j = \bar{x}_1^j, 1 \leqslant j \leqslant p$. □

2.2.5 BTF for Systems with Inputs

We now consider the BTF existence conditions for a class of affine control nonlinear systems. Without loss of generality, we assume the input u is a scalar,

$$\begin{aligned} \dot{\zeta} &= f(\zeta) + \varrho(\zeta)u, \\ y &= h(\zeta), \end{aligned} \tag{2.21}$$

where $\varrho(\zeta)$ is an input vector field. It is well-known that for nonlinear systems observability is a function of input [46]. As our objective is observer design we avoid the problem of "bad inputs" by assuming (2.21) is uniformly observable for any input. By defining the observable coordinates $x_k^i = L_f^{k-1} h_i, 0 \leqslant k \leqslant \lambda_i - 1, 1 \leqslant i \leqslant p$, the ith subsystem of system (2.21) is rewritten as

$$\begin{aligned} \dot{x}^i &= \begin{pmatrix} x_2^i \\ x_3^i \\ \vdots \\ \varphi^i(x) \end{pmatrix} + \begin{pmatrix} L_{\varrho(x)} h_i(x) \\ L_{\varrho(x)} L_{f(x)} h_i(x) \\ \vdots \\ L_{\varrho(x)} L_{f(x)}^{\lambda_i - 1} h_i(x) \end{pmatrix} u = f^i(x) + \varrho^i(x) u, \\ y_i &= h_i(x), \end{aligned} \tag{2.22}$$

where $f(x), \varrho(x)$ are expression of $f(\zeta), \varrho(\zeta)$ in x-coordinates respectively. Denote $f^i_j(x), \varrho^i_j(x)$ as the jth component of $f^i(x), \varrho^i(x)$ respectively. Sufficient conditions to ensure (2.22) is uniformly observable can be found in [22, 23].

A system (2.21) is in BTF if its ith subsystem has the form

$$\begin{aligned} \dot{\bar{x}}^i &= \bar{f}^i(\bar{x}^{<i>}, y_{[i+1,p]}) + \bar{\varrho}^i(\bar{x}^{<i>}, y_{[i+1,p]})u, \\ y_i &= \bar{h}_i(\bar{x}). \end{aligned} \tag{2.23}$$

We state the BTF existence conditions for the forced case in the following theorem.

Theorem 2.2.10. *Given a uniformly observable system* (2.21) *with*

$$\lambda_i \geqslant \max_{i+1 \leqslant j \leqslant p} \lambda_j - 1$$

, with its first $i-1$ subsystems in BTF (2.23)*, and its last $p-i+1$ subsystems in* (2.22)*, define the prolonged vector field F^i_e and the extended state x^i_e as* (2.5)*. The ith subsystem can be locally transformed into BTF in* (2.23) *by the extended state transformation $\bar{x}^i = \Psi^i(x^i_e)$ and the first $i-1$ subsystems remain identical if and only if in addition to Condition* (2.6)*, the following Lie bracket conditions hold*

$$\left[\mathrm{ad}^q_{F^i_e} \frac{\partial}{\partial y_j^{(\alpha^i_j)}}, \varrho^i_e \right] = 0, \quad 0 \leqslant q \leqslant \alpha^i_j - 1; i+1 \leqslant j \leqslant p,$$

where $\varrho^i_e = (0, (\varrho^i(x^i_e))^T, 0_{1\times L})^T$. The extended state transformation Ψ^i is a solution of the $\lambda_i \cdot L$ PDEs (2.7).

2.3 BTF Existence Conditions : Parameter Approach

Theorem 2.2.1 gives BTF existence conditions for an observable system (2.1) provided that its indices satisfy

$$\lambda_i \geqslant \max_{i+1 \leqslant j \leqslant p} \lambda_j - 1, \quad 1 \leqslant i \leqslant p.$$

This condition potentially restricts the choice of indices, and the objective of this section is therefore to remove this constraint.

Noticing $\alpha^i_j, i+1 \leqslant j \leqslant p$ instead of λ_j plays a crucial role in determining the expressions of the prolonged vector fields Z^j_k and therefore the extended state transformation to a BTF, which in turn affects the BTF existence conditions. We divide the existence problem of a BTF into two cases in terms of α^i_j.

Case 1: $\lambda_i \geqslant \alpha_j^i, \quad i+1 \leqslant j \leqslant p.$

Case 2: $\lambda_i < \alpha_j^i, \quad i+1 \leqslant j \leqslant p.$

Since **Case 1** differs slightly from conditions in Theorem 2.2.8 for BTF with linear output, we state the BTF existence conditions for **Case 1** in the following proposition without proof.

Proposition 2.3.1. *Given an observable system* (2.1) *whose first* $i-1$ *subsystems are in BTF, the ith subsystem* $1 \leqslant i \leqslant p-1$ *in* (2.3) *can be locally transformed into BTF* (2.4) *by* $\Psi^i(x_e^i)$ *leaving the remaining subsystems unchanged if* $\alpha_j^i \leqslant \lambda_i, i+1 \leqslant j \leqslant p$*, and the Lie bracket conditions in Theorem 2.2.1 hold.*

Next, we focus on **Case 2**. As shown by the following example, allowing $\alpha_j^k > \lambda_k$ makes y_k a general function of $\bar{x}$ in BTF coordinates. For the ith subsystem, the complexity of the state approach increases drastically because of the dependence on $y_{[2,i-1]}$ in the first $i-1$ subsystems. Hence, we treat $y_{[2,i-1]}$ as *parameters*. Since the ordering conditions on observability indices are not satisfied, Theorem 2.2.1 and Remark 2.2.4 are in general not applicable. However, with $y_{[2,i-1]}$ treated as parameters, we verify that the first ν_{i-1} components of the vector fields

$$\mathrm{ad}^k_{-F_e^i} \frac{\partial}{\partial y_j^{(\alpha_j^i)}}, \quad 0 \leqslant k \leqslant \lambda_i; i+1 \leqslant j \leqslant p$$

are zero. This fact allows us to establish the following proposition by imitating the proof of Theorem 2.2.1.

Proposition 2.3.2. *Assume* $y_{[2,i-1]}$ *are parameters and* $\alpha_j^i > \lambda_i$*. For the ith subsystem, the output derivatives* $y_j^{(\alpha_j^i - k)}$, $0 \leqslant k \leqslant \lambda_i$*, of the jth subsystem can be removed by an extended state transformation* Ψ^i *if and only if the Lie bracket conditions of Theorem 2.2.1 are satisfied with the indices* $0 \leqslant q, r \leqslant \lambda_i, i+1 \leqslant j, l \leqslant p$ *and the vector fields* $\mathrm{ad}^k_{-F_e^i} \frac{\partial}{\partial y_j^{(\alpha_j^i)}}, 0 \leqslant k \leqslant \lambda_i, i+1 \leqslant j \leqslant p$ *are independent of* $y_{[2,i-1]}$.

Example 2.3. We consider a five-output observable system with indices $(2,1,1,1,4)$

$$\begin{aligned}
\dot{\bar{x}}^1 &= \bar{f}^1(\bar{x}^1, y_{[2,4]}),\\
\dot{\bar{x}}_1^2 &= \bar{f}^2 = \bar{x}_1^2 + y_3^2,\\
\dot{x}_1^3 &= f^3 = x_1^3 + y_5^{(3)} - \ddot{y}_5 + \bar{x}_1^2,\\
\dot{x}_1^4 &= f^4 = x_1^4 + \bar{x}_1^2\dot{y}_5,\\
\dot{x}^5 &= f^5 = \begin{pmatrix} x_2^5 \\ x_3^5 \\ x_4^5 \\ \varphi^5(x) \end{pmatrix},\\
y &= \begin{pmatrix} \bar{x}_1^1 \\ \bar{x}_1^2 \\ x_1^3 \\ x_1^4 \\ x_1^5 \end{pmatrix}.
\end{aligned}$$

The first two and the last subsystems are in BTF. For the third subsystem, since $\lambda_3 = 1$, Proposition 2.3.2 is applied to determine the existence of a change of coordinates where the second subsystem dynamics is independent of $y_5^{(3)}$. We verify the Lie bracket condition (2.6) with $0 \leqslant q, r \leqslant 1, l = j = 5$, and solve PDEs (2.7) for Ψ^3

$$\bar{x}_1^3 = \Psi^3 = x_1^3 - \ddot{y}_5,$$

which puts the third subsystem into

$$\dot{\bar{x}}_1^3 = \bar{f}^3 = \bar{x}_1^3 + \bar{x}_1^2.$$

For the fourth subsystem, $\alpha_5^4 = 1 = \lambda_4$. If y_3 appearing in $\bar{f}^2$ is treated as a function of state, Proposition 2.3.2 is not applicable to remove $\dot{y}_5$ from the fourth subsystem since $\frac{\partial}{\partial \dot{y}_5}, \mathrm{ad}_{-F_e^4}\frac{\partial}{\partial \dot{y}_5}$ are not the vector fields to be rectified. We illustrate this fact in Example 2.4. On the other hand, treating y_3 as parameter, we apply Proposition 2.3.2 to solve an extended state transformation

$$\Psi^4 = x_1^4 - \bar{x}_1^2 y_4.$$

Ψ^4 puts the fourth subsystem into

$$\dot{\bar{x}}_1^4 = \bar{x}_1^4 + (\bar{x}_1^2)^2 y_4 + (\bar{x}_1^2 + y_3^2) y_4, \tag{2.24}$$

which is independent of output derivatives of the lower subsystem. Note that due to the dependence on y_3 (2.24) is not in the BTF (2.4). □

Example 2.3 demonstrates the use of Proposition 2.3.2 for **Case 2** and the limitation of the BTF (2.4). In the sequel, we consider the existence conditions of a generalized BTF whose ith subsystem is written as

$$\begin{aligned} \dot{\bar{x}}^i &= \bar{f}^i(\bar{x}^{<i>}, y), \\ y_i &= \bar{h}_i(\bar{x}). \end{aligned} \tag{2.25}$$

We can make use of Proposition 2.3.2 to solve the extended state transformation Ψ^i which eliminates the output derivatives $y_j^{(\alpha_j^i - k)}, 0 \leqslant k \leqslant \lambda_i$ from the ith subsystem. The ith subsystem after Ψ^i will depend on $y_j^{<\beta_j^i>}, \beta_j^i = \alpha_j^i - \lambda_i$. Hence, in this section we consider the following problem.

Problem: Given an observable system with the first $i-1$ subsystems in BTF

$$\begin{aligned} \dot{\bar{x}}^k &= \bar{f}^k(\bar{x}^{<k>}, y), \\ y_k &= \bar{h}_k(\bar{x}^{<k>}, y_{[k+1,p]}^{<\beta^k>}), \end{aligned} \quad 1 \leqslant k \leqslant i-1, \tag{2.26}$$

the ith subsystem

$$\begin{aligned} \dot{x}^i &= f^i(\bar{x}^{<i-1>}, x^i, y_{[i+1,p]}^{<\beta^i>}, y), \\ y_i &= h_i(\bar{x}^{<i-1>}, x^i, y_{[i+1,p]}^{<\beta^i>}), \end{aligned} \tag{2.27}$$

and the remaining $p-i$ subsystems are in Observable Form, find the necessary and sufficient conditions to guarantee the existence of an extended state transformation

$$\Psi(\bar{x}^{<i-1>}, x^i, y_{[i+1,p]}^{<\beta^i-1>}) = \begin{pmatrix} \bar{x}^{<i-1>} \\ \Psi^i(\bar{x}^{<i-1>}, x^i, y_{[i+1,p]}^{<\beta^i-1>}) \\ y_{[i+1,p]}^{<\beta^i>} \end{pmatrix} \tag{2.28}$$

such that the ith subsystem is transformed into a BTF (2.25). □

Remark 2.3.3. The ith subsystem (2.27) is not in Observable Form and y_i is a general function of system state. This is different from what is considered in Section 2.2. Proposition 2.3.2 is not applicable to remove $y_{[i+1,p]}^{<\beta^i>}$ from f^i. This is explained by Example 2.4. □

Example 2.4. We consider a three-output observable system with indices $(1,1,3)$

$$\begin{aligned} \dot{x}_1^1 &= f^1 = x_1^1 + y_2, \\ \dot{x}_1^2 &= f^2 = x_1^1 + 2x_1^2 + \ddot{y}_3 + \dot{y}_3, \\ \dot{x}^3 &= f^3 = \begin{pmatrix} x_2^3 \\ x_3^3 \\ \varphi^3(x) \end{pmatrix}, \\ y &= \begin{pmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \end{pmatrix}, \end{aligned} \tag{2.29}$$

where the first and the third subsystems are in BTF. Proposition 2.3.2 is applied to removed $\ddot{y}_3$ from the second subsystem. We verify the Lie bracket condition (2.6) for $q=0, r=1, j=l=3$, and solve the extended state transformation

$$\Psi^{21} = x_1^2 - \dot{y}_3,$$

which yields the first two subsystems

$$\begin{aligned}\dot{x}_1^1 &= x_1^1 + \bar{x}_1^{21} + \dot{y}_3,\\ \dot{\bar{x}}_1^{21} &= \bar{f}^{21} = 3\dot{y}_3 + x_1^1 + 2\bar{x}_1^{21}.\end{aligned}$$

We attempt to remove $\dot{y}_3$ by reapplying Proposition 2.3.2. Defining the extended state and prolonged vector field

$$\begin{aligned}x_e^{21} &= (x_1^1, \bar{x}_1^{21}, y_3, \dot{y}_3)^T,\\ F_e^{21} &= (f^1, \bar{f}^{21}, \dot{y}_3, \ddot{y}_3)^T,\end{aligned}$$

computing the vector field

$$\mathrm{ad}_{-F_e^{21}} \frac{\partial}{\partial \dot{y}_3} = \frac{\partial}{\partial x_1^1} + 3\frac{\partial}{\partial \bar{x}_1^{21}} + \frac{\partial}{\partial y_3},$$

and verifying the Lie bracket condition (2.6) for $q=0, r=1, j=l=3$, we solve the extended state transformation to remove $\dot{y}_3$ from the second subsystem

$$\Psi = \begin{pmatrix}\bar{x}_1^1\\ \bar{x}_1^2\end{pmatrix} = \begin{pmatrix}x_1^1 + \bar{x}_1^{21}\\ -2x_1^1 - \bar{x}_1^{21}\end{pmatrix}.$$

We remark that the change of coordinates Ψ cannot preserve the first subsystem dynamics since the first component of the vector field $\mathrm{ad}_{-F_e^{21}} \frac{\partial}{\partial \dot{y}_3}$ is nonzero. In the new coordinates, the first two subsystems are expressed as

$$\begin{aligned}\dot{\bar{x}}_1^1 &= \bar{f}^1 = \bar{x}_1^1 + y_3 + y_2 - \dot{y}_3,\\ \dot{\bar{x}}_1^2 &= \bar{f}^2 = -5\bar{x}_1^1 - \bar{x}_1^2 - 5y_3,\end{aligned}$$

which are not in BTF (2.25) since the first subsystem dynamics depends on $\dot{y}_3$. On the other hand, the first two subsystems are in BTF with the transformation

$$\bar{\Psi} = \begin{pmatrix}\bar{x}_1^1\\ \bar{x}_1^2\end{pmatrix} = \begin{pmatrix}x_1^1\\ \bar{x}_1^{21} - 3y_3\end{pmatrix}.$$

This fact demonstrates that for **Case 2** vector fields which appear in the Lie bracket conditions (2.6) are not necessarily unit vectors to be rectified. This results from the appearance of y_i in the upper subsystem dynamics. □

2.3.1 Main Result

We first introduce several definitions, then state the main theorem, and finally present the proof of the theorem. For the ith subsystem, the prolonged vector field F_e^i and the extended state x_e^i are redefined as

$$F_e^i(x_e^i) = \sum_{j=1}^{i-1} \bar{f}^j(\bar{x}^{<j>}, y) + f^i(x_e^i) + \sum_{j=i+1}^{p} \sum_{k=0}^{\beta_j^i} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}, \tag{2.30a}$$

$$x_e^i = ((\bar{x}^{<i-1>})^T, (x^i)^T, (y_{[i+1,p]}^{<\beta^i>})^T)^T, \tag{2.30b}$$

and $N = \nu_i + L, L = \sum_{j=i+1}^{p}(\beta_j^i + 1)$. The vector fields $\eta_{j,k}^i, 0 \leqslant k \leqslant \beta_j^i, i+1 \leqslant j \leqslant p$ are defined as

$$\eta_{j,0}^i = \frac{\partial}{\partial y_j^{(\beta_j^i)}}, \quad \eta_{j,k}^i = \mathrm{ad}_{-F_e^i}\eta_{j,k-1}^i - \kappa_{j,k}^i, \tag{2.31}$$

where

$$\kappa_{j,0}^i = 0, \quad \kappa_{j,k}^i = \begin{pmatrix} \dfrac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta_j^i - k+1)}} \\ -\left(\dfrac{\partial \Psi^i}{\partial x^i}\right)^{-1} \dfrac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \dfrac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta_j^i - k+1)}} \\ 0 \end{pmatrix}. \tag{2.32}$$

The vector fields $\tilde{\eta}_{j,0}^i, \tilde{\eta}_{j,1}^i, i+1 \leqslant j \leqslant p$ are defined by the following equations

$$\tilde{\eta}_{j,0}^i = \eta_{j,0}^i \quad \tilde{\eta}_{j,1}^i = \frac{\partial F_e^i}{\partial y_j^{(\beta_j^i)}}. \tag{2.33}$$

We also introduce the following definition.

Definition 2.3.4. U is the set of all transformations Ψ^i such that $\frac{\partial \Psi^i}{\partial x^i}$ is nonsingular, and $\eta_{j,k}^i, 0 \leqslant k \leqslant \beta_j^i, i+1 \leqslant j \leqslant p$ are rectified into unit vector fields $\frac{\partial}{\partial y_j^{(\beta_j^i - k)}}$ respectively. □

Remark 2.3.5. According to the proof of Theorem 2.2.1, the set U is not empty if and only if there exists a state transformation Ψ^i such that $\eta_{j,k}^i$ are linearly independent and commute. It is not difficult to verify that vector fields $\eta_{j,k}^i$ (2.31) are always linearly independent. However, the construction algorithm cannot ensure that the first ν_{i-1} rows of $\eta_{j,k}^i$ are zero, which is required to construct an extended state transformation as (2.28), i.e., $(\eta_{i+1,\beta_{i+1}^i}^i, \ldots, \eta_{j,0}^i, \ldots, \eta_{p,\beta_p^i}^i, \ldots, \eta_{p,0}^i)$ is not of the form

$$\begin{pmatrix} 0_{\nu_{i-1} \times L} \\ * \\ I \end{pmatrix}.$$

As in Section 2.2, the extended state transformation rectifying $\eta^i_{j,k}$ is defined by the PDEs

$$\frac{\partial \Psi}{\partial x^i_e}\eta^i_{j,k} = \frac{\partial}{\partial y_j^{(\beta^i_j - k)}}, \quad 0 \leqslant k \leqslant \beta^i_j; i+1 \leqslant j \leqslant p,$$

and $\Psi^i \in U$ can be solved from the PDEs

$$\frac{\partial \Psi^i}{\partial x^i_e}\eta^i_{j,k} = 0, \quad 0 \leqslant k \leqslant \beta^i_j; i+1 \leqslant j \leqslant p. \tag{2.34}$$

As demonstrated in Example 2.4, the extended state transformation solved from the above PDEs is not necessarily written as (2.28). □

Theorem 2.3.6. *Considering an observable system* (2.1) *with the upper i subsystems given by* (2.26)-(2.27), *and the remaining $p-i$ subsystems in Observable Form, the ith subsystem $1 \leqslant i \leqslant p-1$* (2.27) *can be locally transformed into BTF in* (2.25) *by an extended state transformation* (2.28) *if and only if*

(i) U is non-empty;

(ii) the first ν_{i-1} rows of $\eta^i_{j,k}$ are zero;

(iii) the vector fields $\eta^i_{j,k}$ are independent of the parameters, i.e.,

$$\frac{\partial \eta^i_{j,k}}{\partial y_r} = 0, \quad \begin{cases} 1 \leqslant r \leqslant i-1; \\ 0 \leqslant k \leqslant \beta^i_j; \\ i+1 \leqslant j \leqslant p. \end{cases} \tag{2.35}$$

Remark 2.3.7. Treating $y_{[1,i-1]}$ appearing in the first i subsystems as parameters simplifies the presentation of the proof. Also, it ensures that $y_j^{(\beta^i_j)}$ is the highest derivative of $y_j, i+1 \leqslant j \leqslant p$ appearing in the prolonged vector field F^i_e and extended state definition (2.30b). On the contrary, treating $y_{[1,i-1]}$ as the state results in the appearance of $y_j^{(\beta^k_j)}, 1 \leqslant k \leqslant i-1, k+1 \leqslant j \leqslant p$ in F^i_e, where $y_{[1,i-1]}$ might have higher order time derivatives of lower subsystem outputs than y_i, i.e., $\beta^k_j > \beta^i_j$. The extended state has to include output time derivatives higher order than β^i_j. This will overcomplicate the derivation and statement of the existence conditions. As well, treating $y_{[1,i-1]}$ as a parameter does not significantly restrict the class of systems. The component y_k with $y_k = \bar{x}^k_1$ can be treated as a state, and accordingly the conditions (2.35) with $r = k$ are no longer necessary. The example in Section 2.3.4 illustrates this case. □

Remark 2.3.8. Condition (iii), where $\eta^i_{j,k}$ determine the dependence of Ψ, are imposed by the assumption that Ψ depends on x^i_e. The extended state transformation Ψ has no dependence on $y_r, 1 \leqslant r \leqslant i-1$ only if none of the vector fields $\eta^i_{j,k}$ depends on y_r. □

2.3.2 Proof of Main Result: Theorem 2.3.6

Necessity: If the first i subsystems are in BTF and the remaining subsystems are in Observable Form then there exists a transformation (2.28) to put the ith subsystem into a BTF. We show $\Psi^i \in U$, or equivalently the vector fields $\eta^i_{j,k}$ are linearly independent and commute. First we introduce the prolonged vector field

$$\bar{F}^i_e = \sum_{j=1}^{i} \bar{f}^j(\bar{x}^{<j>}, y) + \sum_{j=i+1}^{p} \sum_{k=0}^{\beta^i_j} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}, \tag{2.36}$$

where the dependence of ith subsystem dynamics on derivatives of $y_{[i+1,p]}$ has been removed. As discussed in Section 2.2, y_i is some function of $\bar{x}^{<i>}, y^{<\beta^i>}_{[i+1,p]}$. Consequently, (2.36) is rewritten as

$$\begin{aligned}\bar{F}^i_e = &\sum_{j=1}^{i-1} \bar{f}^j(\bar{x}^{<i>}, y^{<\beta^i>}_{[i+1,p]}, y_{[j+1,i-1]}, y_{[i+1,p]}) \\ &+ \bar{f}^i(\bar{x}^{<i>}, y_{[2,i-1]}, y_{[i+1,p]}) + \sum_{j=i+1}^{p} \sum_{k=0}^{\beta^i_j} y_j^{(k+1)} \frac{\partial}{\partial y_j^{(k)}}.\end{aligned}$$

Taking $\bar{\eta}^i_{j,k} = \frac{\partial}{\partial y_j^{(\beta^i_j - k)}}, 0 \leqslant k \leqslant \beta^i_j, i+1 \leqslant j \leqslant p$, which are linearly independent unit vectors and commute, we have shown the necessity of Condition (i) in BTF coordinates. Condition (ii)–(iii) hold in BTF coordinates. Given the transformation (2.28), we have

$$\Psi_* = \frac{\partial \Psi}{\partial x^i_e} = \begin{pmatrix} I & 0 & 0 \\ \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} & \frac{\partial \Psi^i}{\partial x^i} & * \\ 0 & 0 & I \end{pmatrix}.$$

With $\eta^i_{j,k} = (\Psi_*)^{-1}\bar{\eta}^i_{j,k}$, where $\eta^i_{j,k}$ denotes $\bar{\eta}^i_{j,k}$ in x^i_e-coordinates, we prove that Conditions (ii)–(iii) are necessary. To show Condition (i) is necessary in x^i_e-coordinates, we need to work out the expression of $\eta^i_{j,k}$. According to the definition of the transformation (2.28), it is trivial to determine $\eta^i_{j,0}$ are unit vectors as well. Next, we show the vector fields $\eta^i_{j,k}, k \neq 0$ are given by (2.31). Given the transformation

(2.28), Ψ_* maps the vector fields $\mathrm{ad}_{-F_e^i}\eta_{j,k-1}^i$ into $\mathrm{ad}_{-\bar{F}_e^i}\bar{\eta}_{j,k-1}^i$ for

$$\Psi_*\left[-F_e^i, \eta_{j,k-1}^i\right] = \left[-\Psi_* F_e^i, \Psi_*\eta_{j,k-1}^i\right] = \left[-\bar{F}_e^i, \bar{\eta}_{j,k-1}^i\right].$$

We have

$$\begin{aligned}\mathrm{ad}_{-\bar{F}_e^i}\bar{\eta}_{j,k-1}^i &= \sum_{r=1}^{i-1}\sum_{s=1}^{\lambda_r}\frac{\partial \bar{f}_s^r}{\partial y_j^{(\beta_j^i-k+1)}}\frac{\partial}{\partial \bar{x}_s^r} + \frac{\partial}{\partial y_j^{(\beta_j^i-k)}} \\ &= \bar{\kappa}_{j,k}^i + \bar{\eta}_{j,k}^i,\end{aligned}$$

and

$$\begin{aligned}\bar{\kappa}_{j,0}^i &= 0, \\ \bar{\kappa}_{j,k}^i &= \mathrm{ad}_{-\bar{F}_e^i}\bar{\eta}_{j,k-1}^i - \bar{\eta}_{j,k}^i = \begin{pmatrix}\frac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta_j^i-k+1)}} \\ 0\end{pmatrix}.\end{aligned}\tag{2.37}$$

On the other hand, it is true that

$$\eta_{j,k}^i = \mathrm{ad}_{-F_e^i}\eta_{j,k-1}^i - \kappa_{j,k}^i,$$

where $\kappa_{j,k}^i$ is the expression of $\bar{\kappa}_{j,k}^i$ in x_e^i-coordinates. Knowing $\kappa_{j,k}^i$ we can determine $\eta_{j,k}^i$. For $\kappa_{j,k}^i$, the following relation holds

$$\Psi_*\kappa_{j,k}^i = \bar{\kappa}_{j,k}^i.\tag{2.38}$$

Evidently, there is a unique vector $\kappa_{j,k}^i$ which satisfies (2.38) with Ψ_* nonsingular. Given $\bar{\kappa}_{j,k}^i$ expressed in (2.37), it is straightforward to verify the solution of $\kappa_{j,k}^i$ is given by (2.32) and $\eta_{j,k}^i$ is given by (2.31).

Sufficiency: Condition (i) of Theorem 2.3.6 guarantees the existence of $\Psi^i \in U$ such that $\eta_{j,k}^i$ satisfies the Lie bracket conditions. Given Conditions (ii)–(iii), the extended state transformation Ψ can be constructed as in the sufficiency proof of Theorem 2.2.1 and is given by (2.28). Ψ rectifies the vector fields $\eta_{j,k}^i, 0 \leqslant k \leqslant \beta_j^i, i+1 \leqslant j \leqslant p$ into unit vector fields $\bar{\eta}_{j,k}^i$. For any $\Psi^i \in U$, it is simple to verify any $\Psi(x_e^i)$ of form (2.28) is a local diffeomorphism, rectifies the prolonged vector fields $\eta_{j,k}^i$, and preserves $\bar{x}^{<i-1>}$ and $y_{[i+1,p]}^{<\beta^i>}$. Next we show Ψ transforms the ith subsystem into BTF. In $\bar{x}$-coordinates, we have

$$\begin{aligned}\Psi_*\mathrm{ad}_{-F_e^i}\eta_{j,k}^i &= \mathrm{ad}_{-\bar{F}_e^i}\bar{\eta}_{j,k}^i \\ &= \Psi_*\eta_{j,k+1}^i + \Psi_*\kappa_{j,k+1}^i \\ &= \frac{\partial}{\partial y_j^{(\beta_j^i-k-1)}} + \bar{\kappa}_{j,k+1}^i,\end{aligned}\tag{2.39}$$

where $\bar{\kappa}^i_{j,k+1}$ is given by (2.37). Since $\eta^i_{j,k}, 0 \leqslant k \leqslant \beta^i_j, i+1 \leqslant j \leqslant p$ are rectified into $\frac{\partial}{\partial y_j^{(\beta^i_j - k)}}$ by Ψ, we have

$$\mathrm{ad}_{-\bar{F}^i_e} \bar{\eta}^i_{j,k} = \sum_{r=1}^{i} \sum_{s=1}^{\lambda_r} \frac{\partial \bar{f}^r_s}{\partial y_j^{(\beta^i_j - k)}} \frac{\partial}{\partial \bar{x}^r_s} + \frac{\partial}{\partial y_j^{(\beta^i_j - k - 1)}}. \tag{2.40}$$

It is clear that (2.39) is equal to (2.40) if and only if

$$\frac{\partial \bar{f}^i}{\partial y_j^{(\beta^i_j - k)}} = 0,$$

which implies $\bar{f}^i$ is independent of $y_j^{(\beta^i_j - k)}$. Thus we have verified that $\bar{f}^i$ is independent of $y_j^{(k)}, 1 \leqslant k \leqslant \beta^i_j, i+1 \leqslant j \leqslant p$, and the ith subsystem is in BTF. Hence, the necessary conditions are sufficient. ■

2.3.3 Lowering the Highest Derivatives

Theoretically, we can apply Theorem 2.3.6 to solve the extended state transformation such that all the derivatives of the lower subsystems are removed from the ith subsystem. However, we have to solve the high order PDEs (2.34) which is non-trivial. Considering the well-developed approaches for solving first order PDEs, it is meaningful to derive the conditions on removing the highest order derivatives of lower subsystem outputs. These conditions are formulated in terms of first order PDEs as in the following Proposition.

Proposition 2.3.9. *Given an observable system* (2.1) *with the upper i subsystems given by* (2.26)–(2.27), *a local extended state transformation* (2.28), *which removes $y^{(\beta^i)}_{[i+1,p]}$ from the ith subsystem* (2.27) *can be solved from the following first-order PDEs*

$$\frac{\partial \Psi^i}{\partial x^i_e} \tilde{\eta}^i_{j,k} = 0, \quad k = 0, 1; i+1 \leqslant j \leqslant p, \tag{2.41}$$

where $\tilde{\eta}^i_{j,k}$ is defined in (2.33) *and satisfies*

$$\frac{\partial \tilde{\eta}^i_{j,k}}{\partial y_r} = 0, \quad 0 \leqslant r \leqslant i-1.$$

Proof: In this case, U is defined by (2.34) with indices $k = 0, 1$. PDEs (2.34) with $k = 0$ is equivalent to PDEs (2.41) with $k = 0$ since $\eta^i_{j,0} = \tilde{\eta}^i_{j,0}$. We only need to verify they are equivalent when $k=1$. First, we calculate $\kappa^i_{j,1}, \eta^i_{j,1}$ based on the definitions (2.31), (2.32)

$$\kappa^i_{j,1} = \begin{pmatrix} \dfrac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta^i_j)}} \\ -\left(\frac{\partial \Psi^i}{\partial x^i}\right)^{-1} \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \dfrac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta^i_j)}} \\ 0 \end{pmatrix}.$$

According to the expression of $\kappa^i_{j,1}$ and $\mathrm{ad}_{-F^i_e}\eta^i_{j,0}$, $\eta^i_{j,1}$ is written as

$$\eta^i_{j,1} = \begin{pmatrix} 0 \\ \varrho^i_{j,1} \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

where

$$\varrho^i_{j,1} = \frac{\partial f^i}{\partial y_j^{(\beta^i_j)}} + \left(\frac{\partial \Psi^i}{\partial x^i}\right)^{-1} \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \frac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta^i_j)}}.$$

From the formula of $\eta^i_{j,1}$, (2.34) with $k=1$ is written as

$$\frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \frac{\partial \bar{f}^{<i-1>}}{\partial y_j^{(\beta^i_j-k+1)}} + \frac{\partial \Psi^i}{\partial x^i} \frac{\partial f^i}{\partial y_j^{(\beta^i_j-k+1)}} + \frac{\partial \Psi^i}{\partial y_j^{(\beta^i_j-k+1)}} = 0,$$

which leads to PDEs (2.41). Hence, the transformation Ψ^i solved from PDEs (2.34) with $k=0,1$ also satisfies PDEs (2.41). On the other hand, it is straightforward to verify that for any transformation Ψ^i which solves PDEs (2.41) also satisfies PDEs (2.34). Therefore, PDEs (2.34) and (2.41) are equivalent in that they solve the same transformation Ψ^i, with $\frac{\partial \Psi^i}{\partial x^i}$ nonsingular. Additionally, from the necessity proof of Theorem 2.3.6, we have the condition that $\tilde{\eta}^i_{j,k}$ has to be independent of $y_r, 0 \leqslant r \leqslant i-1$. □

Remark 2.3.10. The ith subsystem $1 \leqslant i \leqslant p-1$ in (2.3) can be locally transformed to BTF (2.25) if and only if Proposition 2.3.9 can be applied step by step until $\beta^i_j = 0, i+1 \leqslant j \leqslant p$. □

Remark 2.3.11. Considering the existence conditions of the transformation eliminating one of the highest derivatives from the ith subsystem, we notice the difference between Theorems 2.3.6 and 2.2.1. Here, given the ith subsystem in Observable Form (2.3), the function φ^i does not have to be affine in the highest order derivative of the lower subsystem outputs. This is shown by verifying the necessary Lie bracket condition of $\eta^i_{j,0}, \eta^i_{j,1}$

$$[\eta^i_{j,0}, \eta^i_{j,1}] = \begin{pmatrix} 0 \\ \frac{\partial^2 f^i}{\partial (y_j^{(\beta^i_j)})^2} + \left(\frac{\partial \Psi^i}{\partial x^i}\right)^{-1} \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \frac{\partial^2 \bar{f}^{<i-1>}}{\partial (y_j^{(\beta^i_j)})^2} \\ 0 \end{pmatrix} = 0.$$

Considering the fact $\frac{\partial \Psi^i}{\partial x^i}$ is nonsingular, we conclude that

$$\frac{\partial^2 f^i}{\partial (y_j^{(\beta^i_j)})^2} + \left(\frac{\partial \Psi^i}{\partial x^i}\right)^{-1} \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \frac{\partial^2 \bar{f}^{<i-1>}}{\partial (y_j^{(\beta^i_j)})^2} = 0$$

has the same solutions as

$$\frac{\partial \Psi^i}{\partial x^i} \frac{\partial^2 f^i}{\partial (y_j^{(\beta^i_j)})^2} + \frac{\partial \Psi^i}{\partial \bar{x}^{<i-1>}} \frac{\partial^2 \bar{f}^{<i-1>}}{\partial (y_j^{(\beta^i_j)})^2} = 0. \tag{2.42}$$

Hence, φ^i is required to be affine in the highest order derivative if and only if the second term in the left hand side of (2.42) is zero. Compared with results in Theorem 2.2.1, this approach potentially introduces additional freedom when transforming a system to a BTF. $\square$

2.3.4 Example

We consider a three-output system with indices $\lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 5$ and

$$\begin{aligned} \dot{x}^1 &= \bar{f}^1 = \begin{pmatrix} \bar{x}^1_2 \\ x^2_1 \end{pmatrix}, \\ \dot{x}^2 &= \bar{f}^2 = \begin{pmatrix} x^2_2 \\ y_3^{(4)} + \bar{x}^1_1 \dot{y}_3 + x^2_1 \bar{x}^1_1 \end{pmatrix}, \\ \dot{x}^3 &= f^3 = \begin{pmatrix} x^3_2 \\ x^3_3 \\ x^3_4 \\ x^3_5 \\ \varphi^3(x) \end{pmatrix}, \\ y &= \begin{pmatrix} \bar{x}^1_1 \\ x^2_1 \\ x^3_1 \end{pmatrix}. \end{aligned}$$

For a general function $\varphi^3(x)$, if the system is arranged in a different order, for instance using indices $(5, 2, 2)$, although Theorem 2.2.1 is applicable, its conditions do not necessarily hold. Hence, we order the system with indices $(2, 2, 5)$. The first and third subsystems are already in BTF, and the second subsystem is not in BTF for $y_3^{(4)}$ and $\dot{y}_3$ appear in f^2. For the second subsystem, $\alpha_3^2 = 4 > 2 = \lambda_2$, and Theorem 2.2.1 is not applicable. However, Proposition 2.3.2 can be applied to check if $y^{(4)}, y^{(3)}$ can be removed. Defining the prolonged vector field and extended state

$$F_e^{21} = ((\bar{f}^1)^T, (f^2)^T, (f^3)^T)^T,$$
$$x_e^{21} = ((\bar{x}^1)^T, (x^2)^T, (y_3^{<4>})^T)^T,$$

the conditions to eliminate $y_3^{(4)}, y_3^{(3)}$ are

$$\left[\mathrm{ad}^r_{-F_e^2} \frac{\partial}{\partial y_3^{(4)}}, \mathrm{ad}^s_{-F_e^2} \frac{\partial}{\partial y_3^{(4)}}\right] = 0, \quad 0 \leqslant r, s \leqslant 2.$$

The transformation to remove $y_3^{(4)}, y_3^{(3)}$ from f^2 can be solved from

$$\frac{\partial \Psi^{21}}{\partial x_e^{21}} \mathrm{ad}^r_{-F_e^2} \frac{\partial}{\partial y_3^{(4)}} = 0, \quad 0 \leqslant r \leqslant 2.$$

After applying the transformation

$$\Psi^{21}(x_e^{21}) = \begin{pmatrix} x_1^2 - \ddot{y}_3 \\ x_2^2 - y_3^{(3)} \end{pmatrix} = \bar{x}^{21},$$

the first two subsystems are written as

$$\dot{\bar{x}}^1 = \begin{pmatrix} \bar{x}_2^1 \\ \bar{x}_1^{21} + \ddot{y}_3 \end{pmatrix},$$
$$\dot{\bar{x}}^{21} = \begin{pmatrix} \bar{x}_2^{21} \\ (\bar{x}_1^{21} + \ddot{y}_3 + \dot{y}_3)\bar{x}_1^1 \end{pmatrix}.$$

To apply Proposition 2.3.9, we redefine the prolonged vector field and extended state

$$F_e^{22} = ((\bar{f}^1)^T, (f^{21})^T, \dot{y}_3, \ddot{y}_3, y_3^{(3)})^T,$$
$$x_e^{22} = ((\bar{x}^1)^T, (\bar{x}^{21})^T, (y_3^{<2>})^T)^T,$$

where $\dot{\bar{x}}^{21} = f^{21}$. Computing $\tilde{\eta}_{3,0}^{22}, \tilde{\eta}_{3,1}^{22}$

$$\tilde{\eta}_{3,0}^{22} = \frac{\partial}{\partial y_3^{(2)}},$$
$$\tilde{\eta}_{3,1}^{22} = \frac{\partial}{\partial \bar{x}_2^1} + \bar{x}_1^1 \frac{\partial}{\partial \bar{x}_2^{21}} + \frac{\partial}{\partial \dot{y}_3},$$

we can solve the PDEs (2.41) for Ψ^{22}

$$\Psi^{22} = \begin{pmatrix} \bar{x}_1^{21} \\ \bar{x}_2^{21} - \bar{x}_1^1 \bar{x}_2^1 \end{pmatrix} = \bar{x}^{22}.$$

Since $y_1 = \bar{x}_1^1$ is treated as a state instead of a parameter, conditions (2.35) are unnecessary. Ψ^{22} removes $\ddot{y}_3$ from the second subsystem and leads to

$$\dot{\bar{x}}^{22} = \begin{pmatrix} \bar{x}_2^{22} + \bar{x}_1^1 \bar{x}_2^1 \\ \bar{x}_1^1 \dot{y}_3 - (\bar{x}_2^1)^2 \end{pmatrix},$$

where $\dot{y}_3$ still appears. Applying Proposition 2.3.9 once again and iterating the same procedure for solving the transformation Ψ^{23}, we define

$$F_e^{23} = ((\bar{f}^1)^T, (f^{22})^T, \dot{y}_3, \ddot{y}_3)^T,$$
$$x_e^{23} = ((\bar{x}^1)^T, (\bar{x}^{22})^T, (y_3^{<1>})^T)^T,$$

where $\dot{\bar{x}}^{22} = f^{22}$. Computing

$$\tilde{\eta}_{3,0}^{23} = \frac{\partial}{\partial \dot{y}_3},$$
$$\tilde{\eta}_{3,1}^{23} = \bar{x}_1^1 \frac{\partial}{\partial \bar{x}_2^{22}} + \frac{\partial}{\partial y_3}.$$

we solve PDEs (2.41) for transformation Ψ^{23}

$$\Psi^{23} = \begin{pmatrix} \bar{x}_1^{22} \\ \bar{x}_2^{22} - y_3 \bar{x}_1^1 \end{pmatrix} = \bar{x}^2.$$

Finally, the second subsystem in BTF is

$$\dot{\bar{x}}^2 = \begin{pmatrix} \bar{x}_2^2 + \bar{x}_1^1(\bar{x}_2^1 + y_3) \\ -\bar{x}_2^1(\bar{x}_2^1 + y_3) \end{pmatrix}.$$

2.4 Observer Design Examples

Industrially relevant examples which are transformable to BTF include: a Web Machine [100] with $y = (\theta_1, \theta_2, \theta_3, r_1, r_2)^T$ and indices $(2, 2, 1, 1, 1)$, a Synchronous Motor [109, Sec. 1.10.10] with $y = (\delta, i_a, i_b)^T$ and indices $(2, 1, 1)$, a Permanent Magnet Stepper (PMS) motor [40, Sec. 3.2] with output $y = (q, i_1, i_2)^T$ and indices $(2, 1, 1)$, a Brushless DC (BLDC) motor [40, Sec. 4.2] with output $y = (q, i_a, i_b)^T$ and indices $(2, 1, 1)$, and a one degree-of-freedom MAGLEV system with $y = (x_1^1, x_1^2)^T$, indices $(2, 1)^T$, and Observable Form dynamics [110]

$$\begin{aligned} \dot{x} &= \begin{pmatrix} x_2^1 \\ g - \frac{\beta(x_1^2)^2}{m(x_1^1)^2} \\ \frac{2\beta x_1^2 x_2^1}{2\beta x_1^1 + \alpha(x_1^1)^2} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{x_1^1}{\alpha x_1^1 + 2\beta} \end{pmatrix} u, \\ y &= \begin{pmatrix} x_1^1 \\ x_1^2 \end{pmatrix}, \end{aligned} \tag{2.43}$$

where α, β, m, g are positive constants, x_1^1 is position, x_2^1 is velocity, x_1^2 is the current, and $x = (x_1^1, x_2^1, x_1^2)^T$. Many practical systems, such as the ones above-mentioned, are forced and control affine. These examples are a special case of that discussed in Section 2.2.5. This is because the input vector field for the ith subsystem in BTF depends on the output alone. Hence, complexity of the observer design is unaffected by the forced part of the system.

When transformed into an Observable Form, the above-mentioned systems are also in BTF. To illustrate the transformation to BTF for a system whose observable form is not in BTF, we consider the cart-pendulum system

$$\begin{pmatrix} M+m & ml\cos\phi \\ ml\cos\phi & I+ml^2 \end{pmatrix}\begin{pmatrix} \ddot{z} \\ \ddot{\phi} \end{pmatrix}+\begin{pmatrix} b & -ml\dot{\phi}\sin\phi \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \dot{z} \\ \dot{\phi} \end{pmatrix}-\begin{pmatrix} 0 \\ mgl\sin\phi \end{pmatrix}=\begin{pmatrix} \mathcal{N} \\ 0 \end{pmatrix}, \tag{2.44}$$

where M is the mass of the cart, m is the mass of the pendulum, b is a viscous friction coefficient of the cart, $\mathcal{N}$ is the force input applied to the cart, l is the length to the pendulum's center of mass, z is the cart's position, and ϕ is the pendulum angle from vertical position. For simplicity, we take the pendulum inertia $I = 0$ as this does not affect the BTF-based observer design. For the output $y = (z, \phi)^T$, system (2.44) is observable with indices $(2, 2)$. However, it is not transformable to BTF because the first subsystem has dependence on the output derivative of the second subsystem from the term $(\dot{\phi})^2$. This implies the Lie bracket conditions (2.6) in Theorem 2.2.1 do not hold. Reordering the outputs as $y = (\phi, z)^T$, we can verify the forced system (2.44) can be transformed to BTF. Leaving the output as $(\phi, z)^T$ leads to $\dot{\phi}^2$ appearing in the first subsystem of the BTF. This increases the complexity of the observer design. Hence, we introduce the following output transformation to allow for linear error dynamics

$$\begin{aligned} y_1 &= \gamma(\phi) = \sqrt{M}\mathcal{E}\left(\phi, -\frac{m}{M}\right), \\ y_2 &= z, \end{aligned} \tag{2.45}$$

where $\mathcal{E}(\xi, k) = \int_0^\xi \sqrt{1 - k\sin^2 t}\,dt$ is an elliptic integral of the second kind. The Observable Form of system (2.44) with this new output is

$$\begin{aligned} f^1(x) &= \begin{pmatrix} x_2^1 \\ x_2^2\theta_1^1 + \theta_2^1 \end{pmatrix}, & g^1(y) &= \begin{pmatrix} 0 \\ \theta_3^1 \end{pmatrix}, \\ f^2(x) &= \begin{pmatrix} x_2^2 \\ x_2^2\theta_1^2 + (x_2^1)^2\theta_2^2 + \theta_3^2 \end{pmatrix}, & g^2(y) &= \begin{pmatrix} 0 \\ \theta_4^2 \end{pmatrix}, \end{aligned} \tag{2.46}$$

where $x^1 = (x_1^1, x_2^1)^T = (\phi, \dot{\phi})^T, x^2 = (x_1^2, x_2^2)^T = (z, \dot{z})^T$, g^1, g^2 are input vector fields, and

$$\theta_1^1 = \frac{b\cos(\gamma^{-1})}{l\eta}, \quad \theta_2^1 = \frac{-g\sin(\gamma^{-1})(M+m)}{l\eta}, \quad \theta_3^1 = \frac{-\cos(\gamma^{-1})}{l\eta},$$

$$\theta_1^2 = -\frac{b}{\eta^2}, \quad \theta_2^2 = \frac{ml\sin(\gamma^{-1})}{\eta^4}, \quad \theta_3^2 = \frac{mg\cos(\gamma^{-1})\sin(\gamma^{-1})}{\eta^2}, \quad \theta_4^2 = \frac{1}{\eta^2},$$

with $\eta(y_1) = \sqrt{M + m \cdot \sin^2(\gamma^{-1}(y_1))}$. In order to put the system into BTF, we need to remove the second subsystem output derivative $\dot{y}_2 = x_2^2$ from the first subsystem dynamics. Defining the extended state $x_e^1 = (x_1^1, x_2^1, y_2, \dot{y}_2)^T$ and prolonged vector field $F_e^1 = ((f^1)^T, \dot{y}_2, \ddot{y}_2)^T$ we can verify the Lie bracket conditions (2.6) of Theorem 2.2.1 hold. The transformation Ψ^1 can be solved from PDEs (2.7) as $\Psi^1 = (x_1^1, x_2^1 - x_1^2\theta_1^1)^T$. Applying Ψ^1 puts the system (2.46) in BTF

$$\bar{f}^1(\bar{x}^1, y_2) = \begin{pmatrix} \bar{x}_2^1 + \bar{\theta}_1^1 \\ \bar{\theta}_2^1\bar{x}_2^1 + \bar{\theta}_3^1 \end{pmatrix}, \qquad \bar{g}^1(y) = \begin{pmatrix} 0 \\ \bar{\theta}_4^1 \end{pmatrix},$$

$$\bar{f}^2(\bar{x}) = \begin{pmatrix} \bar{x}_2^2 \\ \bar{x}_2^2\bar{\theta}_1^2 + (\bar{x}_2^1)^2\bar{\theta}_2^2 + \bar{x}_2^1\bar{\theta}_3^2 + \bar{\theta}_4^2 \end{pmatrix}, \quad \bar{g}^2(y) = \begin{pmatrix} 0 \\ \bar{\theta}_5^2 \end{pmatrix}$$

with $\bar{h}(\bar{x}) = (\bar{h}_1(\bar{x}), \bar{h}_2(\bar{x}))^T = (\bar{x}_1^1, \bar{x}_1^2)^T$, and where $\bar{x}_1^2 = x_1^2, \bar{x}_2^2 = x_2^2, \bar{\theta}_j^1, 1 \leqslant j \leqslant 4$, $\bar{\theta}_j^2, 1 \leqslant j \leqslant 5$ are some functions of y, and $\bar{g}^1, \bar{g}^2$ are input vector fields. The observer and the error dynamics for the first subsystem are

$$\dot{\hat{\bar{x}}}^1 = \begin{pmatrix} 0 & 1 \\ 0 & \bar{\theta}_2^1 \end{pmatrix} \hat{\bar{x}}^1 + \begin{pmatrix} \bar{\theta}_1^1 \\ \bar{\theta}_3^1 \end{pmatrix} + \begin{pmatrix} 0 \\ \bar{\theta}_4^1 \end{pmatrix} \mathcal{N} + \begin{pmatrix} l_1^1 \\ l_2^1 \end{pmatrix} (y_1 - \hat{y}_1),$$

$$\dot{\tilde{\bar{x}}}^1 = \begin{pmatrix} l_1^1 & 1 \\ l_2^1 & \bar{\theta}_2^1 \end{pmatrix} \tilde{\bar{x}}^1 \triangleq \mathcal{A}_1 \tilde{\bar{x}}^1,$$

where $\tilde{\bar{x}}^1 = \bar{x}^1 - \hat{\bar{x}}^1$. The error dynamics of the first subsystem is LTV, hence one approach to determine an observer gain is to apply [136, Cor. 8.4] so that the maximum eigenvalue $\lambda_{\max}(\mathcal{A}_1 + \mathcal{A}_1^T) < -\epsilon < 0$. This ensures the error dynamics of the first subsystem is uniformly exponentially stable. For the second subsystem, the observer and error dynamics are

$$\dot{\hat{\bar{x}}}^2 = \begin{pmatrix} 0 & 1 \\ 0 & \bar{\theta}_1^2 \end{pmatrix} \hat{\bar{x}}^2 + \begin{pmatrix} 0 \\ (\hat{\bar{x}}_2^1)^2\bar{\theta}_2^2 + \hat{\bar{x}}_2^1\bar{\theta}_3^2 + \bar{\theta}_4^2 \end{pmatrix} + \begin{pmatrix} 0 \\ \bar{\theta}_5^2 \end{pmatrix} \mathcal{N} + \begin{pmatrix} l_1^2 \\ l_2^2 \end{pmatrix} (y_2 - \hat{y}_2),$$

$$\dot{\tilde{\bar{x}}}^2 = \underbrace{\begin{pmatrix} l_1^2 & 1 \\ l_2^2 & \bar{\theta}_1^2 \end{pmatrix}}_{\mathcal{A}_2} \tilde{\bar{x}}^2 + \underbrace{\begin{pmatrix} 0 \\ ((\bar{x}_2^1)^2 - (\hat{\bar{x}}_2^1)^2)\bar{\theta}_2^2 + \tilde{\bar{x}}_2^1\bar{\theta}_3^2 \end{pmatrix}}_{\text{disturbance}},$$

where $\tilde{\bar{x}}^2 = \bar{x}^2 - \hat{\bar{x}}^2$. The error dynamics consists of a LTV part forced by a "disturbance" which converges to zero exponentially assuming the system state is bounded. To ensure the second subsystem error converges to zero, we apply the result of [104, Lem. III.1] and place the eigenvalues of $\mathcal{A}_2 + \mathcal{A}_2^T$ in the open LHP so that the origin of the second subsystem error dynamics is exponentially stable. Finally, we remark that without output transformation (2.45) a BTF-based observer design would still be possible; however, a relatively simple LTV design in BTF coordinates results if output transformation (2.45) is used.

2.5 Summary

This chapter considered existence conditions of an established Block Triangular Form for a broad class of uncontrolled multi-output systems. Given the constraint on the ordering of the observability indices $\lambda_1, \ldots, \lambda_p$, necessary and sufficient conditions were derived in Theorem 2.2.1. Sufficient conditions were given in Theorem 2.2.8 to ensure that the output is linear in BTF coordinates. Next, Theorem 2.3.6 was presented to generalize the work of Theorem 2.2.1 by eliminating the ordering restriction. Existence conditions of a state transformation for lowering the highest order derivative dependence on the lower system outputs were proposed as a special case of Theorem 2.3.6. Finally, a physical example illustrated the BTF construction and a typical observer design.

Chapter 3

A Block Triangular Observer Form (BTOF)

In Chapter 2, a BTF was introduced and its existence conditions were derived. Examples of the advantage of a BTF-based observer design were shown. Although the complexity of a decentralized observer design for a system in BTF is generally lower than a centralized approach, it is relatively difficult to establish error convergence and the design procedure itself is less systematic. This motivates the work in this chapter which presents a Block Triangular Observer Form (BTOF) that leads to a straightforward observer design. The BTOF originally appeared in [134]. A BTOF divides a p-output nonlinear system (2.1) into p subsystems and is a special case of a BTF. The BTOF dynamics of the ith subsystem contain two terms: a linear function of the ith subsystem state, and a term with nonlinear dependence on y and on "upper" subsystem states, i.e., the states of subsystem $j, j < i$. This nonlinear term plays the role of the output injection in EEL (see Section 1.1), and allowing for its dependence on upper subsystem states gives the entire system a triangular structure. The ith component of the output in BTOF coordinates is a linear function of the ith subsystem state. As expected, the existence conditions for a BTOF are more restrictive than a BTF but weaker than an OF [85, 165, 167].

As mentioned above, a motivation for BTOF is a straightforward observer design similar to an EEL design. The difference is due to the upper subsystem states in the nonlinear terms which in general cannot be exactly canceled in the error dynamics. However, the exponentially decaying nonlinear terms which appear in the error dynamics do not complicate the proof of convergence. Related recent work in [141] also considers a triangular cascade structure which allows for a more general nonlinear dependence of the ith subsystem dynamics on the ith subsystem state, and the resulting observer design uses high-gain to prove error convergence. We restrict our attention to the more structured BTOF. Relative to work in [141], the BTOF considered here has the advantage of a simpler error convergence proof and condition on the observer gain. As well, no high-gain approach is required.

In Section 3.1 we present BTOF existence conditions and demonstrate counterexample systems which show the conditions established in previous work [134] are neither necessary nor sufficient. In Section 3.2 we discuss the error convergence. The work in Section 3.1 is further extended in Section 3.3 by removing the parameter assumption used to simplify the derivation and statement of the existence conditions. In Section 3.4 we discuss some examples with engineering relevance and provide the details of a BTOF-based observer design which is compared to the recent method in [82].

3.1 BTOF Existence Conditions I

We begin by defining a BTOF.

Definition 3.1.1. An observable system (2.1) is said to be in a BTOF if its ith subsystem has the form

$$\begin{aligned} \dot{z}^i &= A^i z^i + \gamma^i(z^{<i-1>}, z_1^i, y_{[i+1,p]}), \\ y_i &= C^i z^i, \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} z^i &= (z_1^i, \dots, z_{\lambda_i}^i)^T \in \mathbb{R}^{\lambda_i}, \\ \gamma^i(\cdot) &= (\gamma_1^i(\cdot), \dots, \gamma_{\lambda_i}^i(\cdot))^T \in \mathbb{R}^{\lambda_i}, \\ z^{<i>} &= ((z^1)^T, \dots, (z^i)^T)^T \in \mathbb{R}^{\nu_i}, \end{aligned}$$

and A^i, C^i are defined in (1.7). □

In order to derive the existence conditions for the BTOF, we recall the two methods used for the BTF in Chapter 2. These approaches arise naturally from the decentralized design. The first approach treats a subset of the outputs and states appearing in the current subsystem dynamics as parameters. This assumption restricts the class of applicable systems. However, the existence conditions are easier to derive. The second approach treats all system variables appearing in the current subsystem dynamics as functions of state. This approach enlarges the class of applicable systems while increasing the complexity of the derivation of existence conditions. As with the BTF, in this chapter we first derive the BTOF existence conditions under the following parameter assumption.

Assumption 3.1.2. For the ith subsystem, the lower subsystem states $\bar{x}^j$ and outputs $y_j, i+1 \leqslant j \leqslant p$ are treated as constants. □

Later in Section 3.3, a state approach is applied to establish existence conditions, and we demonstrate the limitation of the parameter approach. We make another assumption to simplify the presentation of the existence conditions. This assumption is natural since the BTOF is a special case of the BTF, and the BTOF existence conditions are verified sequentially subsystem-at-a-time.

Assumption 3.1.3. When transforming the ith subsystem, the first $i-1$ subsystem are already in BTOF, and the other subsystems are in BTF. □

3.1.1 Main Result

Theorem 3.1.4. *Let system* (2.1) *be locally observable in* $U_0 \subset \mathbb{R}^n$ *with indices* $\lambda_i, 1 \leqslant i \leqslant p$. *Given Assumptions 3.1.2 and 3.1.3, there exists a local diffeomorphism* $\Phi(\bar{x}_e^i) = ((z^{<i-1>})^T, (\Phi^i(\bar{x}_e^i))^T)^T$ *such that in the new coordinates, the ith subsystem is of the form* (3.1) *if and only if in* U_0

(i) the first $\nu_{i-1} = \sum_{k=1}^{i-1} \lambda_k$ *components of* $\mathrm{ad}^k_{-\bar{F}^i}\bar{g}^i$, $0 \leqslant k \leqslant \lambda_i - 1$ *are zero, where*

$$\bar{F}^i = ((\bar{f}^1)^T, (\bar{f}^2)^T, \dots, (\bar{f}^i)^T)^T \in \mathbb{R}^{\nu_i},$$
$$\bar{x}_e^i = ((z^{<i-1>})^T, (\bar{x}^i)^T)^T \in \mathbb{R}^{\nu_i},$$

and the ν_i-dimensional starting vector $\bar{g}^i$ is the unique solution of

$$L_{\bar{g}^i} L^k_{\bar{F}^i} \bar{h}_i = \delta_{k,\lambda_i - 1}, \quad 0 \leqslant k \leqslant \lambda_i - 1; \tag{3.2}$$

(ii) the Lie bracket conditions are satisfied

$$\left[\mathrm{ad}^r_{-\bar{F}^i}\bar{g}^i, \mathrm{ad}^s_{-\bar{F}^i}\bar{g}^i\right] = 0, \quad 0 \leq r, s \leq \lambda_i - 1;$$

(iii) the vector fields are independent of lower subsystem outputs

$$\frac{\partial}{\partial y_j}\mathrm{ad}^r_{-\bar{F}^i}\bar{g}^i = 0, \quad 0 \le r \le \lambda_i - 1; i+1 \leqslant j \leqslant p;$$

(iv) the ith subsystem output is independent of lower subsystem states

$$\frac{\partial \bar{h}_i}{\partial \bar{x}^j_k} = 0, \quad 1 \leqslant k \leqslant \lambda_j; i+1 \leqslant j \leqslant p.$$

The state transformation Φ^i is the solution of the λ_i^2 PDEs

$$\frac{\partial \Phi^i(\bar{x}^i_e)}{\partial \bar{x}^i_e}\left(\mathrm{ad}^{\lambda_i-1}_{-\bar{F}^i}\bar{g}^i, \dots, \bar{g}^i\right) = I_{\lambda_i}. \tag{3.3}$$

Remark 3.1.5. Given Assumptions 3.1.2 and 3.1.3, Condition (iii) ensures

$$\frac{\partial}{\partial \bar{x}^j_k}\mathrm{ad}^r_{-\bar{F}^i}\bar{g}^i = 0, \quad \begin{cases} 0 \leqslant r \leqslant \lambda_i - 1; \\ 1 \leqslant k \leqslant \lambda_j; \\ i+1 \leqslant j \leqslant p. \end{cases} \tag{3.4}$$

Condition (3.4) means the vector fields $\mathrm{ad}^r_{-\bar{F}^i}\bar{g}^i, 0 \leqslant r \leqslant \lambda_i - 1$ depend on $\bar{x}^i_e$ which in turn guarantees that the state transformation Φ is independent of parameters. Condition (iv) is imposed to ensure the output map $h_i(\bar{x})$ is not a function of $\bar{x}^j, i+1 \leqslant j \leqslant p$ after the transformation Φ^i. A simple example can explain its significance. Considering a system

$$\begin{aligned} \dot{z} &= \begin{pmatrix} A^1 & 0 \\ 0 & A^2 \end{pmatrix} z + \begin{pmatrix} \gamma^1(y) \\ \gamma^2(z^1, y_2) \end{pmatrix}, \\ y &= \begin{pmatrix} z^1_1 + z^2_2 \\ z^2_1 \end{pmatrix}, \end{aligned}$$

we can verify all the conditions of Theorem 3.1.4 hold except Condition (iv). Without Condition (iv), we might falsely conclude the system can be put into a BTOF by an identity transformation which actually cannot transform the system into a BTOF. □

Remark 3.1.6. As in Section 2.2 the solution of (3.3) can be constructed by a composition of flows formula. We denote the vector fields

$$\begin{aligned} Z^j_k &= \frac{\partial}{\partial z^j_k}, \quad 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \\ Y_r &= \mathrm{ad}^{\lambda_i - r}_{-\bar{F}^i}\bar{g}^i, \quad 1 \leqslant r \leqslant \lambda_i \end{aligned}$$

as $W_1, \ldots, W_{\nu_i}$ and order them as follows: the first ν_{i-1} vector fields are the Z_k^j (ordered as $Z_1^1, \ldots, Z_{\lambda_1}^1, \ldots, Z_1^{i-1}, \ldots, Z_{\lambda_{i-1}}^{i-1}$), followed by $Y_1, \ldots, Y_{\lambda_i}$. Given an initial condition

$$s_0 = ((z_0^{<i-1>})^T, (\bar{x}_0^i)^T)^T \in \mathbb{R}^{\nu_i},$$

the change of state coordinates Φ is given by

$$\Phi(\bar{x}_e^i) = \Psi^{-1}(\bar{x}_e^i) + s_0,$$

where

$$\Psi(t_1, \ldots, t_{\nu_i - 1}, t_{\nu_i}) = \psi_{W_{\nu_i}}^{t_{\nu_i}}(\ldots(\psi_{W_2}^{t_2}(\psi_{W_1}^{t_1}(s_0))\ldots)).$$

Thus the solution of the PDE (3.3) can be obtained by integrating ODEs and performing a nonlinear map inversion. Standard numerical routines are available to accomplish this procedure. □

Remark 3.1.7. For a globally observable system (2.1), a globally defined transformation to BTOF exists if

(i) BTF coordinates are globally defined;

(ii) the conditions in Theorem 3.1.4 are satisfied globally;

(iii) the vector fields $\mathrm{ad}_{-\bar{F}^i}^r \bar{g}^i, 0 \leqslant r \leqslant \lambda_i - 1, 1 \leqslant i \leqslant p$ are complete.

□

Remark 3.1.8. A main difference between Theorem 3.1.4 and work in [134] is that the latter uses $\bar{F}^i = \bar{f}^i$ in Conditions (i)–(iii). Also, previously a λ_i-dimensional starting vector $\bar{g}^i$ is defined as in (3.2) with $\bar{F}^i = \bar{f}^i$. Taking $\bar{F}^i = \bar{f}^i$ implies the effect of the upper subsystems is ignored. This leads to the vector fields $\mathrm{ad}_{-\bar{f}^i}^k \bar{g}^i$, $0 \leqslant k \leqslant \lambda_i - 1$ defined on $\mathbb{R}^{\lambda_i}$. However, when applying the transformation Φ^i defined on $\mathbb{R}^{\nu_i}$, $z^{<i-1>}$ and $\bar{x}^i$ are treated as states. The transformation Φ^i defined on $\mathbb{R}^{\nu_i}$ cannot guarantee Condition (ii) (with $\bar{F}^i = \bar{f}^i$), which was defined on $\mathbb{R}^{\lambda_i}$, still holds after transformation. However, if we impose

$$\frac{\partial}{\partial z_k^j} \mathrm{ad}_{-\bar{f}^i}^r \bar{g}^i = 0, \quad \begin{cases} 1 \leqslant k \leqslant \lambda_j; \\ 0 \leqslant r \leqslant \lambda_i - 1; \\ 1 \leqslant j \leqslant i - 1, \end{cases} \tag{3.5}$$

then Condition (ii) (with $\bar{F}^i = \bar{f}^i$) still holds after transformation to BTOF. Condition (3.5) implies the upper subsystem states are treated as parameters and the transformation solved from (3.3) depends only on $\bar{x}^i$. Condition (3.5) limits the applicability of the result to a smaller class of systems. Finally, Condition (iv) does not appear in previous work. □

Remark 3.1.9. The BTOF for the control affine system (2.21) can be generalized as in Section 2.2.5:

$$\begin{aligned} \dot{z}^i &= A^i z^i + \gamma^i(z^{<i-1>}, z_1^i, y_{[i+1,p]}) + \bar{\varrho}^i(z^{<i-1>}, z_1^i, y_{[i+1,p]})u, \\ y_i &= C^i z^i. \end{aligned}$$

The existence conditions of the generalized BTOF above can be established analogously. In addition to the conditions in Theorem 3.1.4, we require

$$\left[\mathrm{ad}^k_{-\bar{F}^i_e}\bar{g}^i, \bar{\varrho}^i_e\right] = 0, \quad 0 \leqslant k \leqslant \lambda_i - 2,$$

where $\bar{\varrho}^i_e = (0, (\bar{\varrho}^i))^T$. □

3.1.2 Counterexamples

In this section two examples illustrate that the existence conditions as previously stated are neither necessary nor sufficient. In order to show existence conditions given in [134] are not necessary, we consider a two-output system

$$\begin{aligned} \dot{\bar{x}}^1 &= \bar{f}^1(\bar{x}) = \begin{pmatrix} \bar{x}_2^1 \\ \bar{x}_3^1 \\ \bar{x}_1^1 + y_2 \end{pmatrix}, \\ \dot{\bar{x}}^2 &= \bar{f}^2(\bar{x}) = \begin{pmatrix} \bar{x}_2^2 \\ -\frac{2\bar{x}_3^2}{(\bar{x}_3^1)^2} \\ \left(\bar{x}_3^2 + (\bar{x}_2^2)^2\right)\bar{x}_3^1 \end{pmatrix}, \\ y &= \begin{pmatrix} \bar{h}_1(\bar{x}) \\ \bar{h}_2(\bar{x}) \end{pmatrix} = \begin{pmatrix} \bar{x}_1^1 \\ \bar{x}_1^2 \end{pmatrix}. \end{aligned}$$

The first subsystem is already in a BTOF. We consider the transformation of the second subsystem. Using (3.2) with $\bar{F}^2 = \bar{f}^2$ we solve for the three-dimensional starting vector $\bar{g}^2$ and check Condition (ii)

$$\bar{g}^2 = \begin{pmatrix} \mathrm{d}\bar{h}_2 \\ \mathrm{d}L_{\bar{f}^2}\bar{h}_2 \\ \mathrm{d}L^2_{\bar{f}^2}\bar{h}_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2}(\bar{x}^1_3)^2 \frac{\partial}{\partial \bar{x}^2_3}, \tag{3.6}$$

$$\left[\mathrm{ad}_{\bar{f}^2}\bar{g}^2, \bar{g}^2\right] = \left[\mathrm{ad}^2_{\bar{f}^2}\bar{g}^2, \bar{g}^2\right] = 0,$$

$$\left[\mathrm{ad}_{\bar{f}^2}\bar{g}^2, \mathrm{ad}^2_{\bar{f}^2}\bar{g}^2\right] = 2\bar{x}^1_3 \frac{\partial}{\partial \bar{x}^2_3} \neq 0.$$

Since Condition (ii) is not satisfied with $\bar{F}^2 = \bar{f}^2$, we falsely conclude that the second subsystem is not transformable to BTOF. On the other hand, using Theorem 3.1.4 we solve (3.2) for the starting vector

$$\bar{g}^2 = -\frac{1}{2}(\bar{x}^1_3)^2 \frac{\partial}{\partial \bar{x}^2_3},$$

and verify the other conditions are satisfied. Therefore, the second subsystem is in fact transformable to a BTOF. The transformation for the second subsystem Φ^2 is determined from the PDEs (3.3) with $\lambda_i = 3, i = 2$

$$\begin{aligned} z^2_1 &= \bar{x}^2_1, \\ z^2_2 &= \frac{1}{\bar{x}^1_3}(\bar{x}^2_2\bar{x}^1_3 + 2\bar{x}^1_2\bar{x}^2_1 + (\bar{x}^2_1)^2 - (\bar{x}^1_3)^2\bar{x}^2_1), \\ z^2_3 &= \frac{1}{(\bar{x}^1_3)^2}\left[\left(\frac{\bar{x}^2_1}{2} - 2 + \bar{x}^1_2\right)\bar{x}^2_1(\bar{x}^1_3)^2 - \bar{x}^2_2(\bar{x}^1_3)^3 + 2\bar{x}^2_2(\bar{x}^1_2 + \bar{x}^2_1)\bar{x}^1_3 - 2\bar{x}^2_3\right. \\ &\quad \left. + 2(\bar{x}^1_2)^2\bar{x}^2_1 + 2\bar{x}^1_2(\bar{x}^2_1)^2 + \frac{2}{3}(\bar{x}^2_1)^3\right], \end{aligned}$$

which puts the second subsystem into a BTOF

$$\begin{aligned} \dot{z}^2_1 &= z^2_2 - \frac{z^2_1}{z^1_3}\left[2z^1_2 + z^2_1 - (z^1_3)^2\right], \\ \dot{z}^2_2 &= z^2_3 - \frac{z^2_1}{6(z^1_3)^2}\left[12(z^1_3)^2 z^1_2 - 24(z^1_3)^2 + 9(z^1_3)^2 z^2_1 + 30 z^1_2 z^2_1 + 10(z^2_1)^2 + 24(z^1_2)^2\right], \\ \dot{z}^2_3 &= -\frac{z^2_1}{3(z^1_3)^3}\left[12(z^1_2)^3 - 6(z^1_3)^2 z^2_1 - 12(z^1_3)^2 z^1_2 - 3(z^1_3)^4 + 24(z^1_2)^2 z^2_1\right. \\ &\quad \left. + 4(z^2_1)^3 + 16 z^1_2 (z^2_1)^2\right]. \end{aligned}$$

Here we have denoted $z^1_2 = \bar{x}^1_2, z^1_3 = \bar{x}^1_3$ as the first subsystem is already in BTOF. Hence, this example shows that the conditions as stated in previous work are not necessary.

Next, we show that the conditions given in [134] are not sufficient by considering the BTF system

$$\begin{aligned}
\dot{\bar{x}}^1 &= \bar{f}^1(\bar{x}) = \begin{pmatrix} \bar{x}_2^1 \\ \bar{x}_3^1 \\ \bar{x}_2^1 + y_2 \end{pmatrix}, \\
\dot{\bar{x}}^2 &= \bar{f}^2(\bar{x}) = \begin{pmatrix} \bar{x}_2^2\bar{x}_3^1 \\ \bar{x}_3^2(\bar{x}_3^1)^2 \\ \bar{x}_3^2\bar{x}_3^1 + \bar{x}_2^2\bar{x}_3^1 \end{pmatrix}, \\
y &= \begin{pmatrix} \bar{h}_1(\bar{x}) \\ \bar{h}_2(\bar{x}) \end{pmatrix} = \begin{pmatrix} \bar{x}_1^1 \\ \bar{x}_1^2 \end{pmatrix}.
\end{aligned} \tag{3.7}$$

The first subsystem is observable and LTI. Hence, the transformation to a BTOF is straightforward. We attempt to transform the second subsystem. As in example system (3.6), we use the incorrect definition of the starting vector to get

$$\bar{g}^2 = \frac{1}{(\bar{x}_3^1)^3}\frac{\partial}{\partial \bar{x}_3^2}.$$

Next we verify that Condition (ii) with $\bar{F}^2 = \bar{f}^2$ and $\bar{g}^2$ holds. Hence, we falsely conclude the system admits a BTOF. However, we can show by direct computation that in fact no transformation exists. For the second subsystem to be in BTOF, we require $z_1^2 = \bar{x}_1^2$ and have the form of the other two coordinates

$$\begin{aligned}
z_2^2 &= \dot{z}_1^2 - \gamma_1^2(\bar{x}^1, \bar{x}_1^2) = \bar{x}_2^2\bar{x}_3^1 - \gamma_1^2(\bar{x}^1, \bar{x}_1^2), \\
z_3^2 &= \dot{z}_2^2 - \gamma_2^2(\bar{x}^1, \bar{x}_1^2) \\
&= \bar{x}_3^2(\bar{x}_3^1)^3 - \bar{x}_2^1\frac{\partial \gamma_1^2}{\partial \bar{x}_1^1} - \bar{x}_3^1\frac{\partial \gamma_1^2}{\partial \bar{x}_2^1} + \bar{x}_2^1\bar{x}_2^2 - \bar{x}_2^1\frac{\partial \gamma_1^2}{\partial \bar{x}_3^1} + \bar{x}_2^2\bar{x}_1^2 \\
&\quad - \bar{x}_1^2\frac{\partial \gamma_1^2}{\partial \bar{x}_3^1} - \bar{x}_2^2\bar{x}_3^1\frac{\partial \gamma_1^2}{\partial \bar{x}_1^2} - \gamma_2^2(\bar{x}^1, \bar{x}_1^2).
\end{aligned}$$

Local diffeomorphism $\Phi^2 = (z_1^2, z_2^2, z_3^2)^T$ puts the second subsystem in the form

$$\begin{aligned}
\dot{z}_1^2 &= z_2^2 + \gamma_1^2, \\
\dot{z}_2^2 &= z_3^2 + \gamma_2^2, \\
\dot{z}_3^2 &= \omega(\bar{x}^1, z_1^2).
\end{aligned}$$

The fact that a BTOF structure requires

$$\frac{\partial \omega}{\partial z_2^2} = \frac{\partial \omega}{\partial z_3^2} = 0 \tag{3.8}$$

provides two PDEs which define γ_1^2, γ_2^2. Using Maple, a solution for γ_1^2, γ_2^2 is

$$\begin{aligned}
\gamma_1^2 &= \frac{2(z_1^2)^2}{\bar{x}_3^1} + \left(\bar{x}_3^1 + \frac{4\bar{x}_2^1}{\bar{x}_3^1}\right) z_1^2, \\
\gamma_2^2 &= \left[\frac{4(\bar{x}_2^1)^2}{(\bar{x}_3^1)^2} - 3\bar{x}_2^1 - 7 + (\bar{x}_3^1)^3 - \frac{6z_2^2}{\bar{x}_3^1}\right] z_1^2 - 2\frac{(z_1^2)^3 + \frac{1}{2}\left[5(\bar{x}_3^1)^2 + 6\bar{x}_2^1\right](z_1^2)^2}{(\bar{x}_3^1)^2}.
\end{aligned}$$

Since γ_2^2 must depend on z_2^2 we cannot transform system (3.7) into a BTOF. As expected, we can verify that the conditions in Theorem 3.1.4 are not satisfied. We use (3.2) to obtain the starting vector

$$\bar{g}^2 = \frac{1}{(\bar{x}_3^1)^3}\frac{\partial}{\partial \bar{x}_3^2}.$$

Condition (ii) is not satisfied for

$$\left[\mathrm{ad}_{\bar{F}^2}^2\bar{g}^2, \mathrm{ad}_{\bar{F}^2}\bar{g}^2\right] = -\frac{6}{(\bar{x}_3^1)^4}\frac{\partial}{\partial \bar{x}_3^2} \neq 0.$$

Hence, this example shows the conditions as stated in previous work [134] are not sufficient.

3.1.3 Proof of Main Result: Theorem 3.1.4

Necessity: Assume there exists a state transformation Φ such that the ith subsystem is put into BTOF, and the first $i-1$ subsystems are unchanged, i.e., Φ defined on a ν_i-dimensional state space takes the form of

$$\Phi = \begin{pmatrix} z^{<i-1>} \\ \Phi^i \end{pmatrix}.$$

This Φ transforms the first i subsystems in BTOF. Given the ith subsystem in BTOF (3.1), we denote the representation of $\bar{g}^i$ in z-coordinates as $\tilde{g}^i$ and define $\tilde{g}^i = \frac{\partial}{\partial z_{\lambda_i}^i}$. Based on the expression of the ith subsystem in BTOF and the assumption that the first $i-1$ subsystems are in BTOF, we have by direct computation

$$\mathrm{ad}_{-\tilde{F}^i}^k\tilde{g}^i = \frac{\partial}{\partial z_{\lambda_i-k}^i}, \quad 0 \leqslant k \leqslant \lambda_i - 1,$$

where $\tilde{F}^i$ is the representation of $\bar{F}^i$ in the z-coordinates. Denoting the ith output function $\bar{h}_i$ expressed in the z-coordinates as $\tilde{h}_i$, we have $\tilde{h}_i(z) = z_1^i$. Hence, we know $\tilde{g}^i$ satisfies

$$\left\langle \mathrm{d}\tilde{h}_i, \mathrm{ad}_{-\tilde{F}^i}^k\tilde{g}^i\right\rangle = \left\langle \mathrm{d}z_1^i, \frac{\partial}{\partial z_{\lambda_i-k}^i}\right\rangle = \delta_{k,\lambda_i-1}, \quad 0 \leqslant k \leqslant \lambda_i - 1. \tag{3.9}$$

From [120, Lem. 6.15], conditions (3.9) are equivalent to

$$L_{\tilde{g}^i}L_{\tilde{F}^i}^k\tilde{h}^i = \delta_{k,\lambda_i-1}, \quad 0 \leqslant k \leqslant \lambda_i - 1, \tag{3.10}$$

which imply $L_{\bar{g}^i} L^k_{\bar{F}^i} \bar{h}_i = \delta_{k,\lambda_i-1}, 0 \leqslant k \leqslant \lambda_i - 1$. We note that the starting vector $\bar{g}^i$, which is the representation of $\tilde{g}^i$ in $\bar{x}$-coordinates, is also given by $\bar{g}^i = (\Phi_*)^{-1}\tilde{g}^i$, where

$$\Phi_* = \frac{\partial \Phi}{\partial \bar{x}^i_e} = \begin{pmatrix} I_{\nu_{i-1}} & 0_{\nu_{i-1}\times\lambda_i} \\ \frac{\partial \Phi^i}{\partial z^{<i-1>}} & \frac{\partial \Phi^i}{\partial \bar{x}^i} \end{pmatrix}. \tag{3.11}$$

It is clear in the $\bar{x}$-coordinates the starting vector is written as

$$\bar{g}^i = \sum_{k=1}^{\lambda_i} \varrho^i_k(\bar{x}^i_e) \frac{\partial}{\partial \bar{x}^i_k},$$

and we have the definition of the starting vector (3.2).

As the vector fields $\mathrm{ad}^s_{-\tilde{F}^i}\tilde{g}^i, 0 \leqslant s \leqslant \lambda_i - 1$ are constant, we have

$$\left[\mathrm{ad}^s_{-\tilde{F}^i}\tilde{g}^i, \mathrm{ad}^k_{-\tilde{F}^i}\tilde{g}^i\right] = 0, \quad 0 \leqslant s, k \leqslant \lambda_i - 1.$$

Also, from the expression of Φ_* in (3.11), we know its inverse is given by

$$(\Phi_*)^{-1} = \begin{pmatrix} I_{\nu_{i-1}} & 0_{\nu_{i-1}\times\lambda_i} \\ * & (\frac{\partial \Phi^i}{\partial \bar{x}^i})^{-1} \end{pmatrix}.$$

Hence, expressed in the $\bar{x}$-coordinates the vector fields $\mathrm{ad}^k_{-\tilde{F}^i}\bar{g}^i$, $1 \leqslant k \leqslant \lambda_i - 1$ are of the form $(0_{1\times\nu_{i-1}}, \varrho^i_{1\times\lambda_i})^T$, with the components of $\varrho^i_{1\times\lambda_i}$ any arbitrary functions depending on $\bar{x}^i_e$. The first ν_{i-1} components of $\mathrm{ad}^k_{-\bar{F}^i}\bar{g}^i$, $1 \leqslant k \leqslant \lambda_i - 1$ are zero. Similarly, as $\mathrm{ad}^s_{-\tilde{F}^i}\tilde{g}^i, 1 \leqslant s \leqslant \lambda_i - 1$ are constant, we have

$$\frac{\partial}{\partial y_j}\mathrm{ad}^r_{-\tilde{F}^i}\tilde{g}^i = 0, \quad 0 \leqslant r \leqslant \lambda_i - 1; i+1 \leqslant j \leqslant p.$$

Conditions (ii) and (iii) in Theorem 3.1.4 are satisfied in the z-coordinates. These conditions hold in $\bar{x}^i_e$-coordinates because they are invariant under change of coordinates which is independent of y_j. In addition, from the expression of $\tilde{h}_i$, we have

$$\frac{\partial \tilde{h}_i}{\partial \bar{x}^j_k} = 0, \quad 1 \leqslant k \leqslant \lambda_j; i+1 \leqslant j \leqslant p.$$

After a transformation Φ^{-1} depending on $z^{<i>}$, we have Condition (iv). Therefore, all conditions in Theorem 3.1.4 hold and necessity is proven.

Sufficiency: Given definition (3.2), local observability ensures the solvability of the starting vector $\bar{g}^i$ in $\bar{x}$-coordinates. Since

$$\begin{pmatrix} \mathrm{d}\bar{h}_i \\ \vdots \\ \mathrm{d}L^{\lambda_i-1}_{\bar{F}^i}\bar{h}_i \end{pmatrix} \left(\mathrm{ad}^{\lambda_i-1}_{-\bar{F}^i}\bar{g}^i, \ldots, \bar{g}^i\right) = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ * & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \ldots & 1 & 0 \\ * & * & \ldots & * & 1 \end{pmatrix} \in \mathbb{R}^{\lambda_i\times\lambda_i},$$

the vector fields $\mathrm{ad}^k_{-\tilde{F}^i}\tilde{g}^i, 0 \leqslant k \leqslant \lambda_i - 1$ are therefore linearly independent. Given Conditions (iii) and (iv) are satisfied, these vector fields are independent of $y_j, \bar{x}^j_k, 1 \leqslant k \leqslant \lambda_j, i+1 \leqslant j \leqslant p$. If Condition (ii) holds, the Simultaneous Rectification Theorem [120, Th. 2.36] guarantees the existence of a change of coordinates Φ defined locally on $\mathbb{R}^{\nu_i}$ such that

$$\Phi_* \left(\mathrm{ad}^{\lambda_i - 1}_{-\bar{F}^i}\bar{g}^i, \ldots, \bar{g}^i\right) = \begin{pmatrix} 0_{\nu_{i-1}\times\lambda_i} \\ I_{\lambda_i} \end{pmatrix} \in \mathbb{R}^{\nu_i \times \lambda_i}. \tag{3.12}$$

Denoting the last λ_i components of Φ as Φ^i we have

$$\frac{\partial \Phi^i}{\partial \bar{x}^i_e}\left(\mathrm{ad}^{\lambda_i - 1}_{-\bar{F}^i}\bar{g}^i, \ldots, \bar{g}^i\right) = I_{\lambda_i}. \tag{3.13}$$

Using the conditions that the first ν_{i-1} components of vector fields $\mathrm{ad}^k_{-\tilde{F}^i}\tilde{g}^i, 0 \leqslant k \leqslant \lambda_i - 1$ are zero, we know the first ν_{i-1} components of Φ can be taken as identity and get $\Phi = ((z^{<i-1>})^T, (\Phi^i)^T)^T$. Next we show Φ is a local diffeomorphism, keeps the first $i-1$ subsystems in BTOF, and transforms the ith subsystem into BTOF. Since Φ^i is solved from (3.13), and

$$\left(\mathrm{ad}^{\lambda_i - 1}_{-\bar{F}^i}\bar{g}^i, \ldots, \bar{g}^i\right) = \begin{pmatrix} 0_{\nu_{i-1}\times\lambda_i} \\ \varrho^i_{\lambda_i\times\lambda_i} \end{pmatrix}, \tag{3.14}$$

where the components of $\varrho^i_{\lambda_i\times\lambda_i}$ are some functions depending on $\bar{x}^i_e$. We know $\frac{\partial \Phi^i}{\partial \bar{x}^i}$ is nonsingular, i.e., Φ_* in (3.11) is nonsingular, hence Φ is a local diffeomorphism defined on some open set of $\mathbb{R}^{\nu_i}$.

Next, we show that the ith subsystem is transformed into BTOF by Φ, i.e.,

$$\begin{aligned} \tilde{f}^i(z) &= A^i z^i + \gamma^i(z^{<i-1>}, z^i_1, y_{[i+1,p]}), \\ \tilde{h}_i(z) &= C^i z^i. \end{aligned}$$

As well, we require $\bar{f}^j, \bar{h}_j, 1 \leqslant j \leqslant i-1$ remain untransformed by Φ. Because the first ν_{i-1} components of the transformation Φ are $z^j_k, 1 \leqslant k \leqslant \lambda_j, 1 \leqslant j \leqslant i-1$, and the first $i-1$ subsystems will not be affected by the transformation Φ^i through their dependence on y_i, the first $i-1$ subsystems remain in BTOF. As for the ith subsystem, we first consider how $\bar{f}^i$ transforms. Given the expression of $\tilde{F}^i$ in the z-coordinates

$$\tilde{F}^i = \sum_{j=1}^{i-1} \tilde{f}^j + \tilde{f}^i_1 \frac{\partial}{\partial z^i_1} + \cdots + \tilde{f}^i_{\lambda_i}\frac{\partial}{\partial z^i_{\lambda_i}},$$

where $\tilde{f}^k, 1 \leqslant k \leqslant i-1$ is the representation of $\bar{f}^k$ in z-coordinates, from Condition (ii), we have

$$\begin{aligned}
&\frac{\partial \tilde{f}^i_j}{\partial z^i_k} = 0, \quad 1 \leqslant j \leqslant \lambda_i; 2 \leqslant k \leqslant \lambda_i; j \neq k-1, \\
&\frac{\partial \tilde{f}^i_j}{\partial z^i_{j+1}} = 1, \quad 1 \leqslant j \leqslant \lambda_i - 1,
\end{aligned}$$

which imply

$$\begin{aligned}
\tilde{f}^i_k &= z^i_{k+1} + \gamma^i_k(z^{<i-1>}, z^i_1, y_{[i+1,p]}), \quad 1 \leqslant k \leqslant \lambda_i - 1, \\
\tilde{f}^i_{\lambda_i} &= \gamma^i_{\lambda_i}(z^{<i-1>}, z^i_1, y_{[i+1,p]}).
\end{aligned}$$

Next we consider the expression of $\tilde{h}_i$. From the definition of the starting vector, we have

$$\begin{aligned}
L_{\bar{g}^i} L^k_{\bar{F}^i} \bar{h}_i &= \left\langle \mathrm{d}\bar{h}_i, \mathrm{ad}^k_{-\bar{F}^i} \bar{g}^i \right\rangle \\
&= \left\langle \mathrm{d}\tilde{h}_i, \mathrm{ad}^k_{-\tilde{F}^i} \tilde{g}^i \right\rangle \\
&= \delta_{k,\lambda_i - 1}, \quad 0 \leqslant k \leqslant \lambda_i - 1,
\end{aligned}$$

which means

$$\mathrm{d}\tilde{h}_i = \mathrm{d}z^i_1 \mod \left\{ \mathrm{d}z^j_k, 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1 \right\},$$

and $\tilde{h}_i = z^i_1 + \kappa(z^{<i-1>})$ for some smooth function κ. Let

$$\bar{\Phi}^i_1 = z^i_1 + \kappa(z^{<i-1>})$$

and $\bar{z}^i_1 = \bar{\Phi}^i_1$ be the first component of Φ^i. With this definition of Φ^i, PDEs (3.13) are still satisfied. With $\bar{\Phi}^i_1$ we have $\tilde{h}_i = z^i_1$, and

$$\begin{aligned}
\dot{\bar{z}}^i_1 &= \dot{z}^i_1 + \sum_{j=1}^{i-1} \frac{\partial \kappa(z^{<i-1>})}{\partial z^j} \bar{f}^j(z^{<j>}, y_{[j+1,p]}) \\
&= z^i_2 + \gamma^i_1(z^{<i-1>}, z^i_1, y_{[j+1,p]}) + \sum_{j=1}^{i-1} \frac{\partial \kappa(z^{<i-1>})}{\partial z^j} \bar{f}^j(z^{<j>}, y_{[j+1,p]}) \\
&= z^i_2 + \bar{\gamma}^i_1(z^{<i-1>}, z^i_1, y_{[j+1,p]}).
\end{aligned}$$

Therefore, $\bar{f}^j, 1 \leqslant j \leqslant i-1$ remains in BTOF and the ith subsystem is transformed into BTOF. This concludes the proof of sufficiency. ■

Remark 3.1.10. We show that if $\bar{h}_i(\bar{x}) = \bar{x}_1^i$, Condition (i) of Theorem 3.1.4 is satisfied. From (3.2), $\bar{g}^i$ is solved from the following equation

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathrm{d}\bar{h}_i \\ \vdots \\ \mathrm{d}L_{\bar{F}^i}^{\lambda_i-2}\bar{h}_i \\ \mathrm{d}L_{\bar{F}^i}^{\lambda_i-1}\bar{h}_i \end{pmatrix} \bar{g}^i = \left(\begin{array}{cccc|cccc} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ * & * & \cdots & * & * & * & \cdots & * \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ * & * & \cdots & * & * & * & \cdots & * \end{array}\right) \bar{g}^i,$$

and the starting vector is

$$\bar{g}^i = \begin{pmatrix} 0_{1\times\nu_{i-1}} & 0 & * & \cdots & * \end{pmatrix}^T, \tag{3.15}$$

or the first $\nu_{i-1}+1$ components have to be zero (since $y_i = \bar{x}_1^i$ implies $\mathrm{d}\bar{h}_i = \mathrm{d}\bar{x}_1^i$). Based on (3.15), we find the special structure of vector fields $\mathrm{ad}_{-\bar{F}^i}^k\bar{g}^i$, $1 \leqslant k \leqslant \lambda_i - 2$. Using induction we have for $k = 1$

$$\begin{aligned} \mathrm{ad}_{-\bar{F}^i}\bar{g}^i &= \begin{pmatrix} * & * & 0_{\nu_{i-1}\times\lambda_i-1} \\ \frac{\partial \bar{f}_1^i}{\partial z^{<i-1>}} & \frac{\partial \bar{f}_1^i}{\partial \bar{x}_1^i} & \frac{\partial \bar{f}_1^i}{\partial \bar{x}_{[2,\lambda_i]}^i} \\ * & * & * \end{pmatrix} \bar{g}^i - \begin{pmatrix} 0_{\nu_{i-1}\times\nu_i} \\ 0_{1\times\nu_i} \\ * \end{pmatrix} \bar{F}^i \\ &= \begin{pmatrix} 0_{\nu_{i-1}\times 1} \\ \langle \mathrm{d}L_{\bar{F}^i}\bar{h}_i, \bar{g}^i \rangle \\ * \end{pmatrix} - \begin{pmatrix} 0_{\nu_{i-1}\times 1} \\ 0 \\ * \end{pmatrix} \\ &= \begin{pmatrix} 0_{\nu_{i-1}\times 1} \\ 0 \\ * \end{pmatrix}, \end{aligned}$$

where $\bar{x}_{[2,\lambda_i]}^i = (\bar{x}_2^i, \ldots, \bar{x}_{\lambda_i}^i)^T$ and $\langle \mathrm{d}L_{\bar{F}^i}\bar{h}_i, \bar{g}^i \rangle = 0$. Hence, $\mathrm{ad}_{-\bar{F}^i}\bar{g}^i$ has a special structure, i.e., its first $\nu_{i-1}+1$ components have to be zero. Assuming $\mathrm{ad}_{-\bar{F}^i}^k\bar{g}^i$ has this special structure, we show $\mathrm{ad}_{-\bar{F}^i}^{k+1}\bar{g}^i, 2 \leqslant k \leqslant \lambda_i - 3$ does as well.

$$\begin{aligned} \mathrm{ad}_{-\bar{F}^i}^{k+1}\bar{g}^i &= \begin{pmatrix} * & * & 0_{\nu_{i-1}\times\lambda_i-1} \\ \frac{\partial \bar{f}_1^i}{\partial z^{<i-1>}} & \frac{\partial \bar{f}_1^i}{\partial \bar{x}_1^i} & \frac{\partial \bar{f}_1^i}{\partial \bar{x}_{[2,\lambda_i]}^i} \\ * & * & * \end{pmatrix} \mathrm{ad}_{-\bar{F}^i}^k\bar{g}^i - \begin{pmatrix} 0_{\nu_{i-1}\times\nu_i} \\ 0_{1\times\nu_i} \\ * \end{pmatrix} \bar{F}^i \\ &= \begin{pmatrix} 0_{\nu_{i-1}\times 1} \\ \left\langle \mathrm{d}L_{\bar{F}^i}\bar{h}_i, \mathrm{ad}_{-\bar{F}^i}^k\bar{g}^i \right\rangle \\ * \end{pmatrix} \\ &= \begin{cases} \begin{pmatrix} 0_{1\times\nu_{i-1}} & 0 & * \end{pmatrix}^T, & 1 \leqslant k \leqslant \lambda_i - 3, \\ \begin{pmatrix} 0_{1\times\nu_{i-1}} & 1 & * \end{pmatrix}^T, & k = \lambda_i - 2, \end{cases} \end{aligned}$$

where

$$\left\langle \mathrm{d}L_{\bar{F}^i}\bar{h}_i, \mathrm{ad}_{-F^i}^k\bar{g}^i \right\rangle = \delta_{k,\lambda_i-2}.$$

Hence, the first $\nu_{i-1}+1$ components of $\mathrm{ad}^{k+1}_{-\bar{F}^i}\bar{g}^i$, $0 \leqslant k \leqslant \lambda_i - 2$ are zero. According to the definition of $\bar{g}^i$, we know only the ν_{i-1}th component of $\mathrm{ad}^{\lambda_i-1}_{-\bar{F}^i}\bar{g}^i$ is non-zero for

$$\left\langle \mathrm{d}L_{\bar{F}^i}\bar{h}_i, \mathrm{ad}^{\lambda_i-2}_{-\bar{F}^i}\bar{g}^i \right\rangle = L_{\bar{g}^i}L^{\lambda_i-1}_{\bar{F}^i}\bar{h}_i = 1.$$

Finally, we have (3.14) and Condition (i) of Theorem 3.1.4 is satisfied. □

3.2 Observer and Error Dynamics Stability

We consider the following observer structure in BTOF coordinates

$$\dot{\hat{z}}^i = A^i\hat{z}^i + \hat{\gamma}^i + L^i(y_i - C^i\hat{z}^i), \quad 1 \leqslant i \leqslant p, \tag{3.16}$$

where we have simplified notation with $\hat{\gamma}^k = \gamma^k(\hat{z}^{<k-1>}, z_1^k, y_{[k+1,p]}), 1 \leqslant k \leqslant p$, and $L^k \in \mathbb{R}^{\lambda_k}, 1 \leqslant k \leqslant p$ are constant observer gains to be designed below. The error dynamics for (3.16) is

$$\dot{\tilde{z}}^i = (A^i - L^iC^i)\tilde{z}^i + \gamma^i - \hat{\gamma}^i, \quad 1 \leqslant i \leqslant p, \tag{3.17}$$

where the state estimate errors are $\tilde{z}^k = z^k - \hat{z}^k, 1 \leqslant k \leqslant p$, and

$$\gamma^k = \gamma^k(z^{<k-1>}, z_1^k, y_{[k+1,p]}), \quad 1 \leqslant k \leqslant p.$$

3.2.1 Global Exponential Stability of Error Dynamics

Theorem 3.2.1. *Assume a BTOF exists globally for* (2.1) *and* $\gamma^i, 1 \leqslant i \leqslant p$ *are globally Lipschitz in* $z^{<i-1>}$ *uniformly in* $y_{[i,p]}$ *w.r.t. any norm, i.e., there exist constants* $M_k^i, 1 \leqslant k \leqslant i-1$ *such that*

$$\left\|\gamma^i - \hat{\gamma}^i\right\| \leqslant \sum_{k=1}^{i-1} M_k^i \left\|\tilde{z}^k\right\|, \quad \forall y_{[i,p]} \in \mathbb{R}^{p-i+1}, \forall z^{<i-1>}, \hat{z}^{<i-1>} \in \mathbb{R}^{\nu_{i-1}}. \tag{3.18}$$

Then provided the spectrum of $A^i - L^iC^i, 1 \leqslant i \leqslant p$ *lies in* $\mathbb{C}^- = \{s \in \mathbf{C} : \Re(s) < 0\}$, *the zero solution of* (3.17) *is Globally Exponentially Stable (GES).*

Proof: We first show that the stability for the error dynamics of the first two subsystems (3.17)

$$\dot{\tilde{z}}^1 = (A^1 - L^1C^1)\tilde{z}^1, \tag{3.19a}$$

$$\dot{\tilde{z}}^2 = (A^2 - L^2C^2)\tilde{z}^2 + \gamma^2 - \hat{\gamma}^2, \tag{3.19b}$$

then induction is applied to show stability for the entire error dynamics. For the first subsystem the error dynamics (3.19a) is LTI, thus given (A^1, C^1) observable there exists L^1 such that the zero solution of the error dynamics is GES. For the second subsystem, its error dynamics (3.19b) consists of a LTI part and a "disturbance" term $\gamma^2 - \hat{\gamma}^2$. Denoting $u^2(t) = \gamma^2 - \hat{\gamma}^2$, (3.19b) is written

$$\dot{\tilde{z}}^2 = (A^2 - L^2C^2)\tilde{z}^2 + B^2u^2(t), \tag{3.20}$$

where $B^2 = I_{\lambda_2}$. Provided (3.18) holds and $\tilde{z}^1$ converges exponentially to zero, $u^2(t)$ converges exponentially to zero. Also, with (A^2, C^2) observable, L^2 can be chosen to make the LTI part of (3.20) exponentially convergent. Therefore, according to [104, Lem. III.1] we know the zero solution of the error dynamics of the second subsystem is GES. For the ith, $3 \leqslant i \leqslant p$, subsystem, after $u^i(t), B^i, L^i$ are similarly chosen, [104, Lem. III.1] can be reapplied to show the zero solution of the error dynamics of the ith subsystem is GES. Therefore, the zero solution of the error dynamics of the first i subsystems is GES, and by induction we have shown the zero solution of the entire system's error dynamics is GES. □

Remark 3.2.2. A global Lipschitz property is important to ensure that the zero solution is GES. We consider a simple quadratic scalar system example

$$\dot{\tilde{x}} = -\mu\tilde{x} + \tilde{x}^2u,$$

where $\tilde{x} \in \mathbb{R}, u = e^{-kt}$, and $\mu > 0$ is constant. The system is not globally Lipschitz in $\tilde{x}, u$. As shown in [87], using the change of dependent variable $w = 1/\tilde{x}$ gives the solution of $\tilde{x}(t)$

$$\tilde{x}(t) = \frac{(\mu + k)\tilde{x}_0}{(\mu + k - \tilde{x}_0)e^{kt} + \tilde{x}_0e^{-kt}},$$

where $\tilde{x}_0 = \tilde{x}(0)$. The solution $\tilde{x}(t)$ has a finite escape time when the initial condition satisfies $\tilde{x}_0 \geqslant \mu + k$. □

Remark 3.2.3. The global Lipschitz condition on γ^i is not always necessary. We consider the Jet Engine Example [87, Sec. 2.4]

$$\begin{pmatrix} \dot{R} \\ \dot{\phi} \\ \dot{\psi} \end{pmatrix} = \begin{pmatrix} -\sigma R^2 - \sigma R(2\phi + \phi^2) \\ -\psi - \frac{3}{2}\phi^2 - \frac{1}{2}\phi^3 - 3R\phi - 3R \\ -u \end{pmatrix},$$

where $R \geqslant 0$ is the normalized stall cell squared amplitude, ϕ is the mass flow, ψ is the pressure rise, σ is a constant positive parameter, and u is the control input.

The homogenous part of the R subsystem is globally asymptotically stable (GAS). It is shown in [87] R has a GAS equilibrium at $R = 0$ if ϕ decays to 0 as $t \to \infty$. However, the nonlinear term $\sigma R(2\phi + \phi^2)$ is not globally Lipschitz in ϕ. A detailed proof in [87] reveals that Input-to-State Stability (ISS) of the R-subsystem w.r.t. ϕ leads to GAS. Realizing the Lipschitz condition on γ is imposed to make the error dynamics (3.17) ISS with respect to $\tilde{z}^k, 1 \leqslant k \leqslant i-1$ which appears in $\gamma^i - \hat{\gamma}^i$ [76, Lem. 4.6], we can provide a GAS version of Theorem 3.2.1. □

Theorem 3.2.4. *Assume that a BTOF exists globally for* (2.1) *and that the error dynamics* (3.17) *is ISS w.r.t.* $\tilde{z}^k, 1 \leqslant k \leqslant i-1$. *Then provided the spectrum of* $A^i - L^iC^i, 1 \leqslant i \leqslant p$ *lies in* $\mathbb{C}^-$, *the zero solution of* (3.17) *is GAS.*

Remark 3.2.5. The proof of Theorem 3.2.4 is omitted since it is result from a direct application of [76, Lem. 4.7]. Although Theorem 3.2.4 is more general than Theorem 3.2.1, the latter is preferable in some cases since GES is a stronger form of convergence and systematically verifying ISS in practice can be challenging. A sufficient condition for ISS is given in [76, Lem. 4.6]. □

Remark 3.2.6. Another choice of observer design is possible for systems in BTOF. If we take the observer of the ith subsystem as

$$\dot{\hat{z}}^i = A^i\hat{z}^i + \gamma^i(\hat{z}^{<i-1>}, y) - L^i(y_i - C^i\hat{z}^i), \tag{3.21}$$

where we have defined

$$\gamma^i(\hat{z}^{<i-1>}, y) = \gamma^i(z_1^1, \hat{z}^1_{[2,\lambda_1]}, \ldots, z_1^{i-1}, \hat{z}^{i-1}_{[2,\lambda_{i-1}]}, z_1^i, y_{[i+1,p]}),$$
$$\hat{z}^k_{[2,\lambda_k]} = (\hat{z}^k_2, \cdots, \hat{z}^k_{\lambda_k})^T, \quad 1 \leqslant k \leqslant i-1,$$

the error dynamics of the ith subsystem is written as

$$\dot{\tilde{z}}^i = (A^i - L^iC^i)\tilde{z}^i + \gamma^i(z^{<i-1>}, y) - \gamma^i(\hat{z}^{<i-1>}, y).$$

With observer design (3.21), we have a similar stability result to Theorem 3.2.1. However, for (3.21) we only require $\gamma^i, 1 \leqslant i \leqslant p$, to be globally Lipschitz in $z^k_{[2,\lambda_k]} = (z^k_2, \ldots, z^k_{\lambda_k})^T, 1 \leqslant k \leqslant i-1$ uniformly in $y \in \mathbb{R}^p$ w.r.t. any norm, i.e., there exist constants $M^i_k, 1 \leqslant k \leqslant i-1$ such that

$$\left\|\gamma^i(z^{<i-1>}, y) - \gamma^i(\hat{z}^{<i-1>}, y)\right\| \leqslant \sum_{k=1}^{i-1} M^i_k \left\|\tilde{z}^k_{[2,\lambda_k]}\right\|,$$

for $\forall y \in \mathbb{R}^p, \forall \tilde{z}^k_{[2,\lambda_k]} \in \mathbb{R}^{\lambda_k - 1}, 1 \leqslant k \leqslant i-1$. □

Remark 3.2.7. Many practical systems have outputs which lie in some compact set $\mathcal{Y}$. Hence, we only require the Lipschitz condition (3.18) to hold uniformly for all $y_{[i,p]} \in \mathcal{Y} \subset \mathbb{R}^{p-i+1}$. If the system state belongs to a compact set and the Lipschitz condition (3.18) does not hold, convergence cannot be proven as in Theorems 3.2.1 or 3.2.4 due to the observer state leaving the compact set. In this situation, an observer design with semi-globally stable error dynamics is therefore preferred. $\square$

3.2.2 Semi-global Stability of Error Dynamics

To ensure a GES result we require $\gamma^i, 1 \leqslant i \leqslant p$ to be globally Lipschitz. This condition can be restrictive and unnecessary in practice. In fact, the state of many physical systems belongs to $\mathcal{D} \subset \mathbb{R}^n$, where $\mathcal{D}$ is a connected compact subset of $\mathbb{R}^n$ [46]. The set $\mathcal{D}$ is mapped to another connected compact set denoted $\bar{\mathcal{D}} \subset \mathbb{R}^n$ in BTOF coordinates. As discussed in [46, 141], we can extend the dynamics in BTOF coordinates from $\bar{\mathcal{D}}$ to $\mathbb{R}^n$ to obtain a semi-global stability result. We choose a particular Lipschitz extension technique, proposed in [141] and described by the following Lemma.

Lemma 3.2.8. *[141, Lem. 2] Consider a C^1 function $\chi : \mathcal{M} \times \mathbb{R}^r \mapsto \mathbb{R}$ where*

$$\mathcal{M} = \{x \in \mathbb{R}^q : \ |x_i| \leqslant \rho_i, 1 \leqslant i \leqslant q, \rho_i > 0\}.$$

Then $\chi(\sigma(x), y)$ is C^0 on $\mathbb{R}^q \times \mathbb{R}^r$ and equal to $\chi(x, y)$ for all $x \in \mathcal{M}$, and there exists a bounded function $M(y)$ such that

$$|\chi(\sigma(x), y) - \chi(\sigma(\hat{x}), y)| \leqslant M(y) \|x - \hat{x}\|, \quad \forall x, \hat{x} \in \mathbb{R}^q, \ \forall y \in \mathbb{R}^r, \tag{3.22}$$

where $\sigma(x)$ is an element-wise saturation function which is saturated outside $\mathcal{M}$.

We remark that if function χ in Lemma 3.2.8 is defined on $\mathcal{M} \times \mathcal{Y}$ where $\mathcal{Y}$ is a bounded compact subset of $\mathbb{R}^r$, we can extend the definition of χ to $\mathbb{R}^q \times \mathcal{Y}$ in a similar way. Considering the observer design (3.16) for the ith subsystem, $\gamma^i(z^{<i-1>}, z_1^i, y_{[i+1,p]})$ is defined on $\mathcal{U}^i \times \mathcal{Y}^i \subset \bar{\mathcal{D}}$, where

$$\mathcal{U}^i = \{\varsigma : \varsigma = z^{<i-1>} \in \bar{\mathcal{D}}\},$$
$$\mathcal{Y}^i = \{\varsigma : \varsigma = y_{[i,p]} \in \bar{\mathcal{D}}\}.$$

We need to extend the definition of γ^i from $\mathcal{U}^i \times \mathcal{Y}^i$ to $\mathbb{R}^{\nu_{i-1}} \times \mathcal{Y}^i$, $1 \leqslant i \leqslant p$, to design a semi-global observer. According to Lemma 3.2.8, the Lipschitz extension of

$\gamma^i(z^{<i-1>}, z_1^i, y_{[i+1,p]})$, $1 \leqslant i \leqslant p$, could be constructed as $\gamma^i(\sigma(z^{<i-1>}), z_1^i, y_{[i+1,p]})$. The system after extension is

$$\begin{aligned} \dot{z}^i &= A^i z^i + \gamma^i(\sigma(z^{<i-1>}), z_1^i, y_{[i+1,p]}), \\ y_i &= C^i z^i, \end{aligned} \qquad 1 \leqslant i \leqslant p, \tag{3.23}$$

and the corresponding semi-global observer and error dynamics system are

$$\dot{\hat{z}}^i = A^i \hat{z}^i + \gamma^i(\sigma(\hat{z}^{<i-1>}), z_1^i, y_{[i+1,p]}) + L^i(y_i - C^i \hat{z}^i), \tag{3.24a}$$

$$\begin{aligned} \dot{\tilde{z}}^i &= (A^i - L^i C^i)\tilde{z}^i \\ &\quad + \left[\gamma^i(\sigma(z^{<i-1>}), z_1^i, y_{[i+1,p]}) - \gamma^i(\sigma(\hat{z}^{<i-1>}), z_1^i, y_{[i+1,p]})\right], \end{aligned} \tag{3.24b}$$

for $1 \leqslant i \leqslant p$. Using the same method of proof as in Theorem 3.2.1, we can show the following result.

Theorem 3.2.9. *Consider system* (2.1) *whose state lies in a connected compact set $\mathcal{D}$, and assume the system is transformable to BTOF on some set enclosing $\mathcal{D}$. Denote the set $\bar{\mathcal{D}}$ as the image of $\mathcal{D}$ in BTOF coordinates, and assume $\gamma^i, 1 \leqslant i \leqslant p$, is Lipschitz in $\mathcal{U}^i$ uniformly in $(z_1^i, y_{[i+1,p]}) \in \mathcal{Y}^i$. The Lipschitz extension* (3.23) *and observer* (3.24a) *yield error dynamics* (3.24b). *Provided the spectrum of $A^i - L^i C^i, 1 \leqslant i \leqslant p$ lies in $\mathbb{C}^-$, the zero solution of* (3.24b) *is GES.*

Remark 3.2.10. The stability results are stated in the BTOF coordinates. If the inverse map $\zeta = \Theta(z)$ is Lipschitz on $\bar{\mathcal{D}}$, the error dynamics (3.24b) are GES in the original coordinates. □

Remark 3.2.11. Based on observer (3.21), another semi-global observer and semi-global stability result can be derived if we redefine

$$\begin{aligned} \mathcal{U}^i &= \{\varsigma : \varsigma = ((z^1_{[2,\lambda_1]})^T, \cdots, (z^{i-1}_{[2,\lambda_{i-1}]})^T)^T\}, \\ \mathcal{Y}^i &= \{\varsigma : \varsigma = (y_1, \cdots, y_p)^T\}, \end{aligned}$$

and extend the definition of γ^i from $\mathcal{U}^i \times \mathcal{Y}^i$ to $\mathbb{R}^{\nu_{i-1}-i+1} \times \mathcal{Y}^i$. □

3.3 BTOF Existence Conditions II

The previous section gives the existence conditions using a decentralized approach, in which a system is decomposed into several observable subsystems based on the

p-tuple of observability indices. Each subsystem is then treated separately and transformed into a BTF by an extended state transformation. Finally each subsystem is sequentially transformed into a BTOF. As indicated above, Assumption 3.1.2 is imposed to simplify the presentation of the existence conditions and its elimination is therefore desirable to enlarge the class of applicable systems.

3.3.1 An Illustrative Example

To demonstrate the challenge in removing the assumption, we consider a two-output system with indices $(2, 2)$

$$\begin{aligned} \dot{z}^1 &= A^1 z^1 + \gamma^1(y), \\ \dot{z}^2 &= A^2 z^2 + \gamma^2(z^1, z_1^2), \\ y &= \begin{pmatrix} z_1^1 \\ z_1^2 \end{pmatrix}. \end{aligned} \tag{3.25}$$

The system is in a BTOF. We discuss the necessary conditions of a BTOF, and explore the differences from the OF case. The Simultaneous Rectification Theorem [120, Th. 2.36] is applied to derive the BTOF existence conditions. We begin by defining the unit vector fields to be rectified

$$\eta_1^1 = \frac{\partial}{\partial z_1^1}, \quad \eta_2^1 = \frac{\partial}{\partial z_2^1}, \quad \eta_1^2 = \frac{\partial}{\partial z_1^2}, \quad \eta_2^2 = \frac{\partial}{\partial z_2^2}.$$

Starting vectors are required to construct these unit vectors. For the second subsystem, we take $g^2 = \eta_2^2$ as the starting vector and have

$$\eta_1^2 = \mathrm{ad}_{-f} g^2.$$

Taking the starting vector for the first subsystem as $g^1 = \eta_2^1$, we have

$$\begin{aligned} \eta_1^1 &= \mathrm{ad}_{-f} g^1 - \frac{\partial \gamma_1^2}{\partial z_2^1} \frac{\partial}{\partial z_1^2} - \frac{\partial \gamma_2^2}{\partial z_2^1} \frac{\partial}{\partial z_2^2} \\ &= \mathrm{ad}_{-f} g^1 - \frac{\partial \gamma_1^2}{\partial z_2^1} \eta_1^2 - \frac{\partial \gamma_2^2}{\partial z_2^1} \eta_2^2. \end{aligned}$$

We notice the vector field η_1^1 depends on unknown functions γ_1^2, γ_2^2, i.e., unit vectors are not merely determined by system dynamics f and output mapping h. This leads to the difficulty in verifying the new existence conditions given later. To compute the starting vectors, it is straightforward to verify $g^2 = \eta_2^2$ is defined by

$$L_{g^2} L_f^k h_i = \delta_{i,2} \delta_{k,1}, \quad 0 \leqslant k \leqslant 1; 1 \leqslant i \leqslant 2.$$

This is equivalent to

$$\left\langle \mathrm{d}h_i, \mathrm{ad}^k_{-f} g^2 \right\rangle = \delta_{i,2}\delta_{k,1}.$$

We note the definition of g^2 is the same as that in the OF. However, we cannot derive a similar definition of the starting vector for the first subsystem. Following the standard procedure of computing the starting vector g^1, we obtain $L_f h_1 = z_2^1 + \gamma_1^1, L_f h_2 = z_2^2 + \gamma_1^2$, and have

$$\begin{aligned} L_{g^1} h_1 = 0, &\quad L_{g^1} h_2 = 0, \\ L_{g^1} L_f h_1 = 1, &\quad L_{g^1} L_f h_2 = \frac{\partial \gamma_1^2}{\partial z_2^1} \neq 0, \end{aligned}$$

which implies the starting vector depends on the unknown function γ_1^2. Without the knowledge of the starting vectors, it is difficult to construct the vector fields to be rectified. Therefore, it is desirable to develop a new definition for the starting vectors which can be solved directly.

Relative to an OF, the extra freedom introduced by a BTOF is evident. For a two-output system with indices $(2,2)$ to be transformed into an OF, the necessary conditions are

$$\left[\mathrm{ad}^k_{-f} g^l, \mathrm{ad}^s_{-f} g^j\right] = 0, \quad 0 \leqslant k, s \leqslant 1; 1 \leqslant l, j \leqslant 2.$$

On the other hand, transforming to a BTOF does not require

$$\left[\mathrm{ad}^k_{-f} g^1, g^1\right] = 0. \tag{3.26}$$

Instead, this condition is replaced by

$$\left[\mathrm{ad}^k_{-f} g^1 - L_{g^1}\gamma_1^2 \mathrm{ad}_{-f} g^2 - L_{g^1}\gamma_2^2 g^2, g^1\right] = 0. \tag{3.27}$$

Allowing the dependence of functions γ_1^2, γ_2^2 on z_2^1 makes (3.27) easier to satisfy.

3.3.2 Construction of Vectors

Starting vectors The previous example shows how the derivation of the starting vectors in the BTOF coordinates is not straightforward. However, we can simplify the derivation using a *parameter assumption.* That is, when calculating $L_f^k h_l, 0 \leqslant k \leqslant \lambda_i - 1, 1 \leqslant l \leqslant i$ for the ith subsystem, the outputs of lower subsystems

$y_{[i+1,p]}$ are treated as parameters. This new assumption can be made without loss of generality. Imposing the assumption we have

$$\begin{aligned} L_f h_l &= z_2^l + \gamma_1^l, \\ &\vdots \qquad\qquad\qquad\qquad\qquad 1 \leqslant l \leqslant i, \\ L_f^{\lambda_i - 1} h_l &= z_{\lambda_i}^l + \varrho_l(z^{<i-1>}, z^i_{[1,\lambda_i-1]}, y), \end{aligned}$$

and $L_f^k h_l = 0$ for $0 \leqslant k \leqslant \lambda_i - 1, i < l \leqslant p$. Hence, the starting vector $g^i = \frac{\partial}{\partial z^i_{\lambda_i}}$ follows its traditional definition [165, 109]

$$L_{g^i} L_f^k h_l = \delta_{i,l}\delta_{k,\lambda_i-1}, \quad 0 \leqslant k \leqslant \lambda_i - 1; 1 \leqslant l \leqslant p. \tag{3.28}$$

It is worth noting that in the BTOF case, g^i is solved from $i\lambda_i$ equations since the last $(p-i)\lambda_i$ equations are trivially satisfied, while the number of equations for g^i is $p\lambda_i$ in the OF case. There is more freedom in choosing the starting vectors for the BTOF than for the OF. Given a system in Observable Form, a typical solution of (3.28) for the starting vector g^i is

$$g^i = \frac{\partial}{\partial x^i_{\lambda_i}} + \sum_{l=i+1}^{p} \sum_{k=1}^{\lambda_l} \varrho_k^l(x) \frac{\partial}{\partial x_k^l}, \tag{3.29}$$

where $\varrho_k^l(x)$ is an arbitrary smooth function of x. The unknown functions $\varrho_k^l(x)$ is result from treating $y_{[i+1,p]}$ as parameters.

Unit vectors Next we discuss the construction of unit vectors given the starting vector g^i defined in (3.28). For the ith subsystem, we proceed as in the example system (3.25) and construct η_j^i from g^i and f

$$\begin{aligned} \eta^i_{\lambda_i} &= g^i, \\ \eta^i_s &= \mathrm{ad}_{-f}\eta^i_{s+1} - \sum_{k=i+1}^{p} \sum_{l=1}^{\lambda_k} \frac{\partial \gamma_l^k}{\partial z^i_{s+1}} \frac{\partial}{\partial z_l^k}, \quad 1 \leqslant s \leqslant \lambda_i - 1. \end{aligned} \tag{3.30}$$

When $i = p$, we have $\eta_k^p = \mathrm{ad}_{-f}^{\lambda_i - k} g^i, 1 \leqslant k \leqslant \lambda_p$. Since η_k^i has dependence on unknown functions $\gamma_k^l, 1 \leqslant k \leqslant \lambda_l, i+1 \leqslant l \leqslant p$, it is impossible for us to verify in advance whether η_k^i are unit vectors. Also, it is not clear whether $\gamma_k^i, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ solved from (3.30) ensure that the Jacobian matrix defined by

$$\Omega(\zeta) = \left(\eta_1^1, \cdots, \eta_{\lambda_1}^1, \cdots, \eta_1^p, \cdots, \eta_{\lambda_p}^p\right)$$

is nonsingular.

3.3.3 Main Result

Given the starting vectors and the unit vectors defined by (3.28) and (3.30) respectively, we state a theorem for existence of a BTOF without Assumption 3.1.2.

Theorem 3.3.1. *Let system* (2.1) *be locally observable in* $U_0 \subset \mathbb{R}^n$ *with indices* $\lambda_i, 1 \leqslant i \leqslant p$. *There exists a state transformation* $z = \Phi(\zeta)$ *defined locally such that in the new coordinates, system* (2.1) *is of the form* (3.1) *if and only if in* U_0

(i) for $1 \leqslant i \leqslant p$, *there exists a starting vector* g^i *satisfying* (3.28);

(ii) there exist functions $\gamma_k^i, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ *such that* n *vectors* $\eta_{i,r}, 1 \leqslant r \leqslant \lambda_i, 1 \leqslant i \leqslant p$ *given by* (3.30), *are linearly independent and commute:*

$$[\eta_r^i, \eta_s^j] = 0, \quad \begin{cases} 1 \leqslant r \leqslant \lambda_i; \\ 1 \leqslant s \leqslant \lambda_j; \\ 1 \leqslant i, j \leqslant p; \end{cases}$$

(iii) the functions $\gamma_k^i, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ *satisfying Condition* (ii) *are such that*

$$L_{\eta_k^i} h_l = \delta_{k,1}\delta_{l,i}, \quad 1 \leqslant k \leqslant \lambda_i; 1 \leqslant i, l \leqslant p.$$

The transformation Φ *is the solution of the* n^2 *PDEs*

$$\frac{\partial \Phi(\zeta)}{\partial \zeta} \Omega(\zeta) = I_n. \tag{3.31}$$

Remark 3.3.2. The BTOF coordinates are globally defined if the system is globally observable, the theorem conditions hold in $\mathbb{R}^n$, and the vector fields $\eta_r^i, 1 \leqslant r \leqslant \lambda_i, 1 \leqslant i \leqslant p$, are complete. □

Remark 3.3.3. The unknown functions ϱ_k^l (defined in (3.29)) and γ_j^i are restricted by Conditions (ii)–(iii). To simplify computation, the starting vectors can be typically taken as $g^i = \frac{\partial}{\partial x_{\lambda_i}^i}$ in the observable coordinates. □

Remark 3.3.4. Unlike the OF case, the vector fields $\mathrm{ad}_{-f}^r g^i, 1 \leqslant r \leqslant \lambda_i, 1 \leqslant i \leqslant p$ are linearly independent. This is guaranteed by the observability assumption on the original system and can be verified by

$$\begin{pmatrix} \mathrm{d}h_1 \\ \vdots \\ \mathrm{d}L_f^{\lambda_1 - 1} h_1 \\ \vdots \\ \mathrm{d}L_f^{\lambda_p - 1} h_p \end{pmatrix} \bar{\Omega}(\zeta) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ * & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & 1 & 0 \\ * & * & \cdots & * & 1 \end{pmatrix},$$

where

$$\bar{\Omega}(\zeta) = \left(\mathrm{ad}_{-f}^{\lambda_1 - 1} g^1, \cdots, g^1, \cdots, \mathrm{ad}_{-f}^{\lambda_p - 1} g^p, \cdots, g^p\right).$$

However, we cannot arrive at the same conclusion if $\bar{\Omega}(\zeta)$ is replaced by $\Omega(\zeta)$. □

Proof: **Necessity:** Given the system in BTOF, the starting vectors are taken as $g^i = \eta^i_{\lambda_i} = \frac{\partial}{\partial z^i_{\lambda_i}}, 1 \leqslant i \leqslant p$. It can be verified that the starting vectors satisfy (3.28) provided the outputs of the lower subsystems are treated as parameters. Furthermore, the vectors $\eta^i_k, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ can be calculated as proposed in (3.30). The vector fields $\eta^i_k, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ are unit vectors and therefore commute. Hence, Condition (ii) holds with some functions γ^i_k. On the other hand, according to the expression of h in BTOF-coordinates we conclude

$$\frac{\partial h_i}{\partial z^l_k} = \left\langle \mathrm{d}h_i, \eta^l_k \right\rangle = L_{\eta^l_k} h_i = \delta_{k,1}\delta_{i,l},$$

for $1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$. The necessity of Conditions (i)–(iii) is therefore shown.

Sufficiency: To prove the sufficiency, we need to verify the existence of the state transformation to BTOF coordinates and f, h under the new coordinates is in BTOF (3.1). According to Conditions (i)–(ii), we can construct the vector fields $\eta^i_k, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$. Since Lie bracket conditions (ii) hold and the vector fields are linearly independent, the Simultaneous Rectification Theorem ensures the existence of a state transformation $\Phi(\zeta)$ which rectifies η^i_k into a unit vector field. Induction is employed to verify the expression of f in the new coordinates. According to (3.31), the starting vector of the pth subsystem is denoted as $\frac{\partial}{\partial z^p_{\lambda_p}}$ in BTOF-coordinates. Following the standard approach in Section 3.1, it can be verified that the dynamics of the pth subsystem is written as $f^p = A^p z^p + \gamma^p(z^{<p-1>}, z^p_1)$. The dynamics of the other subsystems, denoted by $f^i, 1 \leqslant i \leqslant p-1$, have no dependence on the pth subsystem state z^p except on its output z^p_1. We assume the last $p - i + 1$ subsystems' dynamics have the BTOF structure in the transformed coordinates and check the expression for f^i in the new coordinates. Denoting the unit vectors $\eta^i_k = \frac{\partial}{\partial z^i_k}, 1 \leqslant k \leqslant \lambda_i$ and considering (3.30) we have

$$\begin{aligned}
\mathrm{ad}_{-f}\eta^i_{k+1} &= \eta^i_k - \sum_{k=i+1}^{p} \sum_{l=1}^{\lambda_k} \frac{\partial \gamma^k_l}{\partial z^i_{s+1}} \frac{\partial}{\partial z^k_l} \\
&= \frac{\partial}{\partial z^i_k} - \sum_{k=i+1}^{p} \sum_{l=1}^{\lambda_k} \frac{\partial \gamma^k_l}{\partial z^i_{s+1}} \frac{\partial}{\partial z^k_l}, \quad 1 \leqslant s \leqslant \lambda_i - 1,
\end{aligned}$$

for $1 \leqslant k \leqslant \lambda_i - 1$, which implies that

$$\frac{\partial f_s^i}{\partial z_{k+1}^i} = \begin{cases} 1, & s = k, \\ 0, & 1 \leqslant s \leqslant \lambda_i; s \neq k, \end{cases}$$
$$\frac{\partial f_s^j}{\partial z_{k+1}^i} = 0, \quad 1 \leqslant s \leqslant \lambda_j; 1 \leqslant j \leqslant i-1.$$

Similarly, we can verify that γ^i depends on $z^{<i-1>}, z_1^i, y_j = z_1^j, i+1 \leqslant j \leqslant p$, and have $f^i = A^i z^i + \gamma^i$. Enumerating i from 1 to p, we can verify the expression for f in the new coordinates has a BTOF structure. Condition (iii) can be applied to verify the expression for h in the new coordinates. □

Remark 3.3.5. Theorem 3.1.4 gives conditions which are very similar to the result we can obtain by following the generalized characteristic equation approach taken in [74]. Both are based on the construction of n functions γ_k^i. □

3.4 Observer Design Examples

We consider the observer design for a Lorenz system

$$\begin{aligned} \dot{\zeta} = f(\zeta) &= \begin{pmatrix} \sigma(\zeta_2 - \zeta_1) \\ \rho\zeta_1 - \zeta_2 - \zeta_1\zeta_3 \\ \zeta_1\zeta_2 - \delta\zeta_3 \end{pmatrix}, \\ y = h(\zeta) &= \begin{pmatrix} \zeta_1 \\ \zeta_3 \end{pmatrix}, \end{aligned} \tag{3.32}$$

where $\zeta = (\zeta_1, \zeta_2, \zeta_3)^T, y = (y_1, y_2)^T$, and σ, δ, ρ are constants. Provided that $\sigma \neq 0$, the system (3.32) is globally observable with the indices $(2, 1)$. The conditions of Theorem 3.1.4 are satisfied and the transformation of the first subsystem is a linear transformation of the observable coordinates $x_j^i = L_f^{j-1} h_i$: $\Phi^1 = (x_1^1, x_2^1 + (\sigma+1)x_1^1)^T$. The system in BTOF coordinates is

$$\begin{aligned} \dot{z} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} z + \begin{pmatrix} -(\sigma+1)z_1^1 \\ \sigma z_1^1(\rho - y_2 - 1) \\ \frac{1}{\sigma}(z_1^1 z_2^1 - (z_1^1)^2 - \sigma\delta z_1^2) \end{pmatrix} \\ &= \begin{pmatrix} A^1 & 0 \\ 0 & A^2 \end{pmatrix} z + \begin{pmatrix} \gamma^1 \\ \gamma^2 \end{pmatrix}, \\ y &= \begin{pmatrix} z_1^1 \\ z_1^2 \end{pmatrix}. \end{aligned}$$

The error dynamics for the first subsystem is LTI, and the second subsystem has error dynamics

$$\dot{\tilde{z}}_1^2 = -l_1^2 \tilde{z}_1^2 + \frac{1}{\sigma} z_1^1 \tilde{z}_2^1.$$

Applying the observer (3.21) we require the disturbance term to be globally Lipschitz in z_2^1 for GES. This condition is satisfied as z_1^1 is bounded for any initial condition. For the simulations below, the observer is implemented in the ζ-coordinates and we take $L^1 = (2, 1)^T, l_1^2 = 1$.

In order to demonstrate the usefulness of a BTOF-based approach, we compare it to a generally applicable method in [82]. We take the parameters $\sigma = 10, \delta = 8/3, \rho = 28$ and design the observer about the equilibrium

$$\zeta_{e1} = (\sqrt{\delta(\rho - 1)}, \sqrt{\delta(\rho - 1)}, \rho - 1)^T.$$

The system matrix of the linearization $F = \frac{\partial f}{\partial \zeta}(\zeta_{e1})$ has a spectrum

$$\Lambda(F) = \{0.09396 \pm 10.1945i, -13.8546\}.$$

These eigenvalues lie in the Poincaré domain and are non-resonant [5]. We choose A such that its eigenvalues are type (C, ν) w.r.t. $\Lambda(F)$. This is achieved with $A = \mathrm{diag}(-\sqrt{2}, -\sqrt{3}, -1)$. Recall, an n-tuple $\mu = (\mu_1, \ldots, \mu_n)$ of eigenvalues belongs to the *Poincaré domain* if the convex hull of the n points $(\mu_1, \ldots, \mu_n)$ in the complex plane does not contain zero; the n-tuple μ of eigenvalues is said to be resonant if among the eigenvalues there exists an integral relation of the form $\mu_s = (m, \mu)$ where $m = (m_1, \ldots, m_n), m_k \geqslant 0, \sum m_k \geqslant 2$; given an $n \times n$ matrix F with spectrum $\Lambda(F) = \mu$ and constants $C > 0, \nu > 0$, we say a complex number κ is of type (C, ν) w.r.t. $\Lambda(F)$ if for any vector m of nonnegative integers, $|m| = \sum_{i=1}^n m_i > 0$, we have

$$|\kappa - m \cdot \lambda| \geqslant \frac{C}{|m|^\nu}.$$

As the conditions of [83, Main Thm.] are satisfied, there exists a local diffeomorphism θ which transforms the system into $\dot{\theta} = A\theta - \beta(y)$. We choose β to have degree two and compute θ

$$\beta = \begin{pmatrix} -11y_1 + 40y_2 \\ 270y_1 - 10y_1y_2 \\ y_1 - 8/3y_3 - 0.9(y_1)^2 \end{pmatrix},$$

$$\theta^{[1]} = \begin{pmatrix} -0.4447 & 4.6703 & 0.2960 \\ 14.8084 & 1.94072 & -17.6194 \\ 0.3654\text{e-}1 & -0.3058 & -0.4306\text{e-}1 \end{pmatrix} \zeta,$$

$$\theta^{[2]} = \begin{pmatrix} 0.7364\text{e-}1 & 0.6010 & 0.4917 & 0.2279 & 0.1344 & -0.2271 \\ 0.7262 & 3.0887 & 2.1235 & 0.8475 & 1.4636 & -0.9013 \\ -0.1245\text{e-}1 & 0.9849\text{e-}1 & 0.5056\text{e-}1 & 0.3176\text{e-}1 & 0.7445\text{e-}1 & -0.2582\text{e-}1 \end{pmatrix} \zeta^{[2]},$$

where $\zeta^{[2]} = ((\zeta_1)^2, (\zeta_2)^2, (\zeta_3)^2, \zeta_1\zeta_2, \zeta_1\zeta_3, \zeta_2\zeta_3)^T$.

Figure 3.1 shows trajectories of the 2-norm of the error for both observers for three initial conditions

$$\begin{aligned}
&\text{IC1:} \quad \zeta(0) = (6\sqrt{2}-2, 6\sqrt{2}, 27)^T, && \hat{\zeta}(0) = (6\sqrt{2}, 6\sqrt{2}-1, 27)^T,\\
&\text{IC2:} \quad \zeta(0) = (6\sqrt{2}-3, 6\sqrt{2}, 27)^T, && \hat{\zeta}(0) = (6\sqrt{2}, 6\sqrt{2}-1, 27)^T,\\
&\text{IC3:} \quad \zeta(0) = (0.1, 0, 0)^T, && \hat{\zeta}(0) = (10, 10, 10)^T.
\end{aligned}$$

We make a number of observations. First, as expected, the BTOF-based observer converges for all initial conditions. Second, the observer from [82] is local and its convergence depends on which equilibrium point the observer is designed about and the choice of β. Some experimentation was performed to optimize the region of attraction by varying β. However, this is an unguided empirical process which led to local convergence in all cases considered.

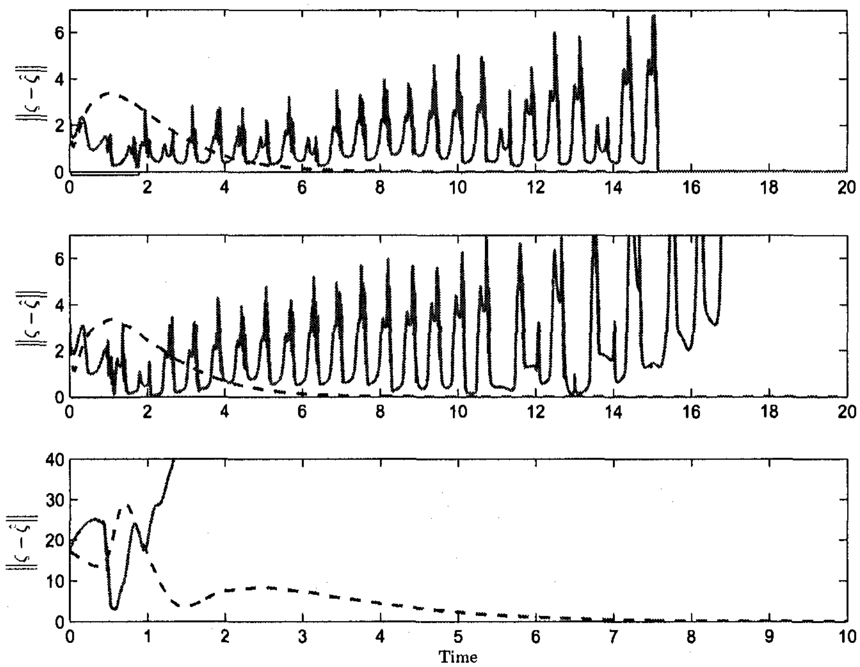

Figure 3.1: 2-Norm of estimate error for BTOF (dashed line) and observer in [82](solid line). Each graph corresponds to a different initial condition.

Other examples to demonstrate the usefulness of the proposed approach include a simple MAGLEV system (2.43). For the indices (2, 1), system (2.43) is already in BTOF provided we include input dependence into γ_1^2. A semi-global observer based on Remark 3.2.11 is applicable if we assume the system state is bounded:

$$\dot{\hat{x}}^1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{x}^1 + \begin{pmatrix} 0 \\ g - \frac{\eta(y_2)^2}{m(x_1^1)^2} \end{pmatrix} + L^1(y_1 - C^1\hat{x}^1),$$
$$\dot{\hat{x}}_1^2 = \frac{2\eta x_1^2 \sigma(\hat{x}_2^1)}{2\eta x_1^1 + \alpha(x_1^1)^2} + \frac{x_1^1}{\alpha x_1^1 + 2\eta} u + l_1^2(y_2 - C^2\hat{x}^2),$$

where $\sigma(\cdot)$ is a saturation function. We remark that an OF does not exist for system (2.43) for x_2^1 appears in the dynamics of x_1^2-subsystem. Further, we note that the approach given in [82] cannot be applied as the spectrum of the system's linearization $F = \frac{\partial f}{\partial x}(x_e)$ about any equilibrium point x_e is not type (C, ν) w.r.t. $\Lambda(F)$, i.e., the last condition of *Main Theorem* in [83] is not satisfied. Other industrially relevant examples for which the BTOF-based design is suitable include: a Synchronous Motor [109, Sec. 1.10.10] with $y = (\delta, i_a, i_b)^T$ and indices $(2, 1, 1)$, a Permanent Magnet Stepper Motor [40, Sec. 3.2] with $y = (q, i_1, i_2)^T$ and indices $(2, 1, 1)$, a Brushless DC Motor [40, Sec. 4.2] with $y = (q, i_a, i_b)^T$ and indices $(2, 1, 1)$, and a Ball-and-Beam system with $y = (\phi, r)^T$ and indices $(2, 2)$ provided a time scale transformation is used for the first subsystem [139, 156]. A straightforward application of the results in this note yields the BTOF coordinates for the above examples. Evidently, for systems which admit a global BTOF, such as the practically relevant ones considered in this section, a BTOF-based observer has significant performance benefits including error dynamics which are GES in BTOF coordinates. In this case, the error dynamics are GAS in the original coordinates. If the inverse map $\zeta = \Theta(z)$ is globally uniformly Lipschitz in z, then the error dynamics are GES.

3.5 Summary

In this chapter, we first discussed the existence conditions of a BTOF for unforced nonlinear multi-output systems. Theorem 3.1.4 removes the upper subsystem state parameter assumption to provide existence conditions. Two counterexamples were provided to demonstrate the difference from the existing work [134, 139]. Condition (iii) in Theorem 3.1.4 is imposed as a parameter assumption. This simplifies the derivation and presentation of the BTOF existence conditions. Theorem 3.1.4 was extended by removing the parameter assumption and new existence conditions of a BTOF were provided in Theorem 3.3.1. Since the vector fields to be rectified depend on unknown functions, conditions in Theorem 3.3.1 were presented as the solvability of PDEs which cannot be verified as easily as those in Theorem 3.1.4. Next BTOF-based observer designs were studied. The BTOF allows an established EEL observer design method to be generalized to a larger class of multi-output systems. Observer design proceeds in a decentralized manner, starting with the uppermost subsystem. Designs for each subsystem effectively treat upper subsystem states as known measurements and are relatively simple given their reduced dimension. The zero solution of the error dynamics was shown to be GES in Theorem 3.2.4 under the global Lipschitz assumption on γ. The global Lipschitz assumption was removed and the semi-global stability of the error dynamics were established in Theorem 3.2.9 by using a saturation technique. Examples illustrated the construction of the BTOF coordinates and the advantages of a BTOF-based design compared to a generally applicable method in [82].

Chapter 4

Time-scaled Observer Form (TOF)

In the previous two chapters, we presented the BTF and BTOF coordinates in which observer design became simplified. These coordinates are generalization of the Observer Form (OF). In this chapter we also consider a generalization, but take a different approach; we show that a time scale transformation can be combined with a state transformation in order to generalize the OF. Recent work in [52, 132] incorporates an output dependent time scale transformation for *single output* nonlinear systems. The extension to multi-output systems form has not appeared in the literature to date. Compared to the time scaling of the single output OF, the multi-output problem allows distinct time scale transformations for each subsystem. We take a different approach relative to [132] in order to simplify the derivation of the existence conditions. This difference in approach is largely due to different time scales for the subsystems.

This chapter considers the existence conditions for a Time-scaled Observer Form (TOF). In Section 4.1 we present two motivational examples and introduce the TOF. In Section 4.2 we discuss both single and multiple time scale transformation cases, and propose the existence conditions of TOF. A comparison between TOF and OF with output transformation [85] is made in Section 4.3. TOF-based observer design ensures global exponential error convergence in the transformed time scales. In Section 4.4 we derive conditions on time scale transformations to preserve global exponential stability in the original time scale.

4.1 Problem Statements

A significant amount of work has been performed on the use of time scale transformations for control design [137, 131, 51]. Early work on time scaling for control in [137] considered a single-input control affine system

$$\frac{d\zeta}{dt} = f(\zeta) + g(\zeta)u, \quad \zeta \in \mathbb{R}^n, u \in \mathbb{R}$$

which is state feedback linearizable in a new time scale. Recall, that a system is said to be locally "state feedback linearizable" if it is locally equivalent to a linear system in Brunovsky controller form by a smooth state feedback $u = k(\zeta) + \beta(\zeta)v, \beta(0) \neq 0$ and a local diffeomorphism $z = T(\zeta)$, where v denotes the auxiliary input. In order to generalize the class of single output systems which admits an OF, an *output dependent* time scale transformation was introduced in [52, 132]:

$$\frac{d\tau}{dt} = s(y(t)) > 0, \quad \tau(t_0) = \tau_0, \tag{4.1}$$

where $s(y)$ is a non-vanishing positive smooth function, called a Time Scale Function (TSF). A TSF uniquely defines a time scale transformation; for simplicity of presentation below, we do not distinguish between a TSF and its associated time scale transformation. As well, we ignore the initial conditions $\tau(t_0) = \tau_0$ since they play no role in our application of the TSF. For the single output case, work in [52, 132] considered the problem of finding a TSF and change of state coordinates $z = T(\zeta)$ to locally transform (2.1) into OF in τ time scale

$$\begin{aligned} \frac{dz}{d\tau} &= Az + \gamma(y), \\ y &= Cz. \end{aligned}$$

When this transformation is possible, a Luenberger-like observer in τ time scale can be designed

$$\frac{d\hat{z}}{d\tau} = A\hat{z} + \gamma(y) + L(y - C\hat{z}),$$

which yields a LTI error dynamics system in τ time scale.

4.1.1 TOF Problem

For the multi output case, we consider distinct time scale transformations for each subsystem

$$\frac{d\tau_i}{dt} = s_i(y(t)) > 0, \quad 1 \leqslant i \leqslant p. \tag{4.2}$$

We define the TOF as an OF in the new time scales

$$\begin{aligned} \frac{dz^i}{d\tau_i} &= A^i z^i + \gamma^i(y), \\ y_i &= C^i z^i, \end{aligned} \qquad 1 \leqslant i \leqslant p, \tag{4.3}$$

where A^i, C^i are given in (1.7). The ith subsystem in the TOF coordinates and t time scale is given by

$$\begin{aligned} \dot{z}^i &= s_i(A^i z^i + \gamma^i(y)), \\ y_i &= C^i z^i, \end{aligned} \qquad 1 \leqslant i \leqslant p,$$

where s_i abbreviates $s_i(y(t))$. Collecting the p subsystems in TOF in the original time we denote the system dynamics as

$$\begin{aligned} \dot{z} &= S(Az + \gamma(y)), \\ y &= Cz, \end{aligned} \tag{4.4}$$

where $S = \text{Blockdiag}\{s_1 I_{\lambda_1}, \cdots, s_p I_{\lambda_p}\}$. Note that the difference between the multi-output and the single-output TOF lies in the matrix S. This difference leads to an alternate and more straightforward approach to deriving the TOF existence conditions. Given TSF (4.2) and TOF (4.3), we then formalize the problem of transformation to TOF.

Definition 4.1.1. The nonlinear system (2.1) locally (globally) observable w.r.t. indices $\lambda_i, 1 \leqslant i \leqslant p$ in Definition A.0.1 is said to be locally (globally) transformable to TOF (4.3) if there exists a local (global) diffeomorphism $z = \Phi(\zeta)$ and time scale transformations (4.2) such that the transformed system in $\tau_k, 1 \leqslant k \leqslant p$ time scales is

$$\frac{dz}{d\tau} = S^{-1} \frac{\partial \Phi(\zeta)}{\partial \zeta} S \hat{f} = Az + \gamma(y), \tag{4.5}$$

where

$$\frac{dz}{d\tau} = \begin{pmatrix} \frac{dz^1}{d\tau_1} \\ \vdots \\ \frac{dz^p}{d\tau_p} \end{pmatrix}, \quad \hat{f} = S^{-1} f.$$

□

System (2.1) is transformable to TOF (4.3) if and only if there exists a change of coordinates $z = \Phi(\zeta)$ and time scale transformations (4.2) such that the system in the t time scale is given by (4.4).

4.1.2 Motivational Examples

The following two examples illustrate that distinct time scale transformations for each subsystem allow for an important degree of freedom in the multi-output case.

Example 4.1. Consider a two-output system in Observable Form with observability indices $(2,2)$ corresponding to the output $y=(y_1,y_2)^T$

$$\dot{x}=\begin{pmatrix} x_2^1 \\ (x_2^1)^2+x_2^1x_2^2 \\ x_2^2 \\ (x_2^2)^2+x_2^1x_2^2 \end{pmatrix}, \qquad y=\begin{pmatrix} x_1^1 \\ x_1^2 \end{pmatrix}. \tag{4.6}$$

We can verify system (4.6) is not transformable to OF since φ^1,φ^2 have terms $(x_2^1)^2, x_2^1x_2^2, (x_2^2)^2$. We introduce the time scale transformation (4.1) with $s(y)=e^{y_1+y_2}$ and rewrite system (4.6) in the τ time scale

$$\frac{dx}{d\tau}=\frac{1}{s}f=\hat{f}, \qquad y=h(x),$$

where s is an abbreviation of $s(y)$. Applying Theorem A.0.2 with f replaced by $\hat{f}$, we solve the starting vectors

$$\hat{g}^1=s\frac{\partial}{\partial x_2^1}, \quad \hat{g}^2=s\frac{\partial}{\partial x_2^2},$$

and verify that the Lie bracket conditions (A.2) are satisfied

$$\left[\mathrm{ad}^k_{-\hat{f}}\hat{g}^r,\mathrm{ad}^l_{-\hat{f}}\hat{g}^q\right]=0, \quad 0\leqslant k,l\leqslant 1; 1\leqslant r,q\leqslant 2.$$

We solve PDEs (A.3) for the new coordinates

$$\Phi(x)=(x_1^1,\frac{x_2^1}{s},x_1^2,\frac{x_2^2}{s})^T, \tag{4.7}$$

in which system (4.6) is a LTI OF

$$\frac{dz}{d\tau}=Az, \qquad y=Cz.$$

□

Example 4.2. We modify the dynamics of the system (4.6) by taking

$$\dot{x}_2^1 = (x_2^1)^2 + x_1^1 x_1^2,$$
$$\dot{x}_2^2 = (x_2^2)^2 + x_1^1.$$

Using the results provided below, we can show the system cannot be put into an OF using a single time scale transformation and a state transformation. If we introduce a different time scale transformation for each subsystem

$$\frac{d\tau_1}{dt} = s_1(y) = e^{y_1},$$
$$\frac{d\tau_2}{dt} = s_2(y) = e^{y_2},$$

the system in the new times can be put into OF

$$\frac{dz}{d\tau} = Az + \begin{pmatrix} 0 \\ e^{-2y_1} y_1 y_2 \\ 0 \\ e^{-2y_2} y_1 \end{pmatrix},$$
$$y = Cz$$

with the local diffeomorphism

$$\Phi(x) = (x_1^1, \frac{x_2^1}{s_1}, x_1^2, \frac{x_2^2}{s_2})^T. \tag{4.8}$$

□

4.2 Existence Conditions

We first present the existence conditions for a TOF when each subsystem has a different time scale transformation. Then necessary and sufficient conditions for a TOF with one time scale transformation are given. These last conditions are presented in a similar form to the established result for OF [167, 132].

4.2.1 Multiple Time Scale Transformation Case

Theorem 4.2.1. *Assume the nonlinear system* (2.1) *is locally observable w.r.t. indices $\lambda_i, 1 \leqslant i \leqslant p$ in Definition A.0.1. The system is locally transformable to TOF* (4.5) *if and only if in U_0*

(i) the TSF of the ith subsystem (4.2) *satisfies the PDEs*

$$\mathrm{d}L_{g^i}L_f^{\lambda_i}h_i = \frac{1}{s_i}\left(l_{\lambda_i}\frac{\partial s_i}{\partial z_1^i}\mathrm{d}L_f h_i + (l_{\lambda_i}-1)\sum_{j=1,j\neq i}^{p}\frac{\partial s_i}{\partial z_1^j}\mathrm{d}L_f h_j\right) \quad (4.9)$$
$$\mod \{\mathrm{d}y\}, \quad 1 \leqslant i \leqslant p,$$

where

$$l_k = \frac{k(k-1)}{2} + 1, \quad 1 \leqslant k \leqslant \lambda_i,$$

and g^i is the starting vector field in the original time and defined by (A.1)*;*

(ii) $Q_i = Q_i \cap Q$*;*

(iii) the Lie brackets conditions are satisfied

$$\left[\eta_r^i, \eta_s^l\right] = 0, \quad \begin{cases} 1 \leq r \leq \lambda_i; \\ 1 \leqslant s \leqslant \lambda_l; \\ 1 \leqslant i, l \leqslant p, \end{cases} \quad (4.10)$$

where for $1 \leqslant i \leqslant p$*,*

$$\eta_1^i = \hat{g}^i, \quad \eta_j^i = \frac{1}{s_i}\mathrm{ad}_{-f}\eta_{j-1}^i, \quad 2 \leqslant j \leqslant \lambda_i, \quad (4.11)$$

and $\hat{g}^i$ are the starting vector fields and defined by

$$L_{\hat{g}^i}L_f^k h_l = s_i^{\lambda_i - 1}\delta_{k,\lambda_i-1}\delta_{l,i}, \quad 0 \leqslant k \leqslant \lambda_i - 1; 1 \leqslant l \leqslant p. \quad (4.12)$$

The transformation $z = \Phi(\zeta)$ is the solution of the n^2 PDEs

$$\frac{\partial \Phi(\zeta)}{\partial \zeta}\left(\eta_{\lambda_1}^1, \cdots, \eta_1^1, \cdots, \eta_{\lambda_p}^p, \cdots, \eta_1^p\right) = I_n. \quad (4.13)$$

Remark 4.2.2. The TOF coordinates are globally defined if the system is globally observable, the theorem conditions hold in $\mathbb{R}^n$, and the vector fields $\eta_j^i, 1 \leqslant j \leqslant \lambda_i, 1 \leqslant i \leqslant p$ are complete. The transformation $\Phi(\zeta)$ can be constructed from the composition of flows of vector fields η_j^i, which is globally defined if the vector fields are complete. □

Remark 4.2.3. Given the nonlinear system (2.1) in Observable Form, we know $L_f^{\lambda_i}h_i = \varphi^i(x), g^i = \partial/\partial x_{\lambda_i}^i$, and $L_f h_k = x_2^k, 1 \leqslant k \leqslant p$. We therefore reformulate Condition (i) as

$$\frac{\partial^2 \varphi^i(x)}{\partial x_2^i \partial x_{\lambda_i}^i} = \frac{l_{\lambda_i}}{s_i}\frac{\partial s_i}{\partial y_i},$$
$$\frac{\partial^2 \varphi^i(x)}{\partial x_2^k \partial x_{\lambda_i}^i} = \frac{l_{\lambda_i}-1}{s_i}\frac{\partial s_i}{\partial y_k}, \quad 1 \leqslant k \leqslant p, k \neq i.$$

Since $s_i > 0$, we introduce the change of variable $\kappa_i = \ln(s_i)$ and rewrite PDEs

$$\frac{\partial \kappa_i}{\partial y_k} = \begin{cases} \frac{1}{l_{\lambda_i}-1}\varphi^i_{k,i}(x), & k \neq i, \\ \frac{1}{l_{\lambda_i}}\varphi^i_{i,i}(x), & k = i, \end{cases}$$

where $\varphi^i_{k,i} = \partial^2\varphi^i(x)/\partial x^k_2 \partial x^i_{\lambda_i}$. A solution of κ_i exists if and only if

$$\frac{\partial^2 \kappa_i}{\partial y_j \partial y_k} = \frac{\partial^2 \kappa_i}{\partial y_k \partial y_j},$$

which imposes conditions on $\varphi^i(x)$:

$$\begin{aligned} \frac{\partial \varphi^i_{k,i}(x)}{\partial y_j} &= \frac{\partial \varphi^i_{j,i}(x)}{\partial y_k}, \quad j \neq i, k \neq i, \\ \frac{1}{l_{\lambda_i}}\frac{\partial \varphi^i_{i,i}(x)}{\partial y_j} &= \frac{1}{l_{\lambda_i}-1}\frac{\partial \varphi^i_{j,i}(x)}{\partial y_i}, \quad k = i, j \neq i. \end{aligned}$$

□

Remark 4.2.4. From definition (4.12), we know $\hat{g}^i = s_i^{\lambda_i - 1} g^i$. Note that the following fact will be used to prove Theorem 4.2.1

$$\begin{aligned} \mathrm{d}h_i \frac{\partial S}{\partial z^i_1}(Az+\gamma) &= \mathrm{d}h_i \frac{\partial S}{\partial z^i_1} S^{-1} \underbrace{S(Az+\gamma)}_{f} \\ &= \frac{\partial s_i}{\partial z^i_1}\frac{1}{s_i}(0_{1\times\nu_{i-1}}, 1, 0_{1\times(n-\nu_{i-1}-1)})f \\ &= \frac{\partial s_i}{\partial z^i_1}\frac{1}{s_i} L_f h_i. \end{aligned}$$

□

Remark 4.2.5. Condition (ii) guarantees the solvability of the starting vectors [167]. This can be illustrated by considering a LTI system with observability indices $(2,1)$

$$\begin{aligned} \dot{\zeta} &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \zeta, \\ y &= \begin{pmatrix} h_1(\zeta) \\ h_2(\zeta) \end{pmatrix} = \begin{pmatrix} \zeta^1_1 \\ \zeta^2_1 \end{pmatrix}. \end{aligned}$$

The starting vector g^1 cannot be solved from

$$\begin{pmatrix} \mathrm{d}h_1(\zeta) \\ \mathrm{d}L_f h_1(\zeta) \\ \mathrm{d}h_2(\zeta) \\ \mathrm{d}L_f h_2(\zeta) \end{pmatrix} g^i = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{with} \quad \begin{pmatrix} \mathrm{d}h_1(\zeta) \\ \mathrm{d}L_f h_1(\zeta) \\ \mathrm{d}h_2(\zeta) \\ \mathrm{d}L_f h_2(\zeta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

since

$$Q_1 = \begin{pmatrix} \mathrm{d}h_1(\zeta) \\ \mathrm{d}h_2(\zeta) \\ \mathrm{d}L_f h_2(\zeta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

$$Q = \begin{pmatrix} \mathrm{d}h_1(\zeta) \\ \mathrm{d}L_f h_1(\zeta) \\ \mathrm{d}h_2(\zeta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$Q_1 \neq Q_1 \cap Q = \begin{pmatrix} \mathrm{d}h_1(\zeta) \\ \mathrm{d}h_2(\zeta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which violates Condition (ii). Hence, the system is not transformable to OF by a linear coordinate transformation. Including a change of output coordinates allows transformation to OF. □

Proof: Necessity: Taking $\eta_1^i = \partial/\partial z_{\lambda_i}^i, 1 \leqslant i \leqslant p$ and following the definition of η_k^i given by (4.11), we verify

$$\eta_k^i = \frac{\partial}{\partial z_{\lambda_i - k+1}^i}, \quad 2 \leqslant k \leqslant \lambda_i.$$

Since $\eta_k^i, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$ are unit vectors and commute, the Lie bracket conditions (4.10) are necessary. Next, we derive the definition of the starting vector $\hat{g}^i$ (4.12). Since $\hat{g}^i = \eta_1^i, 1 \leqslant i \leqslant p$, we have

$$\frac{\partial h_l}{\partial z_{\lambda_i}^i} = 0 \Rightarrow L_{\hat{g}^i} h_l = 0, \quad 1 \leqslant l \leqslant p.$$

Further computation gives

$$\begin{aligned} 0 = \frac{\partial h_l}{\partial z_{\lambda_i - 1}^i} &= \langle \mathrm{d}h_l, \eta_2^i \rangle = \left\langle \mathrm{d}h_l, \frac{1}{s_i}[-f, \hat{g}^i] \right\rangle \\ &= \frac{1}{s_i} \langle \mathrm{d}L_f h_l, \hat{g}^i \rangle - \frac{1}{s_i} L_f \langle \mathrm{d}h_l, \hat{g}^i \rangle \\ &= \frac{1}{s_i} L_{\hat{g}^i} L_f h_l, \end{aligned}$$

for $1 \leqslant l \leqslant p$. Having shown the $k = 0, 1$ cases we can use induction to show

$$L_{\hat{g}^i} L_f^k h_l = \begin{cases} s_i^k \frac{\partial h_l}{\partial z_{\lambda_i - k}^i} = 0, & 0 \leqslant k \leqslant \lambda_i - 2; 1 \leqslant l \leqslant p, \\ s_i^{\lambda_i - 1} \frac{\partial h_i}{\partial z_1^i} = s_i^{\lambda_i - 1}, & k = \lambda_i - 1; 1 \leqslant l \leqslant p. \end{cases} \tag{4.14}$$

Hence, the starting vector $\hat{g}^i$ satisfies (4.12). For the necessity of Condition (ii), it is evident that $Q_i \cap Q \subseteq Q_i$. Thus we only need to show $Q_i \subseteq Q \cap Q_i$. From (4.14), we have

$$Q_i \left(\hat{g}^k, \ldots, \mathrm{ad}_{-f}^{\lambda_1 - \lambda_i - 1} \hat{g}^k \right) = Q_i \left(\eta_1^k, \ldots, \eta_{\lambda_1 - \lambda_i}^k \right) = 0, \quad 1 \leqslant k \leqslant i - 1.$$

Since $\eta_1^k, \ldots, \eta_{\lambda_1-\lambda_i}^k, 1 \leqslant k \leqslant i-1$ are linearly independent and have dimension $\sum_{k=1}^{i-1}(\lambda_k - \lambda_i)$, we conclude

$$\operatorname{rank} Q_i \leqslant n - \sum_{k=1}^{i-1}(\lambda_k - \lambda_i) - 1 = i\lambda_i + \sum_{k=i+1}^{p} \lambda_k - 1.$$

Considering Q is a nonsingular matrix, $\operatorname{rank} Q_i \cap Q = i\lambda_i + \sum_{k=i+1}^{p} \lambda_k - 1$ and $Q_i \subseteq Q_i \cap Q$. We therefore complete the necessity proof of Condition (ii). To derive the condition on TSFs, we first state the equations ensuring the existence of a state transformation $\Phi(\zeta)$, for $1 \leqslant i \leqslant p$

$$s_i \frac{\partial W}{\partial z_{j-1}^i} = \mathrm{ad}_{-f} \frac{\partial W}{\partial z_j^i}, \quad 2 \leqslant j \leqslant \lambda_i, \tag{4.15a}$$

$$\frac{\partial W}{\partial z} \frac{\partial}{\partial z_1^i}(S(Az + \gamma)) = \mathrm{ad}_{-f} \frac{\partial W}{\partial z_1^i}, \quad j = 1, \tag{4.15b}$$

$$\mathrm{d}h_r \frac{\partial W}{\partial z_j^i} = \delta_{j,1}\delta_{r,i}, \quad 1 \leqslant j \leqslant \lambda_i; 1 \leqslant r \leqslant p, \tag{4.15c}$$

where $W = \Phi^{-1}(z)$. $\partial W / \partial z_{\lambda_i}^i$ is verified to be the starting vector since it satisfies the starting vector definition (4.12). Substituting $\hat{g}^i$ into (4.15a), we have

$$\frac{\partial W}{\partial z_j^i} = \eta_{\lambda_i - j+1}^i, \quad 1 \leqslant j \leqslant \lambda_i; 1 \leqslant i \leqslant p.$$

The left hand side of (4.15b) is

$$\frac{\partial W}{\partial z} \frac{\partial}{\partial z_1^i}(S(Az + \gamma)) = \frac{\partial W}{\partial z} \left(\frac{\partial S}{\partial z_1^i}(Az + \gamma) + S \frac{\partial \gamma}{\partial z_1^i} \right).$$

Given the right hand side of (4.15b) in Remark 4.2.6, the multiplication of (4.15b) and $\mathrm{d}h_i(\partial W / \partial z)^{-1}$ gives

$$\mathrm{d}h_i \frac{\partial S}{\partial z_1^i}(Az + \gamma) + \mathrm{d}h_i S \frac{\partial \gamma}{\partial z_1^i} = \mathrm{d}h_i \frac{1}{s_i^{\lambda_i - 1}} \mathrm{ad}_{-f}^{\lambda_i} \hat{g}^i + \mathrm{d}h_i \frac{\sum_{j=1}^{\lambda_i - 1} j}{s_i^{\lambda_i}} L_f(s_i) \mathrm{ad}_{-f}^{\lambda_i - 1} \hat{g}^i. \tag{4.16}$$

According to Remark 4.2.4, we rewrite (4.16) as

$$\frac{\partial s_i}{\partial z_1^i} \frac{1}{s_i} L_f h_i + \rho(y) = \frac{1}{s_i^{\lambda_i - 1}} L_{\mathrm{ad}_{-f}^{\lambda_i} \hat{g}^i} h_i + \frac{\sum_{j=1}^{\lambda_i - 1} j}{s_i^{\lambda_i}} L_f(s_i) L_{\mathrm{ad}_{-f}^{\lambda_i - 1} \hat{g}^i} h_i, \tag{4.17}$$

where $\rho(y) = \mathrm{d}h_i S \partial \gamma / \partial z_1^i$ is some function of y. From [120, Lem. 6.15]

$$L_{\mathrm{ad}_{-f}^{\lambda_i - 1} \hat{g}^i} h_i = L_{\hat{g}^i} L_f^{\lambda_i - 1} h_i = s_i^{\lambda_i - 1},$$

and from [64, Lem. 4.1.2]

$$L_{\mathrm{ad}_{-f}^{\lambda_i}\hat{g}^i} h_i = L_{\hat{g}^i} L_f^{\lambda_i} h_i.$$

Hence, (4.17) is rearranged as

$$\frac{\partial s_i}{\partial z_1^i}\frac{1}{s_i} L_f h_i + \rho(y) = \frac{1}{s_i^{\lambda_i - 1}} L_{\hat{g}^i} L_f^{\lambda_i} h_i - \frac{\lambda_i(\lambda_i - 1)}{2}\frac{1}{s_i} L_f(s_i). \tag{4.18}$$

Collecting the terms of (4.18) and taking the differential, we have

$$\mathrm{d}L_{\hat{g}^i} L_f^{\lambda_i} h_i = l_{\lambda_i} s_i^{\lambda_i - 2}\frac{\partial s_i}{\partial y_i}\mathrm{d}L_f h_i + (l_{\lambda_i} - 1) s_i^{\lambda_i - 2}\sum_{j=1, j\neq i}^{p}\frac{\partial s_i}{\partial y_j}\mathrm{d}L_f h_j \mod \{\mathrm{d}y\}, \tag{4.19}$$

where $y_i = z_1^k, 1 \leqslant k \leqslant p$. Since $\hat{g}^i = s_i^{\lambda_i - 1} g^i$, we have

$$\mathrm{d}L_{\hat{g}^i} L_f^{\lambda_i} h_i = s^{\lambda_i - 1}\mathrm{d}L_{g^i} L_f^{\lambda_i} h_i \mod \{\mathrm{d}y\}.$$

Plugging the above equation into (4.19), we have Condition (i).

Sufficiency: Given the TSF of each subsystem s_i solved from Condition (i), it is readily shown Conditions (ii)–(iii) are sufficient to guarantee the existence of a state transformation $z = \Phi(\zeta)$ which puts system (2.1) into TOF (4.4). This approach was taken in the proof of sufficiency in Theorem 3.1.4. □

Remark 4.2.6. Given (4.15a), we can compute $\partial W/\partial z_j^i, 1 \leqslant j \leqslant \lambda_i - 1$ iteratively and have

$$\frac{\partial W}{\partial z_{\lambda_i - k}^i} = \frac{1}{s_i^k}\mathrm{ad}_{-f}^k \hat{g}^i + \frac{\sum_{j=1}^{k-1} j}{s_i^{k+1}} L_f(s_i)\mathrm{ad}_{-f}^{k-1}\hat{g}^i$$
$$\mathrm{span}\{\mathrm{ad}_{-f}^j \hat{g}^i, 0 \leqslant j \leqslant k-2\}, \quad 1 \leqslant k \leqslant \lambda_i - 1.$$

Further calculation yields the right hand side of (4.15b)

$$\mathrm{ad}_{-f}\frac{\partial W}{\partial z_1^i} = \frac{1}{s_i^{\lambda_i - 1}}\mathrm{ad}_{-f}^{\lambda_i}\hat{g}^i + \frac{\sum_{j=1}^{\lambda_i - 1} j}{s_i^{\lambda_i}} L_f(s_i)\mathrm{ad}_{-f}^{\lambda_i - 1}\hat{g}^i\,\mathrm{span}\{\mathrm{ad}_{-f}^j \hat{g}^i, 0 \leqslant j \leqslant \lambda_i - 2\}.$$

The previous two equations are independent of coordinates and hold provided a TOF exists. Without loss of generality in the above proof we made use of a ζ-coordinate representation in (4.15b). □

Remark 4.2.7. The multiple time scale transformation case has a different TSF for each subsystem. This can be generalized by employing a TSF for each state, i.e., $S = \mathrm{Blockdiag}\{s_1, \cdots, s_n\}$, which leads to the multi-output extension of the output dependent observability linear normal form in [173, 174]. A similar procedure can be followed to obtain the existence conditions of this generalized TOF. □

Example 4.2 (Continued) We recall Example 4.2 and verify the conditions of Theorem 4.2.1, solve for a matrix TSF candidate using Condition (i), and compute the state transformation. The functions $L_{g^i}L_f^2h_i$ and $\mathrm{d}L_fh_i, i = 1,2$ are readily computed and Conditions (i) reduce to

$$2\mathrm{d}x_2^1 = s_1^{-1}(2\frac{\partial s_1}{\partial y_1}\mathrm{d}x_2^1 + \frac{\partial s_1}{\partial y_2}\mathrm{d}x_2^2) \mod \{\mathrm{d}y\},$$
$$2\mathrm{d}x_2^2 = s_2^{-1}(2\frac{\partial s_2}{\partial y_2}\mathrm{d}x_2^2 + \frac{\partial s_2}{\partial y_1}\mathrm{d}x_2^1) \mod \{\mathrm{d}y\},$$

which yield the PDEs

$$\frac{\partial s_1}{\partial y_1} = s_1, \quad \frac{\partial s_1}{\partial y_2} = 0, \quad \frac{\partial s_2}{\partial y_1} = 0, \quad \frac{\partial s_2}{\partial y_2} = s_2.$$

Hence, we solve the TSFs $s_1 = e^{y_1}, s_2 = e^{y_2}$. Computing

$$\eta_1^1 = \frac{\partial}{\partial x_2^1}, \quad \eta_2^1 = \frac{\partial}{\partial x_1^1} + x_2^1\frac{\partial}{\partial x_2^1},$$
$$\eta_1^2 = \frac{\partial}{\partial x_1^2}, \quad \eta_2^2 = \frac{\partial}{\partial x_1^2} + x_2^2\frac{\partial}{\partial x_2^2}$$

and verifying the Lie bracket conditions (4.10) are satisfied for $1 \leqslant r,s,i,l \leqslant 2$, we solve the state transformation (4.8) from (4.13). □

4.2.2 Single Time Scale Transformation Case

The existence conditions are given in the following theorem.

Theorem 4.2.8. *Assume the nonlinear system* (2.1) *is locally observable w.r.t. indices* $\lambda_i, 1 \leqslant i \leqslant p$ *in Definition A.0.1. The system* (2.1) *is locally transformable to TOF* (4.5) *if and only if in* U_0,

(i) Condition (i) in Theorem 4.2.1 with $s = s_i, 1 \leqslant i \leqslant p$ *holds;*

(ii) $Q_i = Q_i \cap Q$*;*

(iii) the Lie brackets conditions are satisfied

$$\left[\mathrm{ad}^k_{-\hat{f}}\hat{g}^r, \mathrm{ad}^l_{-\hat{f}}\hat{g}^q\right] = 0, \quad \left\{ \begin{array}{c} 0 \leq k \leq \lambda_r - 1; \\ 0 \leqslant l \leq \lambda_q - 1; \\ 1 \leqslant r,q \leqslant p, \end{array} \right. \tag{4.20}$$

where $\hat{f} = f/s$, *and* $\hat{g}^i, 1 \leqslant i \leqslant p$, *are the starting vector fields in new time scale and defined by*

$$L_{\hat{g}^i}L_{\hat{f}}^k h_r = \delta_{k,\lambda_i-1}\delta_{i,r}, \quad 0 \leqslant k \leqslant \lambda_i - 1; 1 \leqslant r \leqslant p. \tag{4.21}$$

The transformation $z = \Phi(\zeta)$ is the solution of the n^2 PDEs

$$\frac{\partial \Phi(\zeta)}{\partial \zeta}\left(\mathrm{ad}_{-\hat{f}}^{\lambda_1-1}\hat{g}^1, \cdots, \hat{g}^1, \cdots, \mathrm{ad}_{-\hat{f}}^{\lambda_p-1}\hat{g}^p, \cdots, \hat{g}^p\right) = I_n. \tag{4.22}$$

Proof: Condition (i) in Theorem 4.2.8 can be immediately obtained from Condition (i) in Theorem 4.2.1 by taking $s = s_i$. Given system (2.1) in the new time scale, Conditions (ii)–(iii) in Theorem 4.2.8 are equivalent to Conditions (ii)–(iii) in Theorem A.0.2. Provided the existence of a TSF, Conditions (ii)–(iii) in Theorem 4.2.8 are therefore necessary and sufficient to guarantee the transformation to OF in the new time scale. □

Remark 4.2.9. Theorem 4.2.8 and its proof demonstrate its similarity to Theorem A.0.2. Alternatively, we can show Theorem 4.2.8 by replacing the matrix TSF with a scalar TSF in Theorem 4.2.1. We verify that $\hat{g}^i$ solved from (4.12) is the same as $\hat{g}^i$ solved from (4.21), and $\mathrm{ad}_{-\hat{f}}^{k-1}\hat{g}^i = \eta_k^i, 1 \leqslant k \leqslant \lambda_i, 1 \leqslant i \leqslant p$. Thus Theorem 4.2.8 is proven. When $p = 1$ Theorem 4.2.8 leads to the same existence conditions as stated in [132, Thm. 1]. □

Remark 4.2.10. For a fixed i, it can be verified by induction that for $1 \leqslant r \leqslant p$

$$\mathrm{d}L_{\hat{f}}^k h_r = \begin{cases} s^{-1}\mathrm{d}L_f h_r \mod \{\mathrm{d}h_r\}, & k = 0, \\ s^{-k}\mathrm{d}L_f^k h_r \mod \{\mathrm{d}L_f^j h_r, 0 \leqslant j \leqslant k-1\}, & 1 \leqslant k \leqslant \lambda_i - 1. \end{cases}$$

We therefore conclude that $\hat{g}^i = s^{\lambda_i - 1} g^i$, where $g^i, \hat{g}^i$ are defined by (A.1) and (4.21) respectively. □

Example 4.1 (Continued) We apply Theorem 4.2.8 to solve a scalar TSF, verify the conditions, and compute the state transformation. Given $L_{g^i}L_f^2 h_i, \mathrm{d}L_f h_i, i = 1, 2$ readily obtained, Condition (i) reduces to

$$\begin{aligned} \mathrm{d}(2x_2^1 + x_2^2) &= s^{-1}(2\frac{\partial s}{\partial y_1}\mathrm{d}x_2^1 + \frac{\partial s}{\partial y_2}\mathrm{d}x_2^2) \mod \{\mathrm{d}y\}, \\ \mathrm{d}(2x_2^2 + x_2^1) &= s^{-1}(2\frac{\partial s}{\partial y_2}\mathrm{d}x_2^2 + \frac{\partial s}{\partial y_1}\mathrm{d}x_2^1) \mod \{\mathrm{d}y\}. \end{aligned}$$

The above equations give the PDEs

$$\frac{\partial s}{\partial y_1} = \frac{\partial s}{\partial y_2} = s.$$

Hence, we solve the scalar TSF (4.1). Lie bracket conditions (4.20) are satisfied and the state transformation is (4.7). □

4.3 TOF and OF with Output Transformation

In this section we discuss the difference between TOF and OF with output transformation [85]. Our discussion relies on a system (2.1) being in Observable Form with indices $\lambda_i, 1 \leqslant i \leqslant p$. The starting vector $g^i = \partial/\partial x^i_{\lambda_i}$ and conditions (4.9) are

$$\mathrm{d}\frac{\partial \varphi^i(x)}{\partial x^i_{\lambda_i}} = \frac{1}{s_i}\left(l_{\lambda_i}\frac{\partial s_i}{\partial y_i}\mathrm{d}x^i_2 + (l_{\lambda_i}-1)\sum_{j=1,j\neq i}^{p}\mathrm{d}x^j_2\right) \quad \bmod \{\mathrm{d}y\}, \quad 1 \leqslant i \leqslant p$$

Performing coefficient matching of the above equation, we can solve s_i only if φ^i is affine in $x^i_{\lambda_i}$ and the coefficients of $x^i_{\lambda_i}$ in φ^i are of the form $\alpha_1(y)x^j_2$ or $\alpha_2(y)$. However, without the time scale transformation, the necessary condition for OF requires no terms of the form $\alpha_1(y)x^j_2x^i_{\lambda_i}$ in φ^i. This illustrates a benefit of introducing a time scale transformation. If $\lambda_i = 1$ and $\varphi^i(x)$ has dependence on $x^j_k, k \geqslant 2$, no TOF can be solved. We conclude this by verifying the conditions (4.23c) which are written as

$$\frac{\partial^2 \varphi^i}{\partial y_i \partial \dot{y}_j} = 0, \quad j \neq i.$$

If φ^i has linear dependence on x^j_2, introducing the output transformation leads to the transformability to OF. On the other hand, if $\varphi^i(x)$ depends on y only, no time scale transformation is required for the ith subsystem which is in OF already. Transformation of the other subsystems to OF will not change the expression of the ith subsystem dynamics. We therefore conclude that a time scale transformation is not helpful when attempting to transform one dimensional subsystems.

We perform the comparison between a time scale transformation and an output transformation by considering a multi-output system with observability indices $\lambda_k = 2, 1 \leqslant k \leqslant p$. Assuming the system is in Observable Form, $\hat{g}^i = s_i\partial/\partial x^i_{\lambda_i}, 1 \leqslant i \leqslant p$ are unit vectors and commute. We expand $[\eta^i_2, \eta^k_1] = \left[s_i^{-1}\mathrm{ad}_{-f}\hat{g}^i, \hat{g}^k\right] = 0$ to derive necessary conditions for a TOF:

$$\begin{aligned}
\left[s_i^{-1}\mathrm{ad}_{-f}\hat{g}^i, \hat{g}^k\right] &= \left[\mathrm{ad}_{-f}g^i - s_i^{-1}L_f(s_i)g^i, s_kg^k\right] \\
&= s_k\left[\mathrm{ad}_{-f}g^i, g^k\right] + L_{\mathrm{ad}_{-f}g^i}(s_k)g^k + \left[-s_i^{-1}L_f(s_i)g^i, s_kg^k\right] \\
&= s_k\left[\sum_{l=1}^{p}\frac{\partial \varphi^l}{\partial \dot{y}_i}\frac{\partial}{\partial \dot{y}_l} + \frac{\partial}{\partial y_i}, \frac{\partial}{\partial \dot{y}_k}\right] + L_{g^i}L_f(s_k)g_k + \frac{s_k}{s_i}L_{g_k}L_f(s_i)g^i \\
&= s_k\sum_{l=1}^{p}\frac{\partial^2 \varphi^l}{\partial \dot{y}_i \partial \dot{y}_k}\frac{\partial}{\partial \dot{y}_l} + \frac{\partial s_k}{\partial y_i}\frac{\partial}{\partial \dot{y}_k} + \frac{s_k}{s_i}\frac{\partial s_i}{\partial y_k}\frac{\partial}{\partial \dot{y}_i},
\end{aligned}$$

which yields the PDEs

$$\frac{\partial^2 \varphi^i}{\partial \dot{y}_i^2} + \frac{2}{s_i}\frac{\partial s_i}{\partial y_i} = 0, \tag{4.23a}$$

$$\frac{\partial^2 \varphi^k}{\partial \dot{y}_i \partial \dot{y}_k} + \frac{1}{s_k}\frac{\partial s_k}{\partial y_i} = 0, \quad k \neq i, \tag{4.23b}$$

$$\frac{\partial^2 \varphi^l}{\partial \dot{y}_i \partial \dot{y}_k} = 0, \quad l \neq k; l \neq i, \tag{4.23c}$$

for $1 \leqslant l, i, k \leqslant p$. For the output transformation case, we define the output transformation $\bar{y} = \psi(y) = (\psi_1, \cdots, \psi_p)^T$ and compute

$$L_f^2 \bar{y}_i = \sum_{l=1}^{p} \frac{\partial \psi_i}{\partial y_l}\varphi^l + \sum_{k=1}^{p}\sum_{i=1}^{p} \frac{\partial^2 \psi_i}{\partial y_k \partial y_i}\dot{y}_i \dot{y}_k.$$

For a system transformable to OF with output transformation we require

$$\begin{aligned} \sum_{l=1}^{p} \frac{\partial \psi_i}{\partial y_l}\frac{\partial^2 \varphi^l}{\partial \dot{y}_k \partial \dot{y}_i} + \frac{\partial^2 \psi_i}{\partial y_i \partial y_k} &= 0, \quad i \neq k, \\ \sum_{l=1}^{p} \frac{\partial \psi_i}{\partial y_l}\frac{\partial^2 \varphi^l}{\partial \dot{y}_i^2} + 2\frac{\partial^2 \psi_i}{\partial y_i^2} &= 0, \quad i = k, \end{aligned} \tag{4.24}$$

for $1 \leqslant i, k, l \leqslant p$. Comparing Conditions (4.23) and (4.24), we have the following observations

1. When $p = 1$, i.e., $i = k = l$ in (4.23) and (4.24), the conditions are equivalent since $s_i = \partial \psi_i / \partial y_i$. We recover the result in [132] that an output transformation is equivalent to a time scale transformation for a two dimensional system.

2. In general, an output transformation is not equivalent to a time scale transformation even for a p-output system with observability indices $\lambda_k = 2, 1 \leqslant k \leqslant p$. This demonstrates a difference between the single and multi-output cases.

Next, two systems exemplify the difference between TOF and OF with output transformation.

Example 4.3. Consider an observable system with indices $(2, 2, 2)$

$$\begin{aligned} \dot{x} &= \begin{pmatrix} x_2^1 + \gamma_1^1(y_2) \\ \gamma_2^1(y) \\ x_2^2 + \gamma_1^2(y_1) \\ \gamma_2^2(y) \\ x_2^3 \\ x_2^1 x_2^2 + x_2^3 \end{pmatrix}, \\ y &= \begin{pmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \end{pmatrix}. \end{aligned} \tag{4.25}$$

System (4.25) is not transformable to OF without output transformation since $x_2^1x_2^2$ appears in φ^3. Since conditions (4.23c) do not hold for $l = 3, i = 1, k = 2$, system (4.25) does not admit a TOF. This is consistent with the conclusion from checking the conditions in Theorem 4.2.1. However, we can solve an output transformation $\psi_3 = y_1y_2 - 2y_3$, and the system with new output $y = (x_1^1, x_1^2, \psi_3)^T$ is transformable to OF. $\square$

Example 4.4. Consider a system in Observable Form with indices (2, 2)

$$\begin{aligned}\dot{x} &= \begin{pmatrix} x_2^1 \\ x_1^1x_2^1x_2^2 + x_1^2(x_2^1)^2 \\ x_2^2 \\ x_2^1x_1^2x_2^2 + x_1^1(x_2^2)^2 \end{pmatrix}, \\ y &= \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix} = \begin{pmatrix} x_1^1 \\ x_1^2 \end{pmatrix}.\end{aligned} \tag{4.26}$$

System is not transformable to OF since $(x_2^1)^2, x_2^1x_2^2$, and $(x_2^2)^2$ are present in φ^1, φ^2. We apply Theorem 4.2.1 to investigate whether system (4.26) admits a TOF. The starting vectors are trivially solved as $g^1 = \partial/\partial x_2^1, g^2 = \partial/\partial x_2^2$ and we have

$$\begin{aligned} L_{g^1}L_f^2h_1 &= x_1^1x_2^2 + 2x_1^2x_2^1, \\ L_{g^2}L_f^2h_2 &= x_1^2x_2^1 + 2x_1^1x_2^2. \end{aligned}$$

With $L_fh_k = x_2^k, k = 1, 2$, Condition (i) in Theorem 4.2.1 is applied to set up the PDEs

$$\begin{aligned} 2y_2 &= \frac{1}{s_1}\frac{\partial s_1}{\partial y_1}, \quad y_1 = \frac{1}{s_1}\frac{\partial s_1}{\partial y_2}, \\ 2y_1 &= \frac{1}{s_2}\frac{\partial s_2}{\partial y_2}, \quad y_2 = \frac{1}{s_2}\frac{\partial s_2}{\partial y_1}. \end{aligned}$$

Solving the above PDEs gives TSFs

$$s_1 = s_2 = e^{y_2y_1}.$$

With $s_1 = s_2$, Theorem 4.2.8 can be applied. Since $g^k, k = 1, 2$ exist, Condition (ii) holds. Defining $\hat{f} = f/s_1, \hat{g}^k = s_1g^k, k = 1, 2$, we verify Condition (iii) as well. We therefore know system (4.26) is transformable to a TOF. On the other hand, system (4.26) cannot be put into an OF with an output transformation since no output transformation satisfies PDEs (4.24). $\square$

4.4 Error Dynamics Stability

Assuming the existence of a TOF and considering the Luenberger-like observer in TOF coordinates and the new time scales

$$\frac{d\hat{z}}{d\tau} = A\hat{z} + \gamma(y) + L(y - C\hat{z}), \tag{4.27}$$

we have the LTI error dynamics

$$\frac{d\tilde{z}}{d\tau} = (A - LC)\tilde{z},$$

whose zero solution is GES. The error dynamics of the ith subsystem in the original time scale is written as

$$\dot{\tilde{z}}_i = s_i(A^i - L^iC^i)\tilde{z}^i, \tag{4.28}$$

which is LTV. We study the stability of the error dynamics (4.28) by examining the stability of the LTV system

$$\dot{e} = \varrho(t)A_c e, \quad \varrho(t) > 0, \forall t \geqslant t_0, \tag{4.29}$$

where $e = (e_1, \cdots, e_n)^T$, and $A_c \in \mathbb{R}^{n\times n}$ is Hurwitz. Since the observer gain allows for arbitrary eigenvalue assignment, A_c is assumed diagonalizable. We first give the stability result of (4.29) when $n = 1$.

Proposition 4.4.1. *Given a one dimensional system*

$$\dot{x} = -\sigma\varrho(t)x, \quad x_0 = x(t_0), \sigma > 0, \varrho(t) > 0, \forall t \geqslant t_0, \tag{4.30}$$

its equilibrium point $x = 0$ is GES if and only if there exist positive constants t_0, T_0, and $\epsilon > 0$ such that

$$\int_t^{t+T_0} \varrho(\xi)\mathrm{d}\xi \geqslant \epsilon, \quad \forall t \geqslant t_0. \tag{4.31}$$

Proof: For the LTV system (4.30), its solution is

$$x(t) = e^{-\sigma\int_{t_0}^t \varrho(\xi)\mathrm{d}\xi}x(t_0). \tag{4.32}$$

Sufficiency: From Condition (4.31), we have

$$\left(\frac{(t-t_0)}{T_0} - 1\right)\epsilon \leqslant \int_{t_0}^t \varrho(\xi)\mathrm{d}\xi \leqslant \left(\frac{(t-t_0)}{T_0} + 1\right)\epsilon.$$

Substituting the above equation into (4.32), we know

$$c_1 e^{-m(t-t_0)}|x(t_0)| \leqslant |x(t)| \leqslant c_2 e^{-l(t-t_0)}|x(t_0)|, \tag{4.33}$$

where c_1, m are some positive constants, and $c_2 = \exp(\sigma\epsilon/T_0), l = \sigma\epsilon/T_0$. Hence, the equilibrium point $x = 0$ is GES and the sufficiency of Condition (4.31) is shown. **Necessity:** Since the origin $x = 0$ is GES, the system trajectory satisfies (4.33) where c_1, c_2, m, l are some positive constants. Combining with (4.32), we have

$$\int_{t_0}^{t} \varrho(\xi)\mathrm{d}\xi \geqslant c_3(t - t_0) - c_4, \quad \forall t \geqslant t_0,$$

where c_3, c_4 are some appropriate positive constants. Letting $T_0 > c_4/c_3$ and $\epsilon = c_3 T_0 - c_4 > 0$, and computing the integral from t to $t + T_0$, we have

$$\int_{t}^{t+T_0} \varrho(\xi)\mathrm{d}\xi \geqslant \epsilon, \quad \forall t \geqslant t_0.$$

Thus the proof of the necessity is completed. □

Proposition 4.4.2. *The equilibrium point $e = 0$ of the LTV system* (4.29) *is GES if and only if there exist positive constants t_0, T_0, and $\epsilon > 0$ such that Condition* (4.31) *holds.*

Proof: Since A_c is assumed diagonalizable into $A_d = \mathrm{Diag}\{\sigma_1, \cdots, \sigma_n\}$ by a linear transformation $\bar{e} = (\bar{e}_1, \ldots, \bar{e}_n)^T = He$, system (4.29) is transformed into n decoupled scalar systems

$$\dot{\bar{e}}_i = -\sigma_i \varrho(t)\bar{e}_i, \quad 1 \leqslant i \leqslant n,$$

whose equilibrium points $\bar{e}_i = 0$ are GES. According to Proposition 4.4.1, condition (4.31) is necessary and sufficient to ensure $e_i = 0$ is GES. □

Finally, we state the theorem without proof for the stability of the error dynamics (4.28), $1 \leqslant i \leqslant p$.

Theorem 4.4.3. *Assume system* (2.1) *is globally transformed to TOF* (4.3). *Given the observer* (4.27) *with $A - LC$ Hurwitz, the zero solution of the error dynamics in the original time* (4.28), *$1 \leqslant i \leqslant p$, is GES if and only if there exist positive constants t_0, T_0, and $\epsilon > 0$ such that Condition* (4.31) *holds with $\varrho(\xi)$ replaced by $s_i(y(\xi))$.*

Remark 4.4.4. A non-vanishing positive TSF is necessary and sufficient to preserve the error dynamics stability in the sense of Lyapunov. For a LTV system

$$\dot{x} = A(t)x, \quad x(t_0) = x_0,$$

$x = 0$ is a globally uniformly asymptotically stable (GUAS) equilibrium if and only if $x = 0$ is an GES equilibrium point. Hence, GUAS is guaranteed by Condition (4.31). □

4.5 Summary

The TOF for uncontrolled nonlinear continuous-time systems was considered in this chapter. The multiple and single time scale transformation cases were considered and necessary and sufficient existence conditions were provided in Theorem 4.2.1 and Theorem 4.2.8, respectively. For each case, the unit vector fields are constructed differently. Compared to the existing single output time scaling work in [132], the proof of Theorem 4.2.1 does not directly make use of the OF result. On the other hand, Theorem 4.2.8 can be shown by directly applying the OF result and is therefore analogous in method to the single output time scaling case. Since the time scale transformation affects the error dynamics stability, necessary and sufficient conditions on TSFs to preserve GES and GUAS of the error dynamics were presented.

Chapter 5

Time-scaled Block Triangular Observer Form (TBTOF)

In Chapter 4, a time scale transformation was considered to enlarge the class of systems admitting OF in the new time scales. In this chapter, to further broaden the class of systems allowing EEL design, we proposes a Time-scaled Block Triangular Observer Form (TBTOF) by incorporating time scale transformations with a BTOF. Thanks to its block triangular structure, the TBTOF is more general than the TOF considered in Chapter 4. The TSFs do not necessarily depend on outputs only. Instead, for the ith subsystem, the TSF can be a function of the upper subsystem states $z^k, 1 \leqslant k \leqslant i-1$ and outputs $y_k, 1 \leqslant k \leqslant p$. A system is in TBTOF if in new time scales there exists a state transformation to BTOF. Theorem 3.1.4 is applied to establish the existence conditions of a TBTOF. Since each subsystem is transformed into BTOF sequentially while leaving the upper subsystem unchanged, similar to the single time scale transformation case in Section 4.2, we derive the necessary conditions on the TSF of the ith subsystem. The motivation for considering the TBTOF originated from [135] where it was briefly described.

In Section 5.1 we consider TSFs and the TBTOF. The TBTOF existence conditions are given in Section 5.2. In Section 5.3 we discuss TBTOF-based observer design and error convergence. Section 5.4 presents a Ball-and-Beam system and a mathematical example which admit a TBTOF but not a BTOF; TBTOF-based observer designs are given.

5.1 Problem Statement

Theorem 3.1.4 requires that the original system in new time scales must be in BTF. This fact imposes restriction on the form of TSF for each subsystem. For the ith subsystem, its TSF can only depend on the first i subsystem states and all outputs. On the other hand, a TSF should only depend on measured signals for the purpose of observer design. Considering the observer design for a system in BTOF is performed sequentially, for the ith subsystem, the previous $i-1$ subsystem states can be treated as measurements since the convergence of the error dynamics of the first $i-1$ subsystems can be established independently. Hence, a time scale transformation for the ith subsystem is taken as

$$\frac{d\tau_i}{dt} = s_i(z^{<i-1>}, y_{[i,p]}), \quad 1 \leqslant i \leqslant p. \tag{5.1}$$

Similar to the TOF case, work in [157] chose an output dependent TSF for a TBTOF which led to a straightforward analysis of the error dynamics stability. However, the generalized form (5.1) provides an additional degree of freedom in design. This will be illustrated by the mathematical example in Section 5.4.

After introducing the time scale transformation (5.1), we define the TBTOF whose ith subsystem is

$$\begin{aligned} \dot{z}^i &= s_i(A^i z^i + \gamma^i), \\ y_i &= C^i z^i, \end{aligned} \quad 1 \leqslant i \leqslant p, \tag{5.2}$$

where $\gamma^i = \gamma^i(z^{<i-1>}, z_1^i, y_{[i+1,p]})$. This subsystem can be expressed in BTOF in τ_i time scale

$$\begin{aligned} \frac{dz^i}{d\tau_i} &= A^i z^i + \gamma^i, \\ y_i &= C^i z^i. \end{aligned} \quad 1 \leqslant i \leqslant p. \tag{5.3}$$

Provided that a TBTOF exists, observer design in the new time scales can be performed as described below in Section 5.3. We formalize the TBTOF problem to be solved.

Definition 5.1.1. The nonlinear system (2.1) locally (globally) observable w.r.t. indices $\lambda_i, 1 \leqslant i \leqslant p$ given by Definition 2.1.1 is said to be locally (globally) transformable to a TBTOF (5.2) if there exists a local (global) diffeomorphism $z = \Phi(\zeta)$ and time scale transformations (5.1) such that the transformed system in the τ time scales is in BTOF. $\square$

5.2 Existence Conditions

In order to derive the existence conditions of a TBTOF and a necessary condition on the TSF we follow the method of proof for the BTOF existence conditions in Section 3.1. That is, we assume the first $i-1$ subsystems are in TBTOF and consider the transformation of the ith subsystem from BTF to TBTOF. Specifically, starting with

$$\begin{aligned} \begin{pmatrix} \dot{z}^1 \\ \vdots \\ \dot{z}^{i-1} \\ \dot{\bar{x}}^i \end{pmatrix} &= \bar{F}^i = \begin{pmatrix} s_1(A^1 z^1 + \gamma^1) \\ \vdots \\ s_{i-1}(A^{i-1} z^{i-1} + \gamma^{i-1}) \\ \bar{f}^i \end{pmatrix}, \\ \begin{pmatrix} y_1 \\ \vdots \\ y_{i-1} \\ y_i \end{pmatrix} &= \begin{pmatrix} C^1 z^1 \\ \vdots \\ C^{i-1} z^{i-1} \\ \bar{h}_i(\bar{x}) \end{pmatrix}, \end{aligned} \tag{5.4}$$

we want to transform into

$$\begin{aligned} \dot{z}^k &= s_k(A^k z^k + \gamma^k), \\ y_k &= C^k z^k, \end{aligned} \qquad 1 \leqslant k \leqslant i. \tag{5.5}$$

Expressing (5.4) and (5.5) in the τ_i time scale as

$$\begin{aligned} \begin{pmatrix} \frac{dz^1}{d\tau_i} \\ \vdots \\ \frac{dz^{i-1}}{d\tau_i} \\ \frac{d\bar{x}^i}{d\tau_i} \end{pmatrix} &= \hat{F}^i = \begin{pmatrix} \frac{s_1}{s_i}(A^1 z^1 + \gamma^1) \\ \vdots \\ \frac{s_{i-1}}{s_i}(A^{i-1} z^{i-1} + \gamma^{i-1}) \\ \frac{1}{s_i}\bar{f}^i \end{pmatrix}, \\ \begin{pmatrix} y_1 \\ \vdots \\ y_{i-1} \\ y_i \end{pmatrix} &= \begin{pmatrix} C^1 z^1 \\ \vdots \\ C^{i-1} z^{i-1} \\ \bar{h}_i(\bar{x}) \end{pmatrix}, \end{aligned} \tag{5.6}$$

and

$$\begin{aligned} \frac{dz^k}{d\tau_i} &= \frac{s_k}{s_i}(A^k z^k + \gamma^k), \\ y_k &= C^k z^k, \end{aligned} \qquad 1 \leqslant k \leqslant i, \tag{5.7}$$

respectively, we can apply Theorem 3.1.4 directly on (5.6) to check if it admits a TBTOF (5.7).

Theorem 5.2.1. *Let system* (2.1) *be locally observable in* $U_0 \subset \mathbb{R}^n$ *with indices* $\lambda_i, 1 \leqslant i \leqslant p$ *given by Definition 2.1.1. There exist a change of time scale* (5.1) *and and a local diffeomorphism* $\Phi(z^{<i-1>}, \bar{x}^i)$ *to transform the first* i *subsystems* (5.6) *into* (5.7), *if and only if in* U_0

(i) the TSF (5.1) *must satisfy*

$$\mathrm{d}L_{\bar{g}^i} L_{\bar{F}^i}^{\lambda_i} \bar{h}_i = l_{\lambda_i} \frac{s_i'}{s_i} \mathrm{d}L_{\bar{F}^i} \bar{h}_i \mod \{\mathrm{d}z_k^j, 1 \leqslant k \leqslant \lambda_j, 1 \leqslant j \leqslant i-1; \mathrm{d}z_1^i\}, \quad (5.8)$$

where $s_i' = \partial s_i / \partial y_i$, $\bar{F}^i$ *is defined in* (5.4), *and the* ν_i*-dimensional starting vector* $\bar{g}^i$ *is the unique solution of*

$$L_{\bar{g}^i} L_{\bar{F}^i}^k \bar{h}_i = \delta_{k,\lambda_i - 1}, \quad 0 \leqslant k \leqslant \lambda_i - 1; \quad (5.9)$$

(ii) the first $\nu_{i-1} = \sum_{k=1}^{i-1} \lambda_k$ *components of* $\mathrm{ad}_{\hat{F}^i}^k \hat{g}^i$, $0 \leqslant k \leqslant \lambda_i - 1$ *are zero, where* $\hat{F}^i$ *is defined by* (5.6) *and the* ν_i*-dimensional starting vector* $\hat{g}^i$ *is the unique solution of*

$$L_{\hat{g}^i} L_{\hat{F}^i}^k \bar{h}_i = \delta_{k,\lambda_i - 1}, \quad 0 \leqslant k \leqslant \lambda_i - 1; \quad (5.10)$$

(iii) the Lie bracket conditions are satisfied

$$\left[\mathrm{ad}^r_{-\hat{F}^i} \hat{g}^i, \mathrm{ad}^s_{-\hat{F}^i} \hat{g}^i\right] = 0, \quad 0 \leq r, s \leq \lambda_i - 1;$$

(iv) the vector fields are independent of lower subsystem outputs

$$\frac{\partial}{\partial y_j} \mathrm{ad}^r_{-\hat{F}^i} \hat{g}^i = 0, \quad 0 \leq r \leq \lambda_i - 1; i+1 \leqslant j \leqslant p;$$

(v) the ith subsystem output is independent of lower subsystem states

$$\frac{\partial \bar{h}_i}{\partial \bar{x}_k^j} = 0, \quad 1 \leqslant k \leqslant \lambda_j; i+1 \leqslant j \leqslant p.$$

The state transformation Φ^i *is the solution of the* λ_i^2 *PDEs*

$$\frac{\partial \Phi^i(\bar{x}_e^i)}{\partial \bar{x}_e^i} \left(\mathrm{ad}^{\lambda_i - 1}_{-\hat{F}^i} \hat{g}^i, \ldots, \hat{g}^i\right) = I_{\lambda_i}, \quad (5.11)$$

where $\bar{x}_e^i = ((z^{<i-1>})^T, (\bar{x}^i)^T)^T \in \mathbb{R}^{\nu_i}$.

Remark 5.2.2. By induction, we obtain

$$\mathrm{d}L_{\hat{F}^i}^k \bar{h}_i = \frac{1}{s_i^{k-1}} \mathrm{d}L_{\bar{F}^i}^k \bar{h}_i \mod \{\mathrm{d}z_q^j, 1 \leqslant q \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \mathrm{d}L_{\bar{F}^i}^r \bar{h}_i, 0 \leqslant r \leqslant k-1\}$$

for $0 \leqslant k \leqslant \lambda_i - 1$. The starting vectors $\hat{g}^i, \bar{g}^i$ computed from (5.9) and (5.10) have the formula $\hat{g}^i = s_i^{\lambda_i - 1} \bar{g}^i$. We remark that when calculating $L_{\hat{F}^i}^k \bar{h}_i$ and $L_{\bar{F}^i}^k \bar{h}_i$, $\bar{x}_e^i$ is treated as the state, and $y_{[i+1,p]}$ is treated as a parameter. The same methodology is followed in computing $L_{\hat{F}^i}^k \bar{h}_i, 0 \leqslant k \leqslant \lambda_i$ and the proof of Theorem 5.2.1 below. This is consistent with the proof of Theorem 3.1.4. □

Proof: Conditions (ii)–(v) are apparent by applying Theorem 3.1.4 with the vector fields $\bar{F}^i$, $\bar{g}^i$ replaced by $\hat{F}^i, \hat{g}^i$. Next, we focus on the derivation of the necessary conditions on a TSF, i.e., Condition (i). We assume (5.4) is transformable into (5.5). Based on (5.9) and (5.5), we compute the starting vector expressed in the original time scale and z-coordinates denoted by $\tilde{g}^i$. First, since $\tilde{h}_i(z) = z_1^i$ and defining $\tilde{F}^i$ as

$$
\begin{aligned}
\tilde{F}^i &= \sum_{j=1}^{i} \tilde{f}^j, \\
\tilde{f}^j &= s_j \left(\sum_{k=1}^{\lambda_j - 1} (z_{k+1}^j + \gamma_k^j) \frac{\partial}{\partial z_k^j} + \gamma_{\lambda_j}^j \frac{\partial}{\partial z_{\lambda_j}^j} \right),
\end{aligned}
$$

we obtain $L_{\tilde{F}^i} \tilde{h}_i = s_i(z_2^i + \gamma_1^i)$. Further computation gives

$$
\begin{aligned}
L_{\tilde{F}^i}^2 \tilde{h}_i &= (z_2^i + \gamma_1^i) L_{\tilde{F}^i} s_i + s_i \left(L_{\tilde{F}^i} z_2^i + L_{\tilde{F}^i} \gamma_1^i \right) \\
&= s_i^2 z_3^i + (z_2^i)^2 s_i' s_i + \varrho_2 z_2^i + \beta_2,
\end{aligned}
$$

where ϱ_2, β_2 depend on $(z^{<i-1>}, y_{[i,p]})$. By induction we have

$$
L_{\tilde{F}^i}^k \tilde{h}_i = s_i^k z_{k+1}^i + l_k s_i' z_2^i z_k^i s_i^{k-1} + \varrho_k z_k^i + \beta_k(z^{<i-1>}, z_{[2,k-1]}^i, y_{[i,p]}) \tag{5.12}
$$

holds for $3 \leqslant k \leqslant \lambda_i - 1$. It can also be verified that $\varrho_k, 2 \leqslant k \leqslant \lambda_i - 1$ depend on $(z^{<i-1>}, y_{[i,p]})$. In particular

$$
\begin{aligned}
L_{\tilde{F}^i}^{\lambda_i - 1} \tilde{h}_i =& s_i^{\lambda_i - 1} z_{\lambda_i}^i + l_{\lambda_i - 1} s_i' z_2^i z_{\lambda_i - 1}^i s_i^{\lambda_i - 2} \\
&+ \varrho_{\lambda_i - 1} z_{\lambda_i - 1}^i + \beta_{\lambda_i - 1}(z^{<i-1>}, z_{[2,\lambda_i - 1]}^i, y_{[i,p]}).
\end{aligned} \tag{5.13}
$$

Given $L_{\tilde{F}^i}^k \tilde{h}_i, 0 \leqslant k \leqslant \lambda_i - 1$ and (5.9), the starting vector is solved as

$$
\tilde{g}^i = \frac{1}{s_i^{\lambda_i - 1}} \frac{\partial}{\partial z_{\lambda_i}^i}.
$$

From (5.13), we have

$$
L_{\tilde{F}^i}^{\lambda_i} \tilde{h}_i = l_{\lambda_i} s_i' z_2^i s_i^{\lambda_i - 1} z_{\lambda_i}^i + \varrho_{\lambda_i} z_{\lambda_i}^i + \beta_{\lambda_i},
$$

where ϱ_{λ_i} and β_{λ_i} depend on $(z^{<i-1>}, y_{[i,p]})$ and $(z^{<i-1>}, z_{[2,\lambda_i - 1]}^i, y_{[i,p]})$ respectively. Hence, we have

$$
L_{\tilde{g}^i} L_{\tilde{F}^i}^{\lambda_i} \tilde{h}_i = l_{\lambda_i} s_i' z_2^i + \frac{1}{s_i^{\lambda_i - 1}} \varrho_{\lambda_i}. \tag{5.14}
$$

On the other hand, we know

$$\mathrm{d}L_{\tilde{F}^i}\tilde{h}_i = s_i \mathrm{d}z_2^i \mod \left\{\mathrm{d}z_k^j, 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \mathrm{d}z_1^i\right\}.$$

We therefore simplify (5.14) into

$$\mathrm{d}L_{\bar{g}^i}L_{\tilde{F}^i}^{\lambda_i}\tilde{h}_i = l_{\lambda_i}\frac{s_i'}{s_i}\mathrm{d}L_{\tilde{F}^i}\tilde{h}_i \mod \left\{\mathrm{d}z_k^j, 1 \leqslant k \leqslant \lambda_j; 1 \leqslant j \leqslant i-1, \mathrm{d}z_1^i\right\},$$

which is the expression of (5.8) in z-coordinates. Since it is independent of coordinates, we have shown Condition (i) holds for $\lambda_i \geqslant 4$. By direct computation, Condition (i) holds when $\lambda_i = 2, \lambda_i = 3$. □

Remark 5.2.3. Since (5.8) is a necessary condition on the TSF, its solution can be too general to give an explicit expression of the TSF candidate. This is the case we have in the Ball-and-Beam example in Section 5.4. To avoid this situation, necessary and sufficient conditions for TSFs would be required as in [137] where the feedback linearization problem is considered. On the other hand, conditions on TSFs in [137] are much more complicated and difficult to verify. Hence, no additional conditions on TSFs are considered here. □

Remark 5.2.4. Condition (iv) in Theorem 5.2.1 implicitly impose some constraints on TSFs. In some special cases, Condition (iv) can give particularly straightforward constraints on the TSF candidates. For example, when $\bar{g}^i$ has no dependence on $y_{[i+1,p]}$, the TSF candidate s_i should be independent of $y_{[i+1,p]}$. This is because with the time scale transformation introduced, $\hat{g}^i = s_i^{\lambda_i - 1}\bar{g}^i$, applying Condition (iv) for $r = 0$ in Theorem 5.2.1 with $\bar{g}^i$ replaced by $\hat{g}^i$ we end up with conditions $\partial\hat{g}^i/\partial y_j = 0, i+1 \leqslant j \leqslant p$. On the other hand, if $\bar{g}^i$ is a function of $y_{[i+1,p]}$, the TSF candidate must have dependence on $y_{[i+1,p]}$ such that $\partial\hat{g}^i/\partial y_j = 0, i+1 \leqslant j \leqslant p$. □

Remark 5.2.5. In the single output case, the output transformation is equivalent to the time scale transformation when the system order is two [132]. We have two examples to illustrate this is not the case for the TBTOF. The first example demonstrates that a system admits a TBTOF but does not admit an OF with an output transformation. The second example admits an OF with an output transformation but does not admit a TBTOF. We consider a two-output system with indices $(2,2)$

$$\begin{aligned} \begin{pmatrix} \dot{z}^1 \\ \dot{\bar{x}}_1^2 \\ \dot{\bar{x}}_2^2 \end{pmatrix} &= \begin{pmatrix} A^1 z^1 + \gamma^1(y) \\ \bar{x}_2^2 \\ \sigma(z^1)(\bar{x}_2^2)^2 \end{pmatrix}, \\ y &= \begin{pmatrix} z_1^1 \\ \bar{x}_1^2 \end{pmatrix}. \end{aligned} \tag{5.15}$$

System (5.15) is not transformable to OF or BTOF since $(\bar{x}_2^2)^2$ appears. To find out if it admits a TBTOF, Condition (i) in Theorem 5.2.1 is applied to solve the TSF for the second subsystem. We can solve the TSF $s_2 = e^{\sigma(z^1)\bar{x}_1^2}$ and verify that Conditions (ii)–(v) in Theorem 5.2.1 are satisfied with $\hat{F}^2 = \bar{F}^2/s_2, \hat{g}^2 = s_2\bar{g}^2$. Hence, there exists a state transformation to put system (5.15) into TBTOF. On the other hand, we show by contradiction that the second subsystem with new output $\bar{y}_2$ cannot be put into OF. Assuming the existence of an output transformation $\bar{y}_2 = \psi_2(y)$, the characteristic equation for the second subsystem is

$$L_f^2 \bar{y}_2 = L_f \gamma_1^2(y_1, \bar{y}_2) + \gamma_2^2(y_1, \bar{y}_2),$$

where γ_1^2, γ_2^2 are the components of the output injections. Further computation leads to the necessary condition on the transformability to OF, that is the so-called polynomial condition: $\partial^2 L_f^2 \bar{y}_2 / \partial \dot{\bar{y}}_2{}^2 = 0$. This condition yields the following PDE

$$\frac{\partial^2}{\partial(\bar{x}_1^2)^2}\psi_2(z_1^1, \bar{x}_1^2) + \sigma(z^1)\frac{\partial}{\partial \bar{x}_1^2}\psi_2(z_1^1, \bar{x}_1^2) = 0,$$

which in general is unsolvable for an output transformation $\psi_2(y)$ since z_2^1 appears in σ. Hence, we conclude system (5.15) admits a TBTOF but not necessarily OF with output transformations. Next, we present a perspective dynamic system [33] with indices $(2,1)$

$$\begin{aligned} \dot{\xi} &= \begin{pmatrix} -\omega_1\xi_1\xi_2 + \omega_2(1+\xi_1^2) - \omega_3\xi_2 + (b_1 - b_3\xi_1)\xi_3 \\ \omega_2\xi_1\xi_2 - \omega_1(1+\xi_2^2) + \omega_3\xi_1 + (b_2 - b_3\xi_2)\xi_3 \\ -(\omega_1\xi_2 - \omega_2\xi_1 + b_3\xi_3)\xi_3 \end{pmatrix}, \\ y &= \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \end{aligned} \tag{5.16}$$

where $\xi = (\xi_1, \xi_2, \xi_3)^T$. System (5.16) is locally observable if

$$(b_1 - b_3\xi_1)^2 + (b_2 - b_3\xi_2)^2 \neq 0,$$

which is a commonly obtained expression, referred to as the focus of expansion. Without loss of generality, we assume $b_1 - b_3\xi_1 \neq 0$. We have shown in [39] that system (5.16) can be transformed to OF with the output transformation

$$\begin{aligned}\bar{y}_1 &= \frac{1}{b_3 y_1 - b_1},\\ \bar{y}_2 &= \frac{b_2 - b_3\xi_2}{b_3(b_1 - b_3\xi_1)}.\end{aligned}$$

The first subsystem is not transformable to BTOF since Condition (ii) in Theorem 3.1.4 is not satisfied. We attempt to transform the system into a TBTOF by introducing a time scale transformation for the first subsystem. Applying Condition (i) in Theorem 5.2.1 yields the PDE for s_1

$$\frac{4b_3}{b_3 y_1 - b_1} = \frac{2}{s_1} s_1'$$

Solving this PDE gives the TSF $s_1 = (b_3 y_1 - b_1)^2$. We verify that Condition (iv) in Theorem 5.2.1 is satisfied if and only if

$$\omega_1 b_1 + \omega_3 b_3 = 0.$$

Hence, system (5.16) is in general not transformable to TBTOF. □

5.3 Observer Design and Error Dynamics Stability

5.3.1 Observer Design

Assume the existence of a TBTOF and consider the following observer structure in BTOF coordinates and the original time scale

$$\begin{aligned}\dot{\hat{z}}^1 &= s_1\left[A^1\hat{z}^1 + \gamma^1 + L^1(y_1 - C^1\hat{z}^1)\right],\\ \dot{\hat{z}}^i &= \hat{s}_i\left[A^i\hat{z}^i + \hat{\gamma}^i + L^i(y_i - C^i\hat{z}^i)\right], \quad 2 \leqslant i \leqslant p,\end{aligned} \tag{5.17}$$

where $s_1 = s_1(y), \hat{s}_k = s_k(\hat{z}^{<k-1>}, y_{[k,p]}), \hat{\gamma}^k = \gamma^k(\hat{z}^{<k-1>}, z_1^k, y_{[k+1,p]}), 2 \leqslant k \leqslant p$. The error dynamics for (5.2) and (5.17) is

$$\begin{aligned}\dot{\tilde{z}}^1 &= s_1(A^1 - L^1C^1)\tilde{z}^1,\\ \dot{\tilde{z}}^i &= s_i\left[A^i z^i + \gamma^i\right] - \hat{s}_i\left[A^i\hat{z}^i + \hat{\gamma}^i + L^i(y_i - C^i\hat{z}^i)\right], \quad 2 \leqslant i \leqslant p,\end{aligned} \tag{5.18}$$

where $\tilde{z}^k = z^k - \hat{z}^k, 1 \leqslant k \leqslant p$. For a TBTOF, a time scale transformation is employed as well as a change of state coordinates. The stability of the error dynamics might be affected by the time scale transformation. We first give an example error dynamics which has different asymptotic behavior in different time scales. Then the conditions for a TSF to preserve the error dynamics stability are discussed briefly. Finally, the stability result of the error dynamics (5.18) is given.

5.3.2 Stability Preserving TSF

Example 5.1. We consider

$$\frac{d\tilde{z}}{d\tau} = -\tilde{z}, \quad \tilde{z}(0) = 1,$$

where the τ time scale is defined by

$$\frac{d\tau}{dt} = e^{-t} > 0, \quad \tau_0 = t_0 = 0.$$

Clearly, $\tilde{z}(\tau) = e^{-\tau}, \tau \geqslant 0$ is globally exponentially convergent to the equilibrium $\tilde{z} = 0$. However, $\tilde{z}(t)$ solved from $\mathrm{d}\tilde{z}/\mathrm{d}t = e^{-t}\tilde{z}, z(0) = 1$ gives the solution

$$\tilde{z}(t) = e^{e^{-t}-1}, \quad t \geqslant 0,$$

which converges to e^{-1} as $t \to \infty$. Hence, a C^{∞} and positive TSF does not necessarily ensure that the stability is preserved. Additional conditions on the TSF are required. If we consider the map

$$\tau(t) = \int_0^t s(\epsilon)\mathrm{d}\epsilon : \mathbb{R}^+ \to \mathbb{R}^+,$$

which can be solved analytically $\tau(t) = 1 - e^{-t}, t \geqslant 0$, we notice the time scale transformation defines a homeomorphism $\tau(t)$ mapping $\mathbb{R}^+$ to a finite interval $[0, 1)$. This means the trajectory of $\tilde{z}(t), t \geqslant 0$ is same as $\tilde{z}(\tau), \tau \in [0, 1)$, which explains

$$\lim_{t\to\infty} \tilde{z}(t) = \lim_{\tau\to 1} \tilde{z}(\tau) = e^{-1}.$$

□

We can intuitively give sufficient conditions on the time scale transformation $\tau(t)$ to preserve Uniformly Asymptotic Stability (UAS):

1. $\tau(t)$ is a monotonically increasing function of t;

2. $\lim_{t\to\infty} \tau(t) \to \infty$.

As well, we can show these conditions are necessary. Note the above conditions on the map $\tau(t)$ can only guarantee that UAS in τ time scale implies UAS in t time scale, not the converse. The conditions for the other direction can be similarly derived. Finally, we have the conditions on the time scale transformation to ensure that the zero solutions of the error dynamics in different time scales are UAS.

Lemma 5.3.1. *Given a system in different time scales t, τ,*

$$\frac{d\tilde{x}}{dt} = s(t)f(\tilde{x}), \quad \frac{d\tilde{x}}{d\tau} = f(\tilde{x}),$$

where

$$\frac{d\tau}{dt} = s(t) > 0, \quad \tau(0) = 0, t \geqslant 0,$$

the following statements are equivalent:

1. *the zero solution of system in τ time scale is GUAS $\Leftrightarrow$ it is GUAS in t time scale;*

2. *the maps*

$$\tau(t) = \int_0^t s(\epsilon)\mathrm{d}\epsilon : \mathbb{R}^+ \to \mathbb{R}^+$$
$$t(\tau) = \int_0^\tau \frac{1}{s(\epsilon)}\mathrm{d}\epsilon : \mathbb{R}^+ \to \mathbb{R}^+$$

satisfy

$$\lim_{t\to\infty} \tau(t) \to \infty, \quad \lim_{\tau\to\infty} t(\tau) \to \infty.$$

A sufficient condition for a TSF s to preserve GES is $J \leqslant |s| \leqslant K$ for some positive constants J, K. This can be readily established from the Lyapunov argument. We simplify the following discussion by introducing a *stability preserving TSF* which defines a new time scale such that the stability of an equilibrium point in the new time scale is equivalent to that in the original time scale. We denote the sets $\mathcal{G}_a$ and $\mathcal{G}_e$ consisting of GUAS preserving TSFs and GES preserving TSFs respectively.

5.3.3 Error Dynamics Stability

To simplify the stability analysis of the error dynamics (5.18) we make the following assumptions.

Assumption 5.3.2. The TSF candidates $s_k, s_k^{-1}, 1 \leqslant k \leqslant p$ in the error dynamics (5.18) belong to $\mathcal{G}_a$. □

Assumption 5.3.3. The state of system (2.1) belongs to a compact subset $\mathcal{D} \subset \mathbb{R}^n$, i.e., there exists a constant M such that $\|\zeta(t)\| \leqslant M, t \geqslant 0$. □

Provided that the TBTOF coordinates are defined on a set containing $\mathcal{D}$, Assumption 5.3.3 implies that the state of system (2.1) in TBTOF coordinates lies in a compact subset $\tilde{\mathcal{D}}$ of $\mathbb{R}^n$, i.e., there exists a constant $\tilde{M}$ such that $\|z(t)\| \leqslant \tilde{M}, t \in \mathbb{R}^+$.

Given Assumption 5.3.3 and that s_i, γ^i are smooth, we have s_i and γ^i are Lipschitz in $z^{<i-1>} \in \tilde{\mathcal{D}}$ uniformly in $y_{[i,p]} \in \tilde{\mathcal{D}}$ w.r.t. any norm, i.e., there exist constants $H_k^i, M_k^i,\ 1 \leqslant k \leqslant i-1$ such that

$$
\begin{aligned}
|s_i - \hat{s}_i| &\leqslant \sum_{k=1}^{i-1} H_k^i \left\|\tilde{z}^k\right\|, \\
\left\|\gamma^i - \hat{\gamma}^i\right\| &\leqslant \sum_{k=1}^{i-1} M_k^i \left\|\tilde{z}^k\right\|,
\end{aligned}
\tag{5.19}
$$

for all $y_{[i,p]}, \hat{z}^k, z^k \in \tilde{\mathcal{D}}$. When considering the stability of the ith subsystem error dynamics, the observer state $\hat{z}^{<i-1>}$ might leave $\tilde{\mathcal{D}}$. The Lipschitz condition (5.19) still holds if we introduce an element-wise saturation function $\sigma(\cdot)$ as in Section 3.2. For the ith subsystem observer, we saturate $\hat{z}^{<i-1>}$ to keep it in $\tilde{\mathcal{D}}$, i.e., in (5.17) we take

$$
\begin{aligned}
\hat{s}_i &= \hat{s}_i(\sigma(\hat{z}^{<i-1>}), y_{[i,p]}), \\
\hat{\gamma}^i &= \hat{\gamma}^i(\sigma(\hat{z}^{<i-1>}), z_1^i, y_{[i+1,p]}).
\end{aligned}
$$

We now present the error dynamics stability result.

Theorem 5.3.4. *Let Assumptions 5.3.2, 5.3.3 hold and provided that a TBTOF for system* (2.1) *is well-defined on $\mathbb{R}^n$, we consider the observer* (5.17)*. If $A^i - L^iC^i, 1 \leqslant i \leqslant p$ are Hurwitz, the solution $\tilde{z} = 0$ of the error dynamics* (5.18) *is GUAS.*

Proof: $s_k, s_k^{-1} \in \mathcal{G}_a$ implies $\hat{s}_k, \hat{s}_k^{-1} \in \mathcal{G}_a$. The error dynamics (5.18) being GUAS in $\tau_k, \hat{\tau}_k$ time scales is equivalent to them being GUAS in the t time scale. Hence, we only need to show (5.18) is GUAS in any one of $\tau_k, \hat{\tau}_k, t$ time scales. We use an induction argument to prove that the error dynamics (5.18) is GUAS in $\hat{\tau}_k$ time scale. First, we consider the first subsystem error dynamics

$$
\frac{d\tilde{z}^1}{d\hat{\tau}_1} = (A^1 - L^1C^1)\tilde{z}^1.
$$

We define its Lyapunov function candidate $V_1 = (\tilde{z}^1)^T P^1 \tilde{z}^1$, where $P^1 > 0$ is to be determined. Since (A^1, C^1) is observable we can always solve $L^1, P^1 > 0$ such that

$$\frac{dV_1}{d\hat{\tau}_1} \leqslant -\mu_1 \left\|\tilde{z}^1\right\|^2,$$

where μ_1 is arbitrarily large. Hence, the error dynamics of the first subsystem is GUAS in $\hat{\tau}_1$ time scale.

Assuming the error dynamics of the first $i-1$ subsystems is GUAS in $\hat{\tau}_i$ time scale, we consider the stability of the ith subsystem error dynamics

$$\frac{d\tilde{z}^i}{d\hat{\tau}_i} = (A^i - L^i C^i)\tilde{z}^i + \gamma^i - \hat{\gamma}^i + \frac{s^i - \hat{s}^i}{\hat{s}^i}(A^i z^i + \gamma^i). \tag{5.20}$$

Treating $\tilde{z}^{<i-1>}$ as an input u, and $z^{<i-1>}, y_{[i+1,p]}$ as functions of $\hat{\tau}_i$, (5.20) is rewritten as $\mathrm{d}\tilde{z}^i/\mathrm{d}\hat{\tau}_i = \rho(\tilde{z}^i, u, \hat{\tau}_i)$. Clearly, $\rho(\tilde{z}^i, u, \hat{\tau}_i)$ is Lipschitz in $\tilde{z}^i$ and u on $\tilde{\mathcal{D}}$ which can be extended globally as in [46]. On the other hand, taking $u = 0$, we have $\mathrm{d}\tilde{z}^i/\mathrm{d}\hat{\tau}_i = (A^i - L^i C^i)\tilde{z}^i$, which is GES. Combined with the global Lipschitz property, the error dynamics (5.20) is ISS [76, Lem. 4.6]. Given the first $i-1$ subsystems' error dynamics are GUAS and the ith subsystem error dynamics is ISS, we can apply [76, Lem. 4.7] to show the zero solution of the first i subsystem error dynamics is GUAS in $\hat{\tau}_i$ time scale. Therefore, by induction we have shown the zero solution of the entire system's error dynamics is GUAS. $\square$

Remark 5.3.5. If Assumption 5.3.2 is replaced by $s_k, s_k^{-1} \in \mathcal{G}_e$, with Assumption 5.3.3 and (5.19), the zero solution of the entire system's error dynamics can be shown to be GES. This follows from an induction argument and [104, Lem. III.1]. $\square$

5.4 Observer Design Examples

5.4.1 Ball-and-Beam Example

To illustrate the TBTOF-based observer design, we consider a Ball-and-Beam system. It can be shown that this system does not admit a BTOF. However, it is transformable to a TBTOF. From [139] the dynamics for the system in BTF with indices $\lambda_1 = \lambda_2 = 2$ is

$$\dot{\bar{x}} = \begin{pmatrix} \frac{\bar{x}_2^1}{m(\bar{x}_1^2)^2+J} \\ u - mg\bar{x}_1^2\cos(\bar{x}_1^1) \\ \bar{x}_2^2 \\ \frac{(\bar{x}_2^1)^2\bar{x}_1^1}{(m(\bar{x}_1^2)^2+J)^2} - g\sin(\bar{x}_1^1) \end{pmatrix}, \tag{5.21}$$
$$y = \begin{pmatrix} \bar{x}_1^1 \\ \bar{x}_1^2 \end{pmatrix}.$$

Let $\bar{x}^k = (\bar{x}_1^k, \bar{x}_2^k)^T, y_k = \bar{x}_1^k, k = 1, 2$ denote the state and output of the kth subsystem, respectively. We first verify the conditions of Theorem 3.1.4 to see if the first subsystem can be put into BTOF in t time. Define

$$\begin{aligned} \bar{F}^1 &= \left(\frac{\bar{x}_2^1}{m(\bar{x}_1^2)^2 + J}, -mg\bar{x}_1^2\cos(\bar{x}_1^1)\right)^T, \\ \bar{x}_e^1 &= \bar{x}^1. \end{aligned}$$

The starting vector $\bar{g}^1$ is obtained from

$$\bar{g}^1 = \left(\frac{\partial}{\partial \bar{x}_e^1}\begin{pmatrix} \bar{h}_1 \\ L_{\bar{F}^1}\bar{h}_1 \end{pmatrix}\right)^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = (m(\bar{x}_1^2)^2 + J)\frac{\partial}{\partial \bar{x}_2^1},$$

where $\bar{h}_1 = y_1$. Since $\bar{g}^1$ depends on y_2 which violates Condition (iii), the first subsystem is not transformable to the BTOF. Next we investigate if the first subsystem admits a TBTOF. Using Condition (i) in Theorem 5.2.1, we calculate

$$\begin{aligned} L_{\bar{F}^1}^2\bar{h}_1 &= -\frac{mg\bar{x}_1^2\cos(\bar{x}_1^1)}{m(\bar{x}_1^2)^2 + J}, & L_{\bar{g}^1}L_{\bar{F}^1}^2\bar{h}_1 &= 0, \\ \mathrm{d}L_{\bar{F}^1}\bar{h}_1 &= \frac{1}{m(\bar{x}_1^2)^2 + J}\mathrm{d}\bar{x}_2^1, & \mathrm{d}L_{\bar{g}^1}L_{\bar{F}^1}^2\bar{h}_1 &= 0, \end{aligned}$$

and setup the PDE for s_1

$$l_2\frac{s_1'}{s_1}\mathrm{d}L_{\bar{F}^1}\bar{h}_1 = 0.$$

Solving this PDE gives $s_1 = \varrho(y_2)$, where ϱ is some positive-valued C^∞ function depending on y_2. Hence, we introduce the new time scale for the first subsystem given by $\mathrm{d}\tau_1/\mathrm{d}t = \varrho(y_2)$. The first subsystem in τ_1 time scale is $\hat{F}^1 = \bar{F}^1/\varrho(y_2), y_1 = \hat{h}_1 = \bar{x}_1^1, \bar{x}_e^1 = (\bar{x}_1^1, \bar{x}_2^1)^T$. Solving the starting vector $\hat{g}^1$ in τ_1 time scale from (3.2) gives

$$\hat{g}^1 = \varrho(y_2)(m(\bar{x}_1^2)^2 + J)\frac{\partial}{\partial \bar{x}_2^1}.$$

According to Theorem 5.2.1, Condition (iv) has to hold, i.e., $\partial\hat{g}^1/\partial y_2 = 0$. Taking

$$\varrho(y_2) = \frac{1}{m(\bar{x}_1^2)^2 + J} > 0,$$

which belongs to $\mathcal{G}_e$, we have $\hat{g}^1 = \partial/\partial\bar{x}_2^1$ and $\mathrm{ad}_{-\hat{F}^1}\hat{g}^1 = \partial/\partial\bar{x}_1^1$. It is trivial to verify the other conditions in Theorem 5.2.1 are satisfied. This implies that the first subsystem admits a BTOF in τ_1 time scale. Further, $\hat{g}^1, \mathrm{ad}_{-\hat{F}^1}\hat{g}^1$ are unit vector fields, and the BTF and BTOF coordinates are identical. We observe the second subsystem is in BTOF in t time scale. System (5.21) is therefore transformable to a TBTOF. The observers for the first and second subsystem expressed in τ_1 and t time scales, respectively, are

$$\frac{d\hat{z}^1}{d\tau_1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{z}^1 + \underbrace{\begin{pmatrix} 0 \\ (m(z_1^2)^2 + J)(u - mgz_1^2\cos(z_1^1)) \end{pmatrix}}_{\hat{\gamma}^1(z_1^1, y_2, u)} + L^1(y_1 - C^1\hat{z}^1),$$

$$\dot{\hat{z}}^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{z}^2 + \underbrace{\begin{pmatrix} 0 \\ \frac{(\hat{z}_2^1)^2 z_1^1}{(m(z_1^2)^2 + J)^2} - g\sin(z_1^1) \end{pmatrix}}_{\hat{\gamma}^2(z_1^1, \hat{z}_2^1, y_2)} + L^2(y_2 - C^2\hat{z}^2),$$

where $L^k = (l_1^k, l_2^k)^T, k = 1, 2$ are observer gains. The observer in x-coordinates and t time scale is

$$\dot{\hat{x}} = \begin{pmatrix} \frac{s_1(y_2)}{s_1(\hat{y}_2)} f^1(\hat{x}^1) \\ f^2(\hat{x}^2) \end{pmatrix} + \left(\frac{\partial \hat{z}}{\partial \hat{x}}\right)^{-1} \begin{pmatrix} s_1(y_2)\left(\hat{\gamma}^1 - \hat{\gamma}^{1*} + L^1(y_1 - C^1\hat{z}^1)\right) \\ \hat{\gamma}^2 - \hat{\gamma}^{2*} + L^1(y_2 - C^2\hat{z}^2) \end{pmatrix}, \tag{5.22}$$

where $\hat{\gamma}^{1*} = \hat{\gamma}^1(\hat{z}_1^1, \hat{y}_2, u), \hat{\gamma}^{2*} = \hat{\gamma}^2(\hat{z}^1, \hat{y}_2)$, and

$$z(x) = (x_1^1, x_2^1(m(x_1^2)^2 + J), x_1^2, x_2^2)^T.$$

Take the observer gains $L^1 = (4,4)^T, L^2 = (4,4)^T$, $J = 0.020002\text{kg}\cdot\text{m}^2$, $m = 0.05$ kg, $g = 9.81$ m/s^2, and the initial condition of system (5.21) and its observer (5.22) as

$$x(0) = (0.1, 0.5, 0.1, 0.5)^T,$$
$$\hat{x}(0) = (0.1, 0, 0.1, 0)^T.$$

Given the approximate feedback linearization and estimated state-based feedback

$$\begin{aligned}
u(\hat{x}, y) =& \frac{\nu(\hat{x}, y) - \beta(\hat{x}, y)}{\alpha(\hat{x}, y)}, \\
\nu(\hat{x}, y) =& ab^4\cos(bt) + 8(ab^3\sin(bt) + g\hat{x}_2^1\cos(y_1)) + 24(g\sin(y_1) - ab^2\cos(bt)) \\
& - 32(ab\sin(bt) + \hat{x}_2^2) + 16(a\cos(bt) - y_2), \\
\beta(\hat{x}, y) =& \frac{g\cos(y_1)}{m(\hat{x}_2^1)^2 + J}(2m\hat{x}_2^1\hat{x}_2^2 y_2 + mgy_2\cos(y_1)) + g(\hat{x}_2^1)^2\sin(y_1), \\
\alpha(\hat{x}, y) =& -\frac{g\cos(y_1)}{m(\hat{x}_2^1)^2 + J},
\end{aligned}$$

and the sinusoidal reference $y_1 = a\cos(bt)$ with $a = 1, b = 0.63$ [59], the simulation result is shown in Figure 5.1.

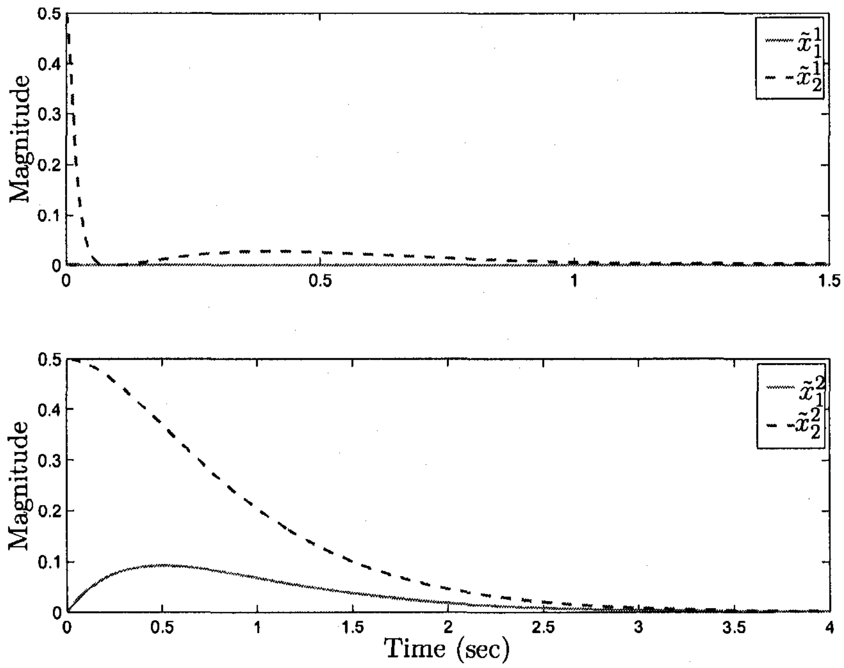

Figure 5.1: Estimate error of TBTOF observer for Ball-and-Beam System

5.4.2 Mathematical Example

For the Ball-and-Beam example, a time scale transformation is solved which is only output dependent. In this section we provide a mathematical example which results in a TSF with more general dependence. We consider a three-output system

$$\begin{aligned}
\dot{\bar{x}}^1 &= \bar{f}^1 = \begin{pmatrix} \bar{x}_2^1 - \bar{x}_1^1 \\ \gamma_2^1(\bar{x}_1^1, \bar{x}_1^2) \end{pmatrix}, \\
\dot{\bar{x}}^2 &= \bar{f}^2 = \begin{pmatrix} \bar{x}_2^2 - 2\bar{x}_1^2 \\ \varrho_2^2(\bar{x}_1^1, \bar{x}_1^2)(\bar{x}_2^2)^2 + \beta_2^2(\bar{x}^1, \bar{x}_1^2, \bar{x}_1^3) \end{pmatrix}, \\
\dot{\bar{x}}^3 &= \bar{f}^3 = \begin{pmatrix} \bar{x}_2^3 - \bar{x}_1^3 \\ \varrho_2^3(\bar{x}^{<2>}, \bar{x}_1^3)(\bar{x}_2^3)^2 + \beta_2^3(\bar{x}^{<2>}, \bar{x}_1^3) \end{pmatrix}, \\
y &= \begin{pmatrix} \bar{x}_1^1 \\ \bar{x}_1^2 \\ \bar{x}_1^3 \end{pmatrix}.
\end{aligned} \tag{5.23}$$

The first subsystem is in BTOF. The second and third subsystems are in BTF in the original time scale. Our objective is to transform the second and third subsystems into BTOF to perform observer design. For the second subsystem, we define the extended vector field and extended state respectively

$$\bar{F}^2 = ((\bar{f}^1)^T, (\bar{f}^2))^T,$$

$$\bar{x}_e^2 = ((\bar{x}^1)^T, (\bar{x}^2)^T)^T.$$

According to (3.2), the starting vector $\bar{g}^2$ of the second subsystem in t time scale has the expression

$$\bar{g}^2 = \varrho_1^2(\bar{x}_e^2)\frac{\partial}{\partial \bar{x}_1^2} + \varrho_2^2(\bar{x}_e^2)\frac{\partial}{\partial \bar{x}_2^2}.$$

where $\varrho_k^2, k = 1, 2$ are smooth functions. Solving the following equations

$$L_{\bar{g}^2}\bar{h}_2 = 0,$$
$$L_{\bar{g}^2}L_{\bar{F}^2}\bar{h}_2 = 1,$$

where $\bar{h}_2 = \bar{x}_1^2, L_{\bar{F}^2}\bar{h}_2 = \bar{x}_2^2 - 2\bar{x}_1^2$, we have

$$\bar{g}^2 = \frac{\partial}{\partial \bar{x}_2^2}.$$

Next, we check Condition (ii) in Theorem 3.1.4

$$\left[\mathrm{ad}_{\bar{F}^2}\bar{g}^2, \bar{g}^2\right] = 2\varrho_2(\bar{x}_1^1, \bar{x}_1^2)\frac{\partial}{\partial \bar{x}_2^2} \neq 0.$$

Hence, the second subsystem is not transformable to BTOF in t time scale. To investigate the existence of a TBTOF, we first solve for a TSF $s_2(\bar{x}^1, y_2, y_3)$ for the second subsystem. We calculate the Lie derivatives

$$L_{\bar{F}^2}^2\bar{h}_2 = \varrho_2^2(\bar{x}_1^1, \bar{x}_1^2)(\bar{x}_2^2)^2 + \beta_2^2(\bar{x}^1, \bar{x}_1^2) - 2(\bar{x}_2^2 - 2\bar{x}_1^2),$$
$$L_{\bar{g}^2}L_{\bar{F}^2}^2\bar{h}_2 = 2\varrho_2^2(\bar{x}_1^1, \bar{x}_1^2)\bar{x}_2^2 - 2,$$

and the differential of $L_{\bar{g}^2}L_{\bar{F}^2}^2\bar{h}_2$

$$\mathrm{d}L_{\bar{g}^2}L_{\bar{F}^2}^2\bar{h}_2 = 2\varrho_2^2(\bar{x}_1^1, \bar{x}_1^2)\mathrm{d}\bar{x}_2^2 \mod \left\{\mathrm{d}\bar{x}_1^1, \mathrm{d}\bar{x}_1^2\right\},$$

where $\mathrm{d}\bar{x}_2^2 = \mathrm{d}L_{\bar{F}^2}\bar{h}_2 + \mathrm{d}\bar{x}_1^2$. From condition (5.8) the PDE defining the TSF is

$$\frac{s_2'}{s_2} = \varrho_2^2(\bar{x}_1^1, y_2),$$

whose general solution is

$$s_2(\bar{x}^1, y_2, y_3) = \kappa_2(\bar{x}^1, y_3)e^{\int \varrho_2^2(\bar{x}_1^1, y_2)\mathrm{d}y_2},$$

where $\kappa_2(\bar{x}^1, y_3)$ is some function depending on $\bar{x}^1, y_3$. For simplicity, we choose $\kappa_2(\bar{x}^1, y_3) = 1$ and the τ_2 time scale is defined by

$$\frac{d\tau_2}{dt} = e^{\int \varrho_2^2(\bar{x}_1^1, y_2)\mathrm{d}y_2}.$$

Given system (5.23) in the τ_2 time scale, we define the vector field $\hat{F}^2 = \bar{F}^2/s_2$ and compute the starting vector $\hat{g}^2 = s_2\bar{g}^2$. Verifying the conditions of Theorem 5.2.1, we know the second subsystem can be put into TBTOF.

Following the analogous procedure as with the second subsystem, we conclude the third subsystem is not transformable to BTOF in t time scale but to a TBTOF with the τ_3 time scale given by

$$\frac{d\tau_3}{dt} = s_3(\bar{x}^{<2>}, y_3) = e^{\int \varrho_2^3(\bar{x}^{<2>}, y_3)dy_3}.$$

We perform an observer design for the case

$$\begin{aligned} \gamma_2^1(\bar{x}_1^1, \bar{x}_1^2) &= -\bar{x}_1^1 + \bar{x}_1^2, \\ \varrho_2^2(\bar{x}_1^1, \bar{x}_1^2) &= -1, \quad \beta_2^2(\bar{x}^1, \bar{x}_1^2) = e^{\bar{x}_1^1}, \\ \varrho_2^3(\bar{x}^1, \bar{x}^2, \bar{x}_1^3) &= -\bar{x}_2^1, \quad \beta_2^3(\bar{x}^1, \bar{x}^2, \bar{x}_1^3) = \bar{x}_1^1\bar{x}_2^1. \end{aligned}$$

The TSFs for the second and third subsystems are

$$\begin{aligned} s_2(\bar{x}^1, y_2, y_3) &= e^{-y_2}, \\ s_3(\bar{x}^{<2>}, y_3) &= e^{-\bar{x}_2^1 y_3}. \end{aligned}$$

Solving the transformation $\Phi^2(\bar{x}_e^2)$ for the second subsystem from the PDEs

$$\frac{\partial\Phi^2(\bar{x}_e^2)}{\partial\bar{x}_e^2}\left(\mathrm{ad}_{-\hat{F}^2}\hat{g}^2, \hat{g}^2\right) = I_2$$

gives the general solution

$$\Phi^2(\bar{x}_e^2) = \begin{pmatrix} \bar{x}_1^2 + \omega_1^2(\bar{x}^1) \\ (\bar{x}_2^2 + 2\bar{x}_1^2 - 2)e^{\bar{x}_1^2} + \omega_2^2(\bar{x}^1) \end{pmatrix}.$$

Taking $\omega_1^2 = \omega_2^2 = 0$, the transformation is a global diffeomorphism. After the transformation, the second subsystem is written as

$$\dot{z}^2 = s_2\begin{pmatrix} z_2^2 + e^{z_1^2}(-4z_1^2 + 2) \\ e^{2z_1^2}\left(-4(z_1^2)^2 + e^{z_1^1}\right) \end{pmatrix}.$$

Solving the following PDEs for the transformation of the third subsystem

$$\frac{\partial\Phi^3(\bar{x})}{\partial\bar{x}}\left(\mathrm{ad}_{-\hat{F}^3}\hat{g}^3, \hat{g}^3\right) = I_2$$

gives the general solution

$$\begin{aligned} \phi_1^3(\bar{x}) &= \bar{x}_1^3 + \omega_1^3(\bar{x}_e^2), \\ \phi_2^3(\bar{x}) &= \frac{1}{(\bar{x}_2^1)^2}\left[(\bar{x}_2^3 + \bar{x}_1^3)(\bar{x}_2^1)^2 + ((\bar{x}_1^1 - \bar{x}_1^2)\bar{x}_1^3 - 1)\bar{x}_2^1 + \bar{x}_1^2 - \bar{x}_1^1\right]e^{\bar{x}_2^1\bar{x}_1^3} + \omega_2^3(\bar{x}_e^2). \end{aligned}$$

Taking $\omega_1^3 = \omega_2^3 = 0$, this particular transformation is well-defined everywhere except $\bar{x}_2^1 = 0$, and it transforms the third subsystem into

$$\dot{z}^3 = s_3(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} z^3 + \gamma^3) = s_3(A^3 z^3 + \gamma^3),$$

where

$$\begin{aligned}
\gamma_1^3 &= e^{z_2^1 z_1^3} \left\{ -2z_1^3 - \frac{1}{(z_2^1)^2} \left[z_2^1(z_1^3 z_1^1 - 1 - z_1^3 z_1^2) + z_1^2 - z_1^1 \right] \right\}, \\
\gamma_2^3 &= e^{3z_2^1 z_1^3} \left\{ -z_1^3 + 2(z_1^3)^2 z_1^1 - (z_1^3)^2 z_1^2 + z_2^1 (z_1^3)^2 - z_2^1 z_1^1 \right. \\
&\quad + \frac{1}{z_2^1} \left[z_1^3 \bar{x}_2^2 + 1 + (z_1^3)^2 (z_1^1)^2 + (z_1^3)^2 (z_1^2)^2 - z_1^3 z_1^2 - 2(z_1^3)^2 z_1^1 z_1^2 \right] \\
&\quad + \frac{1}{(z_2^1)^2} \left[4z_1^3 z_1^1 z_1^2 + z_1^2 - \bar{x}_2^2 - 2z_1^3 (z_1^1)^2 - 2z_1^3 (z_1^2)^2 \right] \\
&\quad \left. + \frac{1}{(z_2^1)^3} \left[2(z_1^1)^2 + 2(z_1^2)^2 - 4z_1^2 z_1^1 \right] \right\}, \\
\bar{x}_2^2 &= e^{-z_1^2}(z_2^2 - 2e^{z_1^2} z_1^2 + 2e^{z_1^2}).
\end{aligned}$$

We design the observer for each subsystem of (5.23) in z-coordinates and t, τ_2, τ_3 time scales respectively. For the first subsystem we take

$$\dot{\hat{z}}^1 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \hat{z}^1 + \begin{pmatrix} 0 \\ z_1^2 \end{pmatrix} + L^1(y_1 - C_1 \hat{z}^1).$$

We choose the observer gains $L^1 = (1, 0)^T$ to place the eigenvalues of $A^1 - L^1 C^1$ at -1. This observer gives the LTI error dynamics

$$\dot{\tilde{z}}^1 = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix} \tilde{z}^1.$$

For the second subsystem we take the observer

$$\dot{\hat{z}}^2 = s_2 \left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \hat{z}^2 + \begin{pmatrix} e^{z_1^2}(-4z_1^2 + 2) \\ e^{2z_1^2}\left(-4(z_1^2)^2 + e^{\hat{z}_1^1}\right) \end{pmatrix} + L^2(y_2 - C^2 \hat{z}^2) \right],$$

where $L^2 = (l_1^2, l_2^2)^T$. The corresponding error dynamics is

$$\dot{\tilde{z}}^2 = s_2 \left[\begin{pmatrix} -l_1^2 & 1 \\ -l_2^2 & 0 \end{pmatrix} \tilde{z}^2 + e^{2z_1^2} \begin{pmatrix} 0 \\ e^{z_1^1} - e^{\hat{z}_1^1} \end{pmatrix} \right]. \tag{5.24}$$

If y_2 is bounded, s_2 has finite lower and upper bound which are positive, the stability of (5.24) is same as

$$\frac{d\tilde{z}^2}{d\tau_2} = \begin{pmatrix} -l_1^2 & 1 \\ -l_2^2 & 0 \end{pmatrix} \tilde{z}^2 + e^{2z_1^2} \begin{pmatrix} 0 \\ e^{z_1^1} - e^{\hat{z}_1^1} \end{pmatrix}. \tag{5.25}$$

To stabilize (5.25), since it consists of a LTI part and a disturbance which converges exponentially to zero, we take the observer gains $L^2 = (2,1)^T$ for eigenvalues at -1 to get exponentially stable error dynamics. For the third subsystem, the observer is taken

$$\frac{d\hat{z}^3}{d\hat{\tau}_3} = A^3\hat{z}^3 + \hat{\gamma}^3 + L^3(y_3 - C^3\hat{z}^3),$$

where $L^3 = (l_1^3, l_2^3)^T$. The error dynamics in t time scale is

$$\begin{aligned}\dot{\tilde{z}}^3 &= s_3(A^3z^3 + \gamma^3) - \hat{s}_3(A^3\hat{z}^3 + \hat{\gamma}^3 + L^3(y_3 - C^3\hat{z}^3)) \\ &= \hat{s}_3\left\{(A^3 - L^3C^3)\tilde{z}^3 + \left[(\gamma^3 - \hat{\gamma}^3) + \frac{(s_3 - \hat{s}_3)}{\hat{s}_3}(A^3z^3 + \gamma^3)\right]\right\}.\end{aligned} \tag{5.26}$$

Since $\hat{s}_3$ depends on $\hat{z}^{<2>}$ which itself is exponentially convergent to a bounded $z^{<2>}$, $\hat{s}_3$ has finite positive lower bound and upper bound. Therefore, the stability of (5.26) is equivalent to that of

$$\frac{d\tilde{z}^3}{d\hat{\tau}_3} = (A^3 - L^3C^3)\tilde{z}^3 + \underbrace{\left[(\gamma^3 - \hat{\gamma}^3) + \frac{(s_3 - \hat{s}_3)}{\hat{s}_3}(A^3z^3 + \gamma^3)\right]}_{disturbance}. \tag{5.27}$$

The disturbance converges exponentially to zero due to the exponential stability of $\tilde{z}^{<2>}$, the Lipschitz dependence of γ^3, s_3, and the boundedness of $A^3z^3 + \gamma^3$. Hence, if the eigenvalues of $A^3 - L^3C^3$ are in the left-half-plane, system (5.27) has an exponentially stable equilibrium. We take $L^3 = (2,1)^T$ to place the eigenvalues of $A^3 - L^3C^3$ at -1.

Since the transformation Φ^3 is not a global diffeomorphism, the state estimate is not globally convergent. If any bounded trajectory $z(t) \in \tilde{\mathcal{D}}$ and estimate trajectory $\hat{z}(t) \in \mathbb{R}^n$ satisfy $z_2^1(t), \hat{z}_2^1(t) \neq 0, \forall t \geqslant 0$, then the error dynamics are exponentially stable. The system (5.23) has three equilibria. One of its equilibria

$$\bar{x}_{e1} = (0.715, 0.715, 0.715, 1.430, 0.846, 0.846)^T$$

is stable. This point is mapped by Φ^2, Φ^3 into

$$z_{e1} = (0.715, 0.715, 0.715, 1.756, 0.846, 0.536)^T.$$

Hence, we choose the system's initial condition

$$z(0) = (-1, 1, -1, 1, -1, 1)^T$$

in the domain of attraction of z_{e1} to ensure $\|z(t)\|$ is bounded and $z(t) \neq 0, t \geqslant 0$. Together with $A^2 - L^2C^2, A^3 - L^3C^3$ Hurwitz we ensure exponential convergence of the estimate error provided the initial conditions gives $\hat{z}_2^1(t) \neq 0, t \geqslant 0$. Taking the initial condition of the observer as

$$\hat{z}(0) = (0, 0.5, 0, 0, 0, 0)^T,$$

the simulation results in z-coordinates and t time scale are shown in Figures 5.2, 5.3, and 5.4.

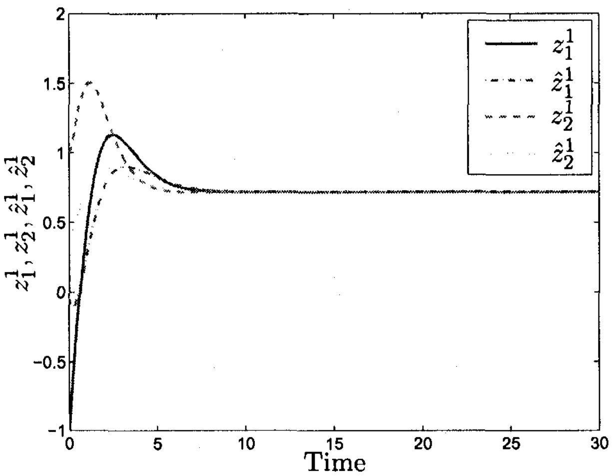

Figure 5.2: Actual and estimated states of the first subsystem

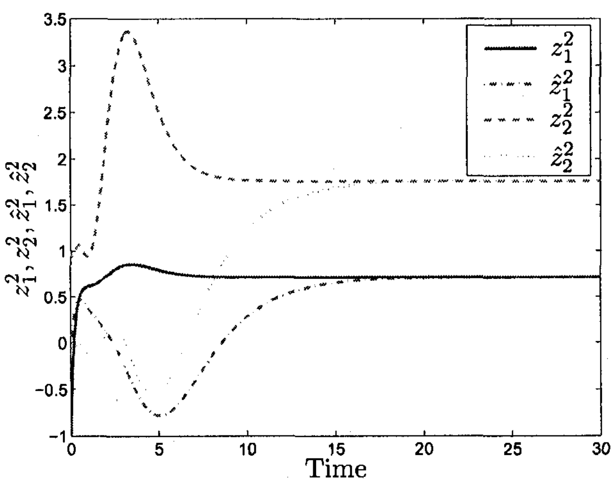

Figure 5.3: Actual and estimated states of the second subsystem

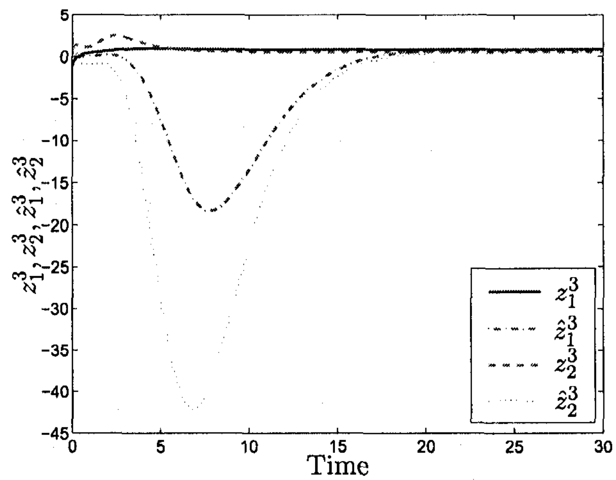

Figure 5.4: Actual and estimated states of the third subsystem

5.5 Summary

This chapter has discussed the existence conditions of a TBTOF for multi-output systems which generalizes an established BTOF by including time scale transformations. Since a block triangular structure leads to sequential observer design, the TSFs allow for more general dependence. That is, dependence on the upper subsystem state as well as outputs. A necessary condition for the TSF was presented. A stability result for the observer error dynamics was given. Two examples which do not admit a BTOF but admit a TBTOF were presented and TBTOF-based observer designs were given.

Chapter 6

Adaptive Observer Design for Nonlinear Systems

The nonlinear adaptive observer design problem was partially addressed by introducing certain Adaptive Observer Forms (AOFs) in [11, 105, 103, 104, 37, 16, 169]. A relatively straightforward stability analysis of the error dynamics resulted for systems with linearly parameterized (LP) dynamics and linear outputs. The linearity requirement restricts the class of systems allowing adaptive observer design. In this chapter, we first consider the adaptive observer design for a LP system but without the linear output constraint. The adaptive observer design is carried out on the basis of the Nonlinear Output Observer Forms (NOOF) in [73, 82]. The existence conditions of NOOF are commonly satisfied [82] and the adaptive observer design can be performed in general. Next, we consider adaptive observer design for a class of NLP systems admitting an OF:

$$\begin{aligned} \dot{z} &= Az + \gamma(y, \theta) \\ y &= Cz \end{aligned} \tag{6.1}$$

where A, C are given by (1.6), $z \in \mathbb{R}^n, \theta \in \mathbb{R}^m$, and $\gamma \in \mathbb{R}^n$ is a C^∞ vector field.

In Section 6.1 we recall some fundamental results, assumptions and stability theorems in the adaptive observer design setting. In Section 6.2, an adaptive observer for a LP NOOF is proposed and the local exponential stability of the error dynamics is proven. In Section 6.3, we present a local adaptive observer for a class of NLP systems in OF and the local exponential stability of its error dynamics are established. In Section 6.4, an example illustrates the adaptive observer design procedures.

6.1 Fundamental Results

The design of an adaptive law and stability analysis are the core parts of adaptive observer design. Indeed, these are not trivial even for linear adaptive observer design since the error dynamics are generally Nonlinear Time-Varying (NLTV). We first introduce some key concepts and results for adaptive systems, then demonstrate the use of these techniques by introducing a typical linear adaptive observer, and finally present a typical nonlinear adaptive observer to illustrate its connection to the linear case. Strictly Positive Real (SPR) functions play a central role in the stability analysis using Lyapunov's method.

Definition 6.1.1. [116, Def. 2.6.1] A rational function $H(s)$ of the complex variable $s = \sigma + j\omega$ is Positive Real (PR) if

(i) $H(s)$ is real for real s;

(ii) $\Re[H(s)] \geqslant 0$ for all $\Re[s] > 0$.

□

Definition 6.1.2. [116, Def. 2.7] A rational function $H(s)$ is SPR if $H(s-\epsilon)$ is PR for some $\epsilon > 0$. □

Alternate definitions of SPR give certain criteria for a SPR function [63, 147]. From the definitions of PR and SPR, it is clear that if $H(s)$ is PR, its phase shift for all frequencies lies in the interval $[-\pi/2, \pi/2]$. Hence, if $H(s)$ is the transfer function of a causal system, the relative degree can only be either zero or one [116]. The connection between a SPR function and the existence of a Lyapunov function is demonstrated by the Lefschetz-Kalman-Yakubovich (LKY) Lemma. The choice of the Lyapunov function is therefore simplified substantially when the transfer function of a linear system is SPR.

Lemma 6.1.3 (LKY Lemma). *[116, Lem. 2.6] Given a scalar $\gamma \geqslant 0$, a vector h, an asymptotically stable matrix A, a vector b such that (A, b) is controllable, and a positive definite matrix L, there exist a scalar $\epsilon > 0$, a vector q and a symmetric positive-definite matrix P satisfying*

$$A^T P + PA = -qq^T - \epsilon L,$$

$$Pb - h = \sqrt{\gamma} q,$$

if and only if

$$H(s) = \frac{1}{2}\gamma + h^T(sI - A)^{-1}b$$

is SPR.

The controllability requirement in LKY Lemma was relaxed in [111]. As shown below, SPR and LKY Lemma are commonly used to prove the uniform stability of the error dynamics in adaptive observer design. To show UAS, Barbalat's Lemma and its corollary are the most commonly used tools.

Lemma 6.1.4 (Barbalat's Lemma). *[138, Lem. 1.2.1] If $f(x)$ is a uniformly continuous function, such that $\lim_{t\to\infty}\int_0^t f(\tau)\mathrm{d}\tau$ exists and is finite, then $f(t) \to 0$ as $t \to \infty$.*

Corollary 6.1.5. *[138, Cor. 1.2.2] If $g, \dot{g} \in \mathcal{L}_\infty$, and $g \in \mathcal{L}_p$, for some $p \in [1, \infty)$, then $g(t) \to 0$ as $t \to \infty$.*

For a time-varying system

$$\dot{x} = f(x, t), \tag{6.2}$$

if $f(x,t)$ is additionally assumed to be bounded for all $t \geqslant t_0$ for any bounded x, the LaSalle-Yoshizawa Theorem [76, Thm. 8.4] can be applied to prove stability.

6.1.1 A Typical Linear Adaptive Observer

Consider a single-input single-output LTI plant of order n

$$\begin{aligned} \dot{x} &= Ax + bu, \\ y &= cx, \end{aligned} \tag{6.3}$$

where input u is a piecewise-continuous uniformly bounded function of time, and (A, b) are unknown parameters. The plant (6.3) is assumed to be observable. The objective is to construct an adaptive observer to estimate both the *parameters* as well as the *state*. Although the triple (c, A, b) contains $n^2 + 2n$ elements, only $2n$ parameters are needed to uniquely determine the input-output relationship. Of these, n correspond to the matrix A while the remaining n are contained in c and b. Different realizations of the plant transfer function lead to different adaptive observer designs. One of the adaptive observers is based on the minimal realization of the plant

$$\begin{aligned} \dot{x} &= [-k, \bar{A}]x + gy + bu, \\ y &= cx = x_1, \end{aligned}$$

where k is chosen such that $K = [-k \ \bar{A}]$ is asymptotically stable, g, b are the unknown vectors, and

$$\bar{A} = \begin{pmatrix} 1_{1\times(n-1)} \\ \mathrm{diag}(-a_2, \cdots, -a_n) \end{pmatrix},$$

where $1_{1\times(n-1)}$ is a $1 \times (n-1)$ row vector of ones, and $a_k > 0, 2 \leqslant k \leqslant n$. To estimate g, b, the observer is chosen

$$\begin{aligned} \dot{\hat{x}} &= K\hat{x} + \hat{g}y + \hat{b}u, \\ \hat{y} &= c\hat{x}, \end{aligned}$$

where $\hat{g}, \hat{b}$ are the estimates of the parameter vectors g, b respectively, and $\hat{x}$ is the estimate of x. The error equation is derived

$$\begin{aligned} \dot{\tilde{x}} &= K\tilde{x} + \tilde{g}y + \tilde{b}u + v_1(t) + v_2(t), \\ \tilde{y} &= c\tilde{x}, \end{aligned}$$

where $v_1(t), v_2(t)$ are auxiliary signals, and

$$\tilde{g} = g - \hat{g}, \quad \tilde{b} = b - \hat{b}.$$

The signals $v_1(t), v_2(t)$ are designed such that the output error asymptotically converges to ϵ_1 which is determined by [116, Thm. 4.2]

$$\begin{aligned} \dot{\epsilon} &= K\epsilon + d\phi^T\omega, \\ \epsilon_1 &= h^T\epsilon, \end{aligned} \tag{6.4}$$

where $\omega = (y, u)^T$, $\phi^T = (\tilde{g}, \tilde{b})$, $h^T(sI - K)^{-1}d$ is SPR. The adaptive observer for the error system (6.4) can be designed based on the following well-established stability result.

Theorem 6.1.6. *[116, Thm. 4.1] Let a dynamical system be represented by the controllable and observable triple (h_1^T, A_1, b_1) where A_1 is an asymptotically stable matrix, and*

$$W(s) = h_1^T(sI - A_1)^{-1}b_1 \text{ is SPR.}$$

Let the elements of a vector $\omega(t)$ be bounded piecewise-continuous functions. Then the origin of the system

$$\begin{aligned} \dot{\epsilon} &= A_1\epsilon + b_1\phi^T\omega, \quad \epsilon_1 = h_1^T\epsilon \\ \dot{\phi} &= -\epsilon_1\omega \end{aligned}$$

is globally uniformly stable.

Proof: Since the transfer function $W(s)$ is SPR, according to LKY Lemma 6.1.3, there exists a positive definite matrix P satisfying

$$\begin{aligned} A_1^T P + P A_1 &= -Q, \\ Pb &= h_1^T, \end{aligned}$$

where Q is positive definite. Taking the Lyapunov function candidate

$$V(\epsilon, \phi) = \epsilon^T P \epsilon + \phi^T \phi,$$

we compute its time derivative

$$\begin{aligned} \dot{V} &= -\epsilon^T Q \epsilon + \epsilon^T P b_1 \phi^T \omega - \phi^T \epsilon_1 \omega \\ &= -\epsilon^T Q \epsilon \leqslant 0, \end{aligned}$$

which implies uniform stability. □

As shown in the proof of Theorem 6.1.6, the boundedness of ϵ and ϕ is concluded from $\dot{V}$ being negative definite. From the boundedness of V, we have $\epsilon \in \mathcal{L}_2$. If $\omega(t)$ is further assumed to be bounded, the boundedness of $\dot{\epsilon}$ is concluded. Applying Barbalat's Lemma and its corollary, we know the zero solution of ϵ is GUAS. Furthermore, global asymptotic convergence of $\phi(t) \to 0$ as $t \to \infty$ can be established under the persistent excitation assumption on $\omega(t)$. Readers can refer to [31, 114, 116] for details on the construction of $v_1(t), v_2(t)$ and the proof of stability.

6.1.2 A Typical Nonlinear Adaptive Observer

In this section, we use a typical nonlinear observer design to exemplify a straightforward application of linear adaptive observer design methods. The adaptive observer design for a class of single-output systems considered in [105]

$$\begin{aligned} \dot{x} &= f(x) + q_0(x, u) + \sum_{i=1}^{m} \theta_i q_i(x, u), \\ y &= h(x), \end{aligned} \tag{6.5}$$

where $u \in \mathbb{R}^l, y \in \mathbb{R}, x \in \mathbb{R}^n$, $q_i : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n$, and $\theta_i \in \mathbb{R}$ are unknown constant parameters. Necessary and sufficient conditions are given in [105] to transform system (6.5) into an AOF

$$\begin{aligned} \dot{z} &= Az + \psi_0(y, u) + \sum_{i=1}^{m} \theta_i \beta_i(y, u, t) b, \\ y &= Cz, \end{aligned} \tag{6.6}$$

where A, C are defined in (1.7), $b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$, $b_n > 0$, and the polynomial $b_1 + b_2 s + \cdots + b_n s^{n-1}$ is Hurwitz, $\beta_i(y, u, t) : \mathbb{R}^1 \times \mathbb{R}^l \times \mathbb{R}^+ \to \mathbb{R}$. The restrictions on b and θ_i make the form (6.6) more conservative than that in [11]. Taking an observer

$$\begin{aligned} \dot{\hat{z}} &= A\hat{z} + \psi_0(y, u) + \sum_{i=1}^{m} \hat{\theta}_i \beta_i(y, u, t) b + L(y - C\hat{z}), \\ \hat{y} &= C\hat{z} \end{aligned} \tag{6.7}$$

yields the error dynamics

$$\dot{\tilde{z}} = (A - LC)\tilde{z} + \sum_{i=1}^{m} \tilde{\theta}_i \beta_i(y, u, t) b,$$

where $\tilde{\theta}_i = \theta_i - \hat{\theta}_i$. Since b is Hurwitz and (A, C) is observable, the transfer function $C(sI - A + LC)^{-1} b$ can be made SPR by choosing a certain L. If $\beta(y, u, t)$ is bounded for any bounded u, y, there exists an adaptive observer that consists of the observer (6.7) and the adaptive law

$$\dot{\hat{\theta}} = -\Gamma \beta(y, u, t) \tilde{y}. \tag{6.8}$$

The proposed adaptive observer gives the asymptotic estimate of z without the need of auxiliary signals $v_1(t), v_2(t)$ as in Section 6.1.1 or any persistent excitation condition as in [11]. Evidently, the adaptive law (6.8) and its associated stability analysis is closely related to the linear system case in Section 6.1.1.

Recall the similar approach taken in the previous linear adaptive observer design case, where auxiliary signals $v_1(t), v_2(t)$ are used to obtain an error dynamics whose output approaches that of (6.4) asymptotically. Later work in [103, 109] introduced a filtered state transformation to put

$$\begin{aligned} \dot{z} &= Az + \psi_0(y, u) + \psi_1(y, u)\theta, \\ y &= Cz \end{aligned}$$

into AOF (6.6). We will see the application of the filtered state transformation in the adaptive observer design in Sections 6.2 and 6.3.

6.2 Adaptive Observer Design Based on a LP NOOF

We first present the definition of a local adaptive observer based on a slight modification of a global adaptive observer [109].

Definition 6.2.1. A local adaptive observer for system (2.1) with the presence of the unknown parameter θ in f is a finite dimensional system

$$\begin{aligned} \dot{w} &= \alpha_1(w,\hat{\theta},y(t)), \quad w \in \mathbb{R}^r,\ r \geqslant n, \\ \dot{\hat{\theta}} &= \alpha_2(w,\hat{\theta},y(t)), \quad \hat{\theta} \in \mathbb{R}^m, \\ \hat{\zeta} &= \alpha_3(w,\hat{\theta},y(t)), \quad \hat{\zeta} \in \mathbb{R}^n \end{aligned}$$

driven by $y(t)$, such that for every $\zeta(0) \in \mathbb{R}^n, w(0) \in U_w \subset \mathbb{R}^r, \hat{\theta}(0) \in U_\theta \subset \mathbb{R}^m$, where U_w, U_θ are the neighborhoods of $\zeta(0), \theta$ respectively, for any value of the unknown parameter θ and for any bounded $\|\zeta(t)\|, \forall t \geqslant 0$:

(i) $\|w(t)\|, \|\hat{\theta}(t)\|$ and $\|\zeta(t) - \hat{\zeta}(t)\|$ are bounded, $\forall t \geqslant 0$.

(ii) $\lim_{t\to\infty} \|\zeta(t) - \hat{\zeta}(t)\| = 0$.

□

Since the state transformation from ζ-coordinates to z-coordinates is locally Lipschitz, the convergence of $\hat{\zeta}(t)$ to $\zeta(t)$ as $t \to \infty$ is guaranteed if $\hat{z}(t) \to z(t)$ as $t \to \infty$. We therefore focus on the adaptive observer design on the z coordinates.

Given the NOOF

$$\begin{aligned} \dot{z} &= Az + \gamma(y), \\ y &= h(z), \end{aligned}$$

we extend it to allow the linear presence of unknown parameters θ

$$\begin{aligned} \dot{z} &= Az + \gamma(y) + \beta(y)\theta, \\ y &= h(z). \end{aligned} \tag{6.9}$$

To simplify the stability analysis of the error dynamics below, we introduce the filtered state transformation which relies on a stable filter

$$\dot{M} = AM + \beta(y). \tag{6.10}$$

We first make an assumption and recall a lemma, then present the local adaptive observer.

Assumption 6.2.2. The auxiliary signal $M(t)$ generated by (6.10) is Persistently Excited (PE), i.e., there exist positive constants α, β, T such that

$$\alpha \mathbf{I} \leqslant \int_t^{t+T} M^T(\tau) C^T C M(\tau) \mathrm{d}\tau \leqslant \beta \mathbf{I}, \quad \forall t \geqslant t_0.$$

□

Lemma 6.2.3. *[1] Let $\phi(t) \in \mathbb{R}^{m\times p}$ be a bounded and piecewise continuous matrix and $\Gamma \in \mathbb{R}^{p\times p}$ be any symmetric positive-definite matrix. If there exist positive constants T, α, β such that*

$$\alpha \mathbf{I}_p \leqslant \int_t^{t+T} x^T \phi^T(\tau)\phi(\tau) x \mathrm{d}\tau \leqslant \beta \mathbf{I}_p, \quad \forall t \geqslant t_0, \|x\| = 1$$

then the system

$$\dot{x} = -\Gamma \phi^T(t)\phi(t)x$$

is GES.

Proof: Taking the Lyapunov function candidate $V(x) = \frac{1}{2}x^T\Gamma^{-1}x$, we compute its time derivative

$$\dot{V} = -x^T\phi(t)^T\phi(t)x.$$

Integrating $\dot{V}$ from t to $t+T$ gives

$$\begin{aligned} V(x(t+T)) - V(x(t)) &= -\int_t^{t+T} x^T(\tau)\phi^T(\tau)\phi(\tau)x(\tau)\mathrm{d}\tau \\ &\leqslant -\alpha x^T x \leqslant -\frac{\alpha}{2\mu}V(x(t)), \end{aligned}$$

where μ is the maximum eigenvalue of Γ. Applying [76, Thm. 8.5] we conclude that the system is GES. □

As shown in Lemma 6.2.3, given Assumption 6.2.2, the zero solution of the LTV system

$$\dot{x} = -\Gamma M^T C^T C M x \tag{6.11}$$

is GES. The existence of the Lyapunov function candidate V is ensured by the following theorem such that

$$\begin{aligned} \frac{\partial V}{\partial t} - \frac{\partial V}{\partial x} M^T C^T C M x \leqslant -\epsilon \|x\|^2 \\ \left\| \frac{\partial V}{\partial x} \right\| \leqslant \alpha_4 \|x\| \end{aligned}$$

where $\epsilon, \alpha_4 > 0$.

Theorem 6.2.4. *[138, Thm. 1.5.1] Assume that $f(t,x) : \mathbb{R}_+ \to \mathbb{R}^n$ has continuous and bounded first partial derivatives in x and is piecewise continuous in t for all $x \in B_h, t \geqslant 0$. Then the following statements are equivalent:*

1. $x = 0$ *is an exponentially stable equilibrium point of*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

2. There exists a function $v(t, x)$, *and some strictly positive constant* $h', \alpha_1, \alpha_2, \alpha_3$, *and* α_4, *such that, for all* $x \in B_{h'}, t \geqslant 0$

$$\begin{aligned}
\alpha_1 \|x\|^2 \leqslant v(t, x) \leqslant \alpha_2 \|x\|^2 \\
\frac{dv(t, x)}{dt} \leqslant -\alpha_3 \|x\|^2 \\
\left\| \frac{\partial v(t, x)}{\partial x} \right\| \leqslant \alpha_4 \|x\|
\end{aligned}$$

Theorem 6.2.5. *Given a locally observable system in LP NOOF* (6.9) *and Assumption 6.2.2 holds with* $C(t) = \frac{\partial h(\hat{z})}{\partial \hat{z}}$, *the system has a local adaptive observer*

$$\begin{aligned}
\dot{\hat{z}} &= A\hat{z} + \gamma(y) + \beta(y)\hat{\theta} + M\dot{\hat{\theta}}, \\
\dot{M} &= AM + \beta(y), \\
\dot{\hat{\theta}} &= M^T \left(\frac{\partial h(\hat{z})}{\partial \hat{z}} \right)^T (y - \hat{y}),
\end{aligned}$$

where $\hat{y} = h(\hat{z})$.

Proof: We apply the filtered state transformation $\eta = z - M\theta$ and have

$$\dot{\eta} = A\eta + \gamma(y).$$

Defining $\tilde{\theta} = \theta - \hat{\theta}, \hat{\eta} = \hat{z} + M\hat{\theta}$, and $\tilde{\eta} = \eta - \hat{\eta}$, the error dynamics are given by

$$\begin{aligned}
\dot{\tilde{\eta}} &= A\tilde{\eta}, \\
\dot{\tilde{\theta}} &= -M^T \left(\frac{\partial h(\hat{z})}{\partial \hat{z}} \right)^T (y - \hat{y}).
\end{aligned} \tag{6.12}$$

Using the Taylor expansion of y along the trajectory of $\hat{z}(t)$, the output error is reformulated as

$$\tilde{y} = h(z) - h(\hat{z}) = \frac{\partial h(\hat{z})}{\partial \hat{z}} \tilde{z} + \frac{1}{2} \tilde{z}^T \frac{\partial^2 h(\xi)}{\partial \xi^2} \tilde{z}, \quad \xi \in [\hat{z}, z],$$

where $\xi = (\xi_1, \ldots, \xi_n)^T$, $\xi \in [\hat{z}, z]$ is used to denote $\xi_k \in [\hat{z}_k, z_k]$ for $1 \leqslant k \leqslant n$, and $\hat{z}_k \leqslant z_k$ is assumed to simplify the presentation. Given Assumption 6.2.2, the LTV part of $\tilde{\theta}$-subsystem is GES. From Theorem 6.2.4, we know there exists a function $V_{\tilde{\theta}}$ such that

$$\frac{\partial V_{\tilde{\theta}}}{\partial t} - \frac{\partial V_{\tilde{\theta}}}{\partial \tilde{\theta}} M^T \left(\frac{\partial h(\hat{z})}{\partial \hat{z}}\right)^T \frac{\partial h(\hat{z})}{\partial \hat{z}} M \tilde{\theta} \leqslant -\epsilon \left\|\tilde{\theta}\right\|^2,$$
$$\left\|\frac{\partial V_{\tilde{\theta}}}{\partial \tilde{\theta}}\right\| \leqslant k_{\tilde{\theta}} \left\|\tilde{\theta}\right\|,$$

where $\epsilon > 0, k_{\tilde{\theta}} \geqslant 0$. Taking the Lyapunov function candidate $V = k\tilde{\eta}^T P \tilde{\eta} + V_{\tilde{\theta}}, k > 0, P > 0$, we have its time derivative

$$\begin{aligned}
\dot{V} =& - k\tilde{\eta}^T Q \tilde{\eta} + \frac{\partial V_{\tilde{\theta}}}{\partial t} - \frac{\partial V_{\tilde{\theta}}}{\partial \tilde{\theta}} M^T \left(\frac{\partial h(\hat{z})}{\partial \hat{z}}\right)^T \frac{\partial h(\hat{z})}{\partial \hat{z}} M \tilde{\theta} \\
&- \frac{\partial V_{\tilde{\theta}}}{\partial \tilde{\theta}} M^T \left(\frac{\partial h(\hat{z})}{\partial \hat{z}}\right)^T \frac{1}{2} \tilde{z}^T \frac{\partial^2 h(\xi)}{\partial \xi^2} \tilde{z} \\
\leqslant& - k\tilde{\eta}^T Q \tilde{\eta} - \epsilon \tilde{\theta}^T \tilde{\theta} + \frac{1}{2} \left\|\tilde{\theta}\right\| k_{\tilde{\theta}} k_M k_{h1} k_{h2} \tilde{z}^T \tilde{z},
\end{aligned}$$

where $\|M(t)\| \leqslant k_M$, $\left\|\frac{\partial h(\hat{z})}{\partial \hat{z}}\right\| \leqslant k_{h1}$, and $\left\|\frac{\partial^2 h(\xi)}{\partial \xi^2}\right\| \leqslant k_{h2}$. Expanding $\tilde{z}^T \tilde{z}$ gives

$$\begin{aligned}
\tilde{z}^T \tilde{z} &= (\tilde{\eta} + M\tilde{\theta})^T (\tilde{\eta} + M\tilde{\theta}) \\
&= \|\tilde{\eta}\|^2 + \tilde{\theta}^T M^T M \tilde{\theta} + 2\tilde{\eta}^T M \tilde{\theta} \\
&\leqslant \|\tilde{\eta}\|^2 + k_M^2 \left\|\tilde{\theta}\right\|^2 + 2k_M \|\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|.
\end{aligned}$$

$\dot{V}$ can be simplified as

$$\begin{aligned}
\dot{V} \leqslant& - (k\mu - 1) \|\tilde{\eta}\|^2 - \epsilon \left\|\tilde{\theta}\right\|^2 \\
&+ \frac{k_{\tilde{\theta}} k_M k_{h1} k_{h2}}{2} \left\|\tilde{\theta}\right\| (\|\tilde{\eta}\|^2 + k_M^2 \left\|\tilde{\theta}\right\|^2 + 2k_M \|\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|) \\
\leqslant& - (k\mu - 1) \|\tilde{\eta}\|^2 - \epsilon \left\|\tilde{\theta}\right\|^2 + \frac{\epsilon}{4} \left\|\tilde{\theta}\right\|^2 + \frac{k_{\tilde{\theta}}^2 k_M^2 k_{h1}^2 k_{h2}^2}{4\epsilon} \|\tilde{\eta}\|^4 \\
&+ \frac{k_{\tilde{\theta}} k_M^3 k_{h1} k_{h2}}{2} \left\|\tilde{\theta}\right\|^3 + \frac{\epsilon}{4} \left\|\tilde{\theta}\right\|^4 + \frac{k_{\tilde{\theta}}^2 k_M^4 k_{h1}^2 k_{h2}^2}{\epsilon} \|\tilde{\eta}\|^2 \\
\leqslant& - (k\mu - 1 - \frac{k_{\tilde{\theta}}^2 k_M^4 k_{h1}^2 k_{h2}^2}{\epsilon} - \frac{k_{\tilde{\theta}}^2 k_M^2 k_{h1}^2 k_{h2}^2}{4\epsilon} \|\tilde{\eta}\|^2) \|\tilde{\eta}\|^2 \\
&- \frac{\epsilon}{4} \left\|\tilde{\theta}\right\|^2 - (\frac{\epsilon}{4} - \frac{k_{\tilde{\theta}} k_M^3 k_{h1} k_{h2}}{2} \left\|\tilde{\theta}\right\|) \left\|\tilde{\theta}\right\|^2 - \frac{\epsilon}{4}(\left\|\tilde{\theta}\right\|^2 - \left\|\tilde{\theta}\right\|^4),
\end{aligned}$$

where μ is the minimum eigenvalue of Q. Given the initial error $\|\tilde{\eta}\| < \infty$, there always exists a sufficiently large $k > 0$ such that

$$k\mu - 1 - \frac{k_{\tilde{\theta}}^2 k_M^4 k_{h1}^2 k_{h2}^2}{\epsilon} - \frac{k_{\tilde{\theta}}^2 k_M^2 k_{h1}^2 k_{h2}^2}{4\epsilon} \|\tilde{\eta}\|^2 > 0.$$

We conclude that $\dot{V}$ is negative definite if the initial error of $\tilde{\theta}$ satisfies

$$\left\|\tilde{\theta}\right\| \leqslant \frac{\epsilon}{2k_{\tilde{\theta}}k_M^3 k_{h1}k_{h2}}. \tag{6.13}$$

Local exponential stability of the error dynamics (6.12) is therefore concluded. □

If system (6.9) has linear outputs, which corresponds to $k_{h2} = 0$, Condition (6.13) becomes $\left\|\tilde{\theta}\right\| < \infty$. This fact implies that GES of state and parameter estimation errors is guaranteed; we thus recover the previous results in work [105, 103] as a special case.

6.3 Adaptive Observer Design Based on a NLP OF

In this section, we consider local adaptive observer design for nonlinear systems having NLP dynamics and linear output maps. This class of systems is more general than that considered in [105, 103, 104, 171]. As indicated below, applying the same filtered state transformation to a NLP system cannot yield a partially decoupled error dynamics, i.e., $\tilde{\eta}, \tilde{\theta}$ still affect each other. Compared to the LP case, where the error dynamics is simplified to be LTV by the filtered state transformation, the error dynamics of a NLP system is NLTV. We first transform a system (2.1) into a NLP OF (6.1), then consider the adaptive observer design for (6.1). The existence conditions of a parameter independent state transformation from (2.1) to (6.1) are given in the following theorem. Its proof is omitted since it is straightforward.

Theorem 6.3.1. *There exists a local diffeomorphism in a neighborhood U_0 of the origin,*

$$z = \Phi(\zeta), \quad z \in \mathbb{R}^n$$

transforming system (2.1) *into* (6.1) *if and only if in U_0 in addition to Conditions (i)-(iii) of Theorem A.0.2,*

$$\partial \mathrm{ad}^k_{-f} g_i / \partial \theta_j = 0, \quad \begin{cases} 0 \leqslant k \leqslant \lambda_i - 1; \\ 1 \leqslant i \leqslant p; \\ 1 \leqslant j \leqslant m. \end{cases} \tag{6.14}$$

The state transformation $\Phi(\zeta)$ is solved from PDEs (A.3).

Remark 6.3.2. Conditions (6.14) ensure the state transformation is independent of parameter vector θ. This is because the state transformation can be constructed by the composition of flows of the vector fields $\mathrm{ad}^k_{-f} g_i$. If none of $\mathrm{ad}^k_{-f} g_i$ has parameter dependence, neither the composition of flows nor the state transformation does. The existence conditions of a parameter dependent diffeomorphism locally transforming system (2.1) into (6.1) can be similarly presented. To simplify the presentation, we only consider a parameter independent state transformation. □

As shown in Section 6.2, a filtered state transformation removes the effect of parameter error $\tilde{\theta}$ from the $\tilde{\eta}$ dynamics. This requires a stable filter driven by the coefficients of the unknown parameter. For NLP systems, we use a Taylor series to approximate the term $\gamma(y, \theta)$ along the trajectory of parameter estimate and define the stable filter

$$\dot{\hat{M}} = (A - LC)\hat{M} + \frac{\partial \hat{\gamma}}{\partial \hat{\theta}}, \tag{6.15}$$

where $\hat{\gamma} = \gamma(y, \hat{\theta})$, and L is an observer gain matrix such that $A - LC$ is Hurwitz. We propose the following theorem which provides a local adaptive observer for system (6.1).

Theorem 6.3.3. *Given Assumption 6.2.2, with $M(t)$ replaced by $\hat{M}(t)$ in* (6.15), *and considering the system*

$$\begin{aligned}
\dot{\hat{z}} &= A\hat{z} + \hat{\gamma} + L(y - C\hat{z}) + \hat{M}\dot{\hat{\theta}}, \\
\dot{\hat{M}} &= (A - LC)\hat{M} + \frac{\partial \hat{\gamma}}{\partial \hat{\theta}}, \\
\dot{\hat{\theta}} &= \hat{M}^T C^T (y - C\hat{z}),
\end{aligned}$$

where $L \in \mathbb{R}^{n \times p}$ is a constant observer gain matrix such that $A - LC$ is Hurwitz, the above system is a local adaptive observer for (6.1).

Proof: Denoting the errors $\tilde{z} = z - \hat{z}, \tilde{\theta} = \theta - \hat{\theta}$, we have the error dynamics

$$\begin{aligned}
\dot{\tilde{z}} &= (A - LC)\tilde{z} + \gamma - \hat{\gamma} - \hat{M}\dot{\hat{\theta}}, \\
\dot{\hat{M}} &= (A - LC)\hat{M} + \frac{\partial \hat{\gamma}}{\partial \hat{\theta}}, \\
\dot{\tilde{\theta}} &= -\hat{M}^T C^T (y - \hat{y}).
\end{aligned}$$

As in Section 6.2, we introduce the filtered state transformation

$$\eta = z - \hat{M}\theta, \quad \hat{\eta} = \hat{z} - \hat{M}\hat{\theta},$$

and parameterize the output error $\tilde{y} = y - C\hat{z}$ as

$$\tilde{y} = C(\tilde{\eta} + \hat{M}\tilde{\theta}),$$

where $\tilde{\eta} = \eta - \hat{\eta}$. We perform the stability analysis in the neighborhoods of $\theta, z(0)$, denoted by $B_\theta \subset \mathbb{R}^m, B_z \subset \mathbb{R}^n$ respectively. Given that $\hat{\theta}(t), y(t)$ remain in B_θ and B_z, $\frac{\partial \hat{\gamma}}{\partial \hat{\theta}}$ and $\hat{M}(t)$ are therefore bounded. The dynamics of $\tilde{\eta}$ and $\tilde{\theta}$ are given by

$$\dot{\tilde{\eta}} = (A - LC)\tilde{\eta} + (\gamma - \hat{\gamma} - \frac{\partial \hat{\gamma}}{\partial \hat{\theta}}\tilde{\theta}),$$
$$\dot{\tilde{\theta}} = -\hat{M}^T C^T C \hat{M} \tilde{\theta} - \hat{M}^T C^T C \tilde{\eta}.$$

Similar to the proof of Theorem 6.2.5, we take the Lyapunov function candidate

$$V(\tilde{\eta}, \tilde{\theta}) = \tilde{\eta}^T P \tilde{\eta} + V_{\tilde{\theta}}.$$

A positive definite matrix P is to be determined such that

$$P(A - LC) + (A - LC)^T P = -Q,$$

where $Q > 0$, and the minimum eigenvalue of Q is sufficiently large. Given an arbitrary Q, such a P always exists if $A - LC$ is Hurwitz. Since γ is smooth and locally has a Taylor expansion along the trajectory of $\hat{\theta}(t)$, we have

$$\gamma - \hat{\gamma} - \frac{\partial \hat{\gamma}}{\partial \hat{\theta}}\tilde{\theta} \leqslant k_\gamma \left\|\tilde{\theta}\right\|^2,$$

where k_γ is the upper bound of $\frac{\partial^2 \hat{\gamma}(\xi)}{\partial \xi^2}$ in B_z, B_θ. Also, for the $\tilde{\eta}$-subsystem, its homogenous part is GES. The time derivative of the Lyapunov function candidate is given by

$$\begin{aligned}
\dot{V} &\leqslant \tilde{\eta}^T Q \tilde{\eta} + 2k_\gamma \|P\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|^2 - \epsilon \left\|\tilde{\theta}\right\|^2 - \frac{\partial V_{\tilde{\theta}}}{\partial \tilde{\theta}} \hat{M}^T C^T C \tilde{\eta} \\
&\leqslant -k_1 \|\tilde{\eta}\|^2 + 2k_\gamma \|P\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|^2 - \epsilon \left\|\tilde{\theta}\right\|^2 + k_{\tilde{\theta}} \left\|\hat{M}C^T C\right\| \cdot \left\|\tilde{\theta}\right\| \cdot \|\tilde{\eta}\| \\
&\leqslant -k_1 \|\tilde{\eta}\|^2 + 2k_\gamma \|P\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|^2 - \epsilon \left\|\tilde{\theta}\right\|^2 + \frac{\epsilon}{2} \left\|\tilde{\theta}\right\|^2 + \frac{k_{\tilde{\theta}}^2 k_{\hat{M}}^2}{2\epsilon} \|\tilde{\eta}\|^2,
\end{aligned}$$

where $k_1 > 0$ can be taken as the minimum eigenvalue of Q, $\left\|\hat{M}^T C^T C\right\| \leqslant k_{\hat{M}}$ since $\hat{\gamma}$ is bounded in B_θ and B_z. Since the order of $\|P\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|^2$ is higher than those of the other terms, this term can be ignored by assuming sufficiently small initial errors $\tilde{\theta}(0), \tilde{\eta}(0)$. Thus we simplify the time derivative of the Lyapunov function candidate to

$$\dot{V} \leqslant -(k_1 - \frac{k_{\tilde{\theta}}^2 k_{\hat{M}}^2}{2\epsilon}) \|\tilde{\eta}\|^2 - \frac{\epsilon}{2} \left\|\tilde{\theta}\right\|^2.$$

By choosing an appropriate matrix P such that the inequality holds

$$k_1 > \frac{k_{\tilde{\theta}}^2 k_{\hat{M}}^2}{2\epsilon},$$

we have $\dot{V}$ is negative definite. The locally exponential convergence of the zero solution of the error dynamics $\tilde{\eta}, \tilde{\theta}$ is therefore established. $\square$

Remark 6.3.4. Given the neighborhoods B_θ and B_z, Assumption 6.2.2 should be satisfied for any $y \in B_z, \hat{\theta} \in B_\theta$. To ensure the dropped term $2\|P\tilde{\eta}\| \cdot \left\|\tilde{\theta}\right\|$ is sufficiently small, we must impose conditions on the initial errors of $\tilde{\eta}$ or $\tilde{\theta}$. That is, a large $\tilde{\eta}(0)$ results in a very small range of $\tilde{\theta}(0)$ allowed. This is different from the adaptive observer design based on a NOOF, where $\tilde{\eta}(0)$ can be large and still not affect the permissible range of $\tilde{\theta}(0)$. $\square$

6.4 An Observer Design Example

We consider a two-output system

$$\begin{aligned} \dot{x} &= \begin{pmatrix} x_2^1 - 2x_1^1 \\ x_1^1 x_2^1 + \sin(x_1^1\theta_1) - \frac{x_1^1}{2} \\ x_2^2 - (1 + e^{x_1^1\theta_2})x_1^2 + x_1^1 \\ \cos(x_1^2\theta_1) - x_1^1 - x_1^2 + \sin(\theta_2 x_1^2) \end{pmatrix}, \\ y &= Cx = \begin{pmatrix} x_1^1 \\ x_1^2 \end{pmatrix}, \end{aligned} \tag{6.16}$$

where θ_1, θ_2 are unknown parameters. It is easy to verify system (6.16) is observable with indices $(2, 2)$. System (6.16) is NLP, thus we first apply Theorem 6.3.1 to check if the system is transformable to the form (6.1). We solve the starting vectors $g^1 = \partial/\partial x_2^1, g^2 = \partial/\partial x_2^2$, and verify that the other conditions in Theorem 6.3.1 are satisfied. System (6.16) can be put into the form (6.1). Applying the state transformation $z = \Phi(x)$ solved from (A.3)

$$z = \begin{pmatrix} x_1^1 \\ x_2^1 - \frac{1}{2}(x_1^1)^2 \\ x_1^2 \\ x_2^2 \end{pmatrix},$$

we rewrite the system (6.16) in z-coordinates

$$\dot{z} = Az + \begin{pmatrix} \frac{1}{2}(y_1)^2 - 2y_1 \\ \sin(y_1\theta_1) - \frac{1}{2}y_1 \\ y_1 - (1 + e^{y_1\theta_2})y_2 \\ \cos(y_2\theta_1) - y_1 - y_2 + \sin(\theta_2 y_2) \end{pmatrix},$$

$$y = Cz = \begin{pmatrix} z_1^1 \\ z_1^2 \end{pmatrix}.$$

We take the adaptive observer as proposed in Theorem 6.3.3

$$\begin{aligned} \dot{\hat{z}} =& A\hat{z} + \begin{pmatrix} \frac{1}{2}y_1^2 - 2y_1 \\ -\frac{1}{2}y_1 \\ y_1 - y_2 \\ -y_1 - y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin(y_1\hat{\theta}_1) \\ -e^{y_1\hat{\theta}_2}y_2 \\ \cos(y_2\hat{\theta}_1) + \sin(\hat{\theta}_2 y_2) \end{pmatrix} \\ &+ L(y - \hat{y}) + M(t)\dot{\hat{\theta}}, \\ \dot{\hat{M}}(t) =& (A - LC)\hat{M}(t) + \begin{pmatrix} 0 & 0 \\ \cos(y_1\hat{\theta}_1)y_1 & 0 \\ 0 & -e^{y_1\hat{\theta}_2}y_1y_2 \\ -\sin(y_2\hat{\theta}_1)y_2 & \cos(\hat{\theta}_2 y_2)y_2 \end{pmatrix}, \\ \dot{\hat{\theta}} =& \hat{M}^T(t)C^T(y - \hat{y}), \end{aligned}$$

where $\theta = (\theta_1, \theta_2)^T$, and

$$L = \begin{pmatrix} 0 & 0 & l_1^2 & l_2^2 \\ l_1^1 & l_2^1 & 0 & 0 \end{pmatrix}^T$$

is the constant observer gain matrix. Assuming $\theta_1 = 1, \theta_2 = 0.5$ and taking $l_j^k = 4, k, j = 1, 2$ to place all eigenvalues of $A - LC$ at -2, with the initial conditions (ICs) $z(0) = (0.3, 0.5, 0.2, 0.3)^T, \hat{M}(0) = 0$, and

$$IC_1 : \hat{z}(0) = (1.2, 1.4, -0.3, 1.2)^T, \quad \hat{\theta}(0) = (0.8, 0.4)^T,$$
$$IC_2 : \hat{z}(0) = (1, 1, -0.3, 0.6)^T, \quad \hat{\theta}(0) = (0.8, 0.4)^T,$$
$$IC_3 : \hat{z}(0) = (0.2, 0.4, 0.3, 0.2)^T, \quad \hat{\theta}(0) = (0.5, 0.4)^T,$$

we have the simulation results as shown in Figures 6.1, 6.2, and 6.3 respectively. The simulation demonstrates that for the error dynamics generated by the adaptive observer in Theorem 6.3.3, the convergence of its zero solution is affected not only by the ICs of $\tilde{\theta}$ but also by the ICs of $\tilde{z}$, and that PE is not sufficient to guarantee the convergence of the error dynamics.

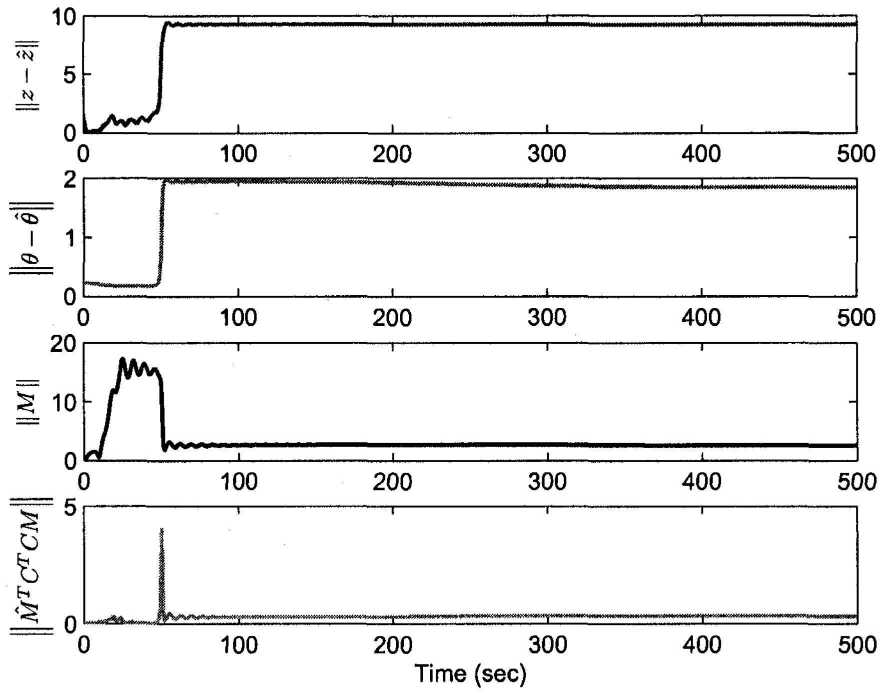

Figure 6.1: Simulation of adaptive observer with IC_1

6.5 Summary

In this chapter, we first considered the adaptive observer design for a LP system in NOOF. Then adaptive observer design for a class of NLP systems admitting OF was studied. Local exponential stability of the error dynamics for both cases was shown. The previous results in [105, 103] were recovered as a special case. The stability of the error dynamics was established using a filtered state transformation, and a PE assumption. The PE assumption was defined on signals which depend not only on the measurements (y, u) but also on unknown parameter estimates $\hat{\theta}$.

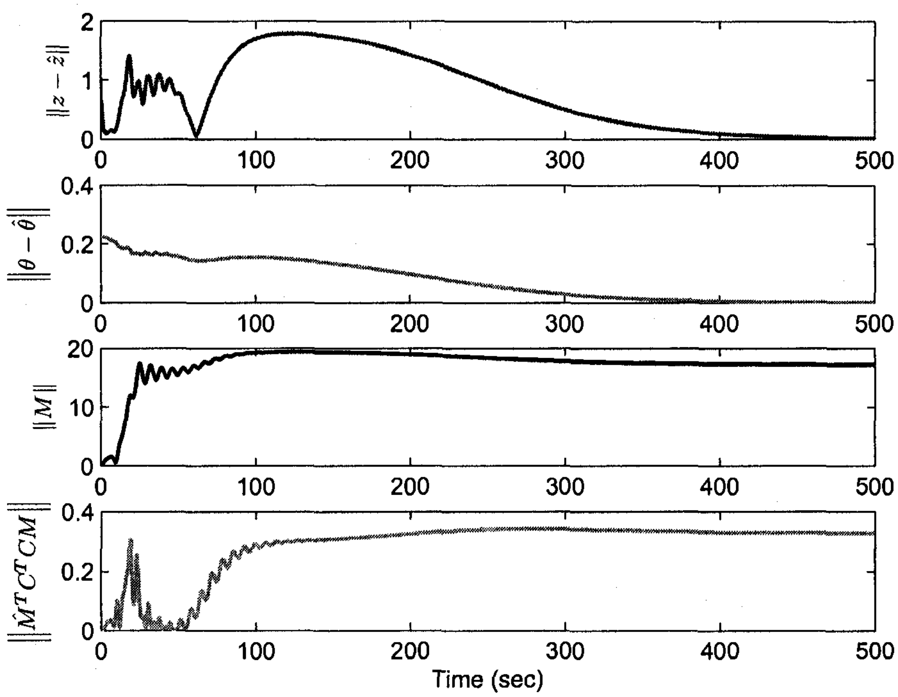

Figure 6.2: Simulation of adaptive observer with IC_2

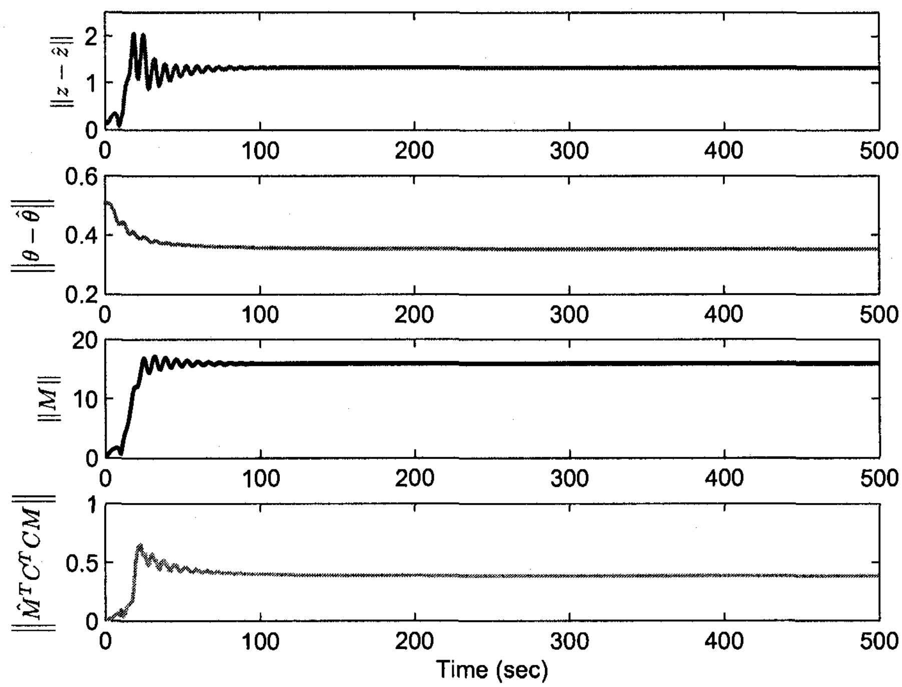

Figure 6.3: Simulation of adaptive observer with IC_3

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis considered observer design for uncontrolled multi-output nonlinear systems using geometric methods. The main contributions lie in four areas: existence conditions for special forms; extensions of OF and BTOF using time scale transformations, observer designs based on the proposed forms and the stability analysis of the resulting error dynamics, adaptive observer designs for two classes of parameterized systems. The main conclusions can be summarized as follows

(i) The existence conditions for a change of coordinates transforming the ith subsystem into BTF or BTOF were presented. An entire system is transformed to BTF or BTOF subsystem-at-a-time by applying these conditions sequentially.

(ii) We extended the OF and BTOF by introducing multiple time scale transformations. Output dependent TSFs were taken in the TOF case. Generalized TSFs, which can be functions of upper subsystems' state and output, were adopted for the TBTOF case. The existence conditions for both cases were established.

(iii) We presented observer designs based on BTOF, TOF, and TBTOF, and performed the stability analysis of the resulting error dynamics. For a BTOF-based observer design, the zero solution of the error dynamics was shown to be GES, GUAS, and semi-globally GES under different conditions. For the TOF and TBTOF cases, we presented necessary and sufficient conditions on TSFs to preserve stability in different time scales.

(iv) We investigated adaptive observer designs for two classes of systems: those transformable to a LP NOOF or a NLP OF. For each class of systems, an adaptive observer was proposed and the resulting error dynamics was shown to be locally GES with a PE assumption.

7.2 Future Work

(i) There are two problems left regarding BTF existence conditions. The first is to develop simplified methods to avoid solving high order PDEs in Theorem 2.3.6. The second is to remove the parameter assumption made when deriving BTF existence conditions in Section 2.3.

(ii) As with the BTF, a more constructive algorithm for computing the BTOF coordinates using Theorem 3.3.1 could be developed.

(iii) As shown in Chapter 2, the complexity of observer design can be reduced by transforming an observable system into BTF. However, currently BTF-based observer design cannot be described in general. Lack of systematic observer design approaches restricts the applicability of the BTF.

(iv) Considering the benefits of BTF and BTOF-based observer design, a question that arises is "can a BTF or BTOF be useful in simplifying control design?". This question is natural given that normal forms are commonly used in nonlinear control, e.g. state feedback stabilization [28, 29, 35, 36, 34], output regulation by state feedback [65, 27], stabilization by output feedback [106, 107, 108], state feedback stabilization by backstepping [87] etc. Particularly relevant work is in [149] where dynamic output feedback stabilizes a restrictive triangular form. Investigating the generalization of this work to a BTF structure would be a logical starting point.

(v) Adaptive observer design for NLP systems is another challenging but interesting problem. Many interesting results on adaptive state feedback control of NLP systems have recently become available. For example, adaptive control of systems having convex or concave parameterizations [78, 142, 79, 3, 9], adaptive control for partially feedback linearizable systems [94], etc. Given these recent developments in adaptive control, it is natural to investigate solving an adaptive observer problem with similar techniques.

Appendix A

Multi-output Observer Form

Multi-output Observer Form relies on the uniquely defined observability indices

Definition A.0.1. [109, Def. 5.4.1] A set of observability indices $\{\lambda_1, \cdots, \lambda_p\}$ is uniquely associated to system (2.1) as follows:

$$\lambda_i = \text{card}\ \{s_j \geqslant i : j \geqslant 0\}, \quad 1 \leqslant i \leqslant p$$

with

$$\begin{aligned}
s_0 &= \text{rank}\ \{\mathrm{d}h_i(\zeta) : 1 \leqslant i \leqslant p\} \\
&\ \vdots \\
s_k &= \text{rank}\ \{\mathrm{d}h_i(\zeta), \cdots, \mathrm{d}L_f^k h_i(\zeta) : 1 \leqslant i \leqslant p\} \\
&\quad - \text{rank}\ \{\mathrm{d}h_i(\zeta), \cdots, \mathrm{d}L_f^{k-1} h_i(\zeta) : 1 \leqslant i \leqslant p\} \\
&\ \vdots \\
s_{n-1} &= \text{rank}\ \{\mathrm{d}h_i(\zeta), \cdots, \mathrm{d}L_f^{n-1} h_i(\zeta) : 1 \leqslant i \leqslant p\} \\
&\quad - \text{rank}\ \{\mathrm{d}h_i(\zeta), \cdots, \mathrm{d}L_f^{n-2} h_i(\zeta) : 1 \leqslant i \leqslant p\}
\end{aligned}$$

□

The existence conditions of (1.5) are given by the following theorem.

Theorem A.0.2. *[109, Thm. 5.4.1] There exists a local diffeomorphism*

$$z = T(\zeta), \quad z \in U_0,$$

transforming system (1.4)*, up to a reordering of the outputs* $y_1, \cdots, y_p$*, into an OF* (1.5) *if, and only if, in* U_0

(i) the system is locally observable and $\lambda_1, \cdots, \lambda_p$ *are constant;*

(ii) $Q_i = Q_i \bigcap Q$ *where the two co-distributions* Q_i, Q *are given, for* $1 \leqslant i \leqslant p$,

$$Q_i = \mathrm{span}\{\mathrm{d}L_f^k h_r, 0 \leqslant k \leqslant \lambda_i - 1, 1 \leqslant r \leqslant p\} \backslash \mathrm{d}L_f^{\lambda_i - 1} h_i,$$
$$Q = \mathrm{span}\{\mathrm{d}L_f^k h_r, 0 \leqslant k \leqslant \lambda_r - 1, 1 \leqslant r \leqslant p\};$$

(iii) the vector fields $g^1, \cdots, g^p$ *satisfying*

$$L_{g^i} L_f^{k-1} h_j = \delta_{i,j} \delta_{k,\lambda_i - 1}, \quad \begin{cases} 1 \leqslant i, j \leqslant p; \\ 0 \leqslant k \leqslant \lambda_i - 1, \end{cases} \tag{A.1}$$

such that

$$\left[\mathrm{ad}_{-f}^k g^i, \mathrm{ad}_{-f}^l g^j \right] = 0, \quad \begin{cases} 1 \leqslant i, j \leqslant p; \\ 0 \leqslant k \leqslant \lambda_i - 1; \\ 0 \leqslant l \leqslant \lambda_j - 1. \end{cases} \tag{A.2}$$

There exists a global diffeomorphism if, and only if, Conditions (i)–(iii) *holds in* $\mathbb{R}^n$ *and, in addition, the vector fields* $\mathrm{ad}_{-f}^k g^i, 0 \leqslant k \leqslant \lambda_i - 1, 1 \leqslant i \leqslant p$ *are complete. The state transformation* $\Phi(\zeta)$ *is obtained by solving PDEs*

$$\frac{\partial \Phi(\zeta)}{\partial \zeta} \left(\mathrm{ad}_{-f}^{\lambda_1 - 1} g^1, \cdots, g^1, \cdots, \mathrm{ad}_{-f}^{\lambda_p - 1} g^p, \ldots, g^p \right) = I_n. \tag{A.3}$$

Appendix B

Derivation of (4.15)

Since $f(\zeta) = W_* f(z)$, taking its partial derivative w.r.t. z_j^i gives

$$\frac{\partial f(\zeta)}{\partial z_j^i} = \frac{\partial W_*}{\partial z_j^i} f(z) + W_* \frac{\partial f(z)}{\partial z_j^i},$$

where $W_* = \partial W / \partial z$. Given $f(z) = S(Az + \gamma)$, we have

$$W_* \frac{\partial f(z)}{\partial z_j^i} = \begin{cases} s_i \frac{\partial W}{\partial z_{j-1}^i}, & 1 \leqslant i \leqslant p; 2 \leqslant j \leqslant \lambda_i, \\ W_* \frac{\partial}{\partial z_1^i} \left(S(Az + \gamma) \right), & 1 \leqslant i \leqslant p; j = 1. \end{cases} \tag{B.1}$$

On the other hand, we know

$$\begin{aligned} \frac{\partial f(\zeta)}{\partial z_j^i} - \frac{\partial W_*}{\partial z_j^i} f(z) &= \frac{\partial f(\zeta)}{\partial \zeta} \frac{\partial W}{\partial z_j^i} - \frac{\partial}{\partial \zeta} \left(\frac{\partial W}{\partial z_j^i} \right) f(\zeta) \\ &= \mathrm{ad}_{-f} \frac{\partial W}{\partial z_j^i}. \end{aligned} \tag{B.2}$$

Combining (B.1) and (B.2), we finally derive (4.15a) and (4.15b). To show (4.15c), we follow the expression of $h(z)$ and have $\partial h(z) / \partial z = C$, i.e.,

$$\frac{\partial h_r(z)}{\partial z_j^i} = \frac{\partial h_r(\zeta)}{\partial \zeta} \frac{\partial W}{\partial z_j^i} = \delta_{j,1} \delta_{r,i}, \quad 1 \leqslant j \leqslant \lambda_i; 1 \leqslant r, i \leqslant p.$$

Bibliography

[1] B. D. Anderson, R. R. Bitmead, C. R. J. Johnson, P. V. Kokotovic, R. L. Kosut, I. M. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis.* Cambridge, MA: MIT Press, 1986.

[2] V. Andrieu and L. Praly, "On the existence of a Kazantzis-Kravaris/Luenberger observer," *SIAM Journal on Control and Optimization*, vol. 45, no. 2, pp. 432–456, 2006.

[3] A. M. Annaswamy, F. P. Skantze, and A.-P. Loh, "Adaptive control of continuous time systems with convex/concave parameterization," *Automatica*, vol. 34, no. 1, pp. 33–49, Jan. 1998.

[4] B. Armstrong-Helouvry, P. Dupont, and C. C. D. Wit, "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083–1138, Jul. 1994.

[5] V. I. Arnol'd, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd ed. New York, NY: Springer-Verlag, 1983.

[6] J. Back and J. H. Seo, "Immersion of nonlinear systems into linear systems up to output injection: characteristic equation approach," *International Journal of Control*, vol. 77, no. 8, pp. 723–734, Aug. 2004.

[7] ——, "An algorithm for system immersion into nonlinear observer form: SISO case," *Automatica*, vol. 42, no. 2, pp. 321–328, Feb. 2006.

[8] J. Back, K. T. Yu, and J. H. Seo, "Dynamic observer error linearization," *Automatica*, vol. 42, no. 12, pp. 2195–2200, Dec. 2006.

[9] R. Bakker and A. M. Annaswamy, "Stability and robustness properties of a simple adaptive controller," *IEEE Transactions on Automatic Control*, vol. AC-41, no. 9, pp. 1352–1358, Sep. 1996.

[10] J. S. Baras, A. Bensoussan, and M. R. James, "Dynamic observers as asymptotic limits of recursive filters: Special cases," *SIAM Journal of Applied Mathematics*, vol. 48, no. 5, pp. 1147–1158, Oct. 1988.

[11] G. Bastin and M. R. Gevers, "Stable adaptive observers for nonlinear time-varying systems," *IEEE Transactions on Automatic Control*, vol. AC-33, no. 7, pp. 650–658, Jul. 1988.

[12] O. Bernard, G. Sallet, and A. Sciandra, "Nonlinear observers for a class of biological systems: Application to validation of a phytoplanktonic growth model," *IEEE Transactions on Automatic Control*, vol. AC-43, no. 8, pp. 1056–1065, Aug. 1998.

[13] G. Besançon, H. Hammouri, and S. Benamor, "State equivalence of discrete-time nonlinear control systems to state affine form up to input/output injection," *Systems and Control Letters*, vol. 33, no. 1, pp. 1–10, Jan. 1998.

[14] G. Besançon, J. D. Leon-Morales, and O. Huerta-Guevara, "On adaptive observers for state affine systems," *International Journal of Control*, vol. 79, no. 6, pp. 581–591, Jun. 2006.

[15] G. Besançon, "On output transformations for state linearization up to output injection," *IEEE Transactions on Automatic Control*, vol. AC-44, no. 10, pp. 1975–1981, Oct. 1999.

[16] ——, "Remarks on nonlinear adaptive observer design," *Systems and Control Letters*, vol. 41, no. 4, pp. 271–280, Nov. 2000.

[17] G. Besançon, G. Bornard, and H. Hammouri, "Observer synthesis for a class of nonlinear control systems," *European Journal of Control*, vol. 2, no. 3, pp. 176–192, 1996.

[18] G. Besançon and A. Ticlea, "An immersion-based observer design for rank-observable nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-52, no. 1, pp. 83–88, Jan. 2007.

[19] D. Bestle and M. Zeitz, "Canonical form observer design for non-linear time-variable systems," *International Journal of Control*, vol. 38, no. 2, pp. 419–431, Aug. 1983.

[20] J. Birk and M. Zeitz, "Extended Luenberger observer for nonlinear multivariable systems," *International Journal of Control*, vol. 47, no. 6, pp. 1823–1836, Jun. 1988.

[21] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemaniann Geometry.* New York, NY: Academic Press, 1975.

[22] G. Bornard and H. Hammouri, "A high gain observer for a class of uniformly observable systems," in *Proceedings of the 30th Conference on Decision and Control*, Brighton, England, 1991, pp. 1494–1496.

[23] ——, "A graph approach to uniform observability of nonlinear multi-output systems," in *Proceedings of the 41th Conference on Decision and Control*, Las Vegas, NV, 2002, pp. 701–706.

[24] S. A. Bortoff and A. F. Lynch, "Synthesis of optimal nonlinear observers," in *Proceedings of the 34th Conference on Decision and Control*, New Orleans, LA, 1995, pp. 95–100.

[25] J. D. Boskovic, "Adaptive control of a class of nonlinearly parameterized plants," *IEEE Transactions on Automatic Control*, vol. AC-43, no. 7, pp. 930–934, Jul. 1998.

[26] M. Boutayeb, H. Rafaralahy, and M. Darouach, "Convergence analysis of the Extended Kalman Filter used as an observer for nonlinear deterministic discrete-time systems," *IEEE Transactions on Automatic Control*, vol. AC-42, no. 4, pp. 581–586, Apr. 1997.

[27] C. I. Byres and A. Isidori, "Limit sets, zero dynamics, and internal models in the problem of nonlinear output regulation," *IEEE Transactions on Automatic Control*, vol. AC-48, no. 10, pp. 1712–1723, Oct. 2003.

[28] C. I. Byrnes and A. Isidori, "Local stabilization of minimum phase nonlinear systems," *Systems and Control Letters*, vol. 11, no. 1, pp. 9–17, Jul. 1988.

[29] ——, "Asymptotic stabilization of minimum phase nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-36, no. 10, pp. 1122–1137, Oct. 1991.

[30] C. Califano, S. Monaco, and D. Normand-Cyrot, "On observer design in discrete-time," *Systems and Control Letters*, vol. 49, pp. 255–265, 2003.

[31] R. L. Carroll and D. P. Lindorff, "An adaptive observer for single-input single-output linear systems," *IEEE Transactions on Automatic Control*, vol. AC-18, no. 5, pp. 428–435, Oct. 1973.

[32] X. Chen and H. Kano, "State observer for a class of nonlinear systems and its application to machine vision," *IEEE Transactions on Automatic Control*, vol. AC-49, no. 11, pp. 2085–2091, Nov. 2004.

[33] ——, "A new state observer for perspective systems," *IEEE Transactions on Automatic Control*, vol. AC-47, no. 4, pp. 658–663, Apr. 2002.

[34] D. Cheng and Y. Guo, "Stabilization of nonlinear systems via the center manifold approach," *Systems and Control Letters*, vol. 57, no. 6, pp. 511–518, Jun. 2008.

[35] D. Cheng and C. Martin, "Stabilization of nonlinear systems via designed center manifold," *IEEE Transactions on Automatic Control*, vol. AC-46, no. 9, pp. 1372–1383, Sep. 2001.

[36] D. Cheng and L. Zhang, "Generalized normal form and stabilization of nonlinear systems," *International Journal of Control*, vol. 76, no. 2, pp. 116–128, Feb. 2003.

[37] Y. M. Cho and R. Rajamani, "A systematic approach to adaptive observer synthesis for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-42, no. 4, pp. 534–537, Apr. 1997.

[38] G. Ciccarella, M. D. Mora, and A. Germani, "A Luenberger-like observer for nonlinear systems," *International Journal of Control*, vol. 57, no. 3, pp. 537–556, Mar. 1993.

[39] O. Dahl, Y. Wang, A. F. Lynch, and A. Heyden, "Observer forms for perspective systems," in *Proceedings of the IFAC 17th World Congress*, Seoul, Korea, Jul. 2008.

[40] D. M. Dawson, J. Hu, and T. C. Burg, *Nonlinear Control of Electric Machinery.* New York, NY: Marcel Dekker, 1998.

[41] E. Delaleau and W. Respondek, "Lowering the orders of derivatives of controls in generalized state space systems," *Journal of Mathematical Systems, Estimation, and Control*, vol. 5, no. 3, pp. 1–27, 1995, (Summary: 375–378).

[42] F. Deza, E. Busvelle, and J. P. Gauthier, "Exponentially converging observers for distillation columns and internal stability of the dynamic output feedback," *Chemical Engineering Science*, vol. 47, no. 15-16, pp. 3935–3941, Oct.-Nov. 1992.

[43] D. Dochain and G. Nastin, "Adaptive identification and control algorithms for nonlinear bacterial growth systems," *Automatica*, vol. 20, no. 5, pp. 621–634, May 1984.

[44] B. Friedland, "A nonlinear observer for estimating parameters in dynamic systems," *Automatica*, vol. 33, no. 8, pp. 1525–1530, Aug. 1997.

[45] J. P. Gauthier, H. Hammouri, and I. Kupka, "Observers for nonlinear systems," in *Proceedings of the 30th Conference on Decision and Control*, Brighton, England, 1991, pp. 1483–1489.

[46] J. P. Gauthier, H. Hammouri, and S. Othman, "A simple observer for nonlinear systems – applications to bioreactors," *IEEE Transactions on Automatic Control*, vol. AC-37, no. 6, pp. 875–880, Jun. 1992.

[47] J. P. Gauthier and I. A. K. Kupka, "Observability and observers for nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 32, no. 4, pp. 975–994, Jul. 1994.

[48] J. P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications.* UK: Cambridge University Press, 2001.

[49] S. S. Ge, C. C. Hang, and T. Zhang, "A direct adaptive controller for dynamic systems with a class of nonlinear parameterization," *Automatica*, vol. 35, no. 4, pp. 741–747, Apr. 1999.

[50] A. Gelb, *Applied Optimal Estimation.* Cambridge, MA: MIT Press, 1984.

[51] M. Guay, "An algorithm for orbital feedback linearization of single-input control affine systems," *Systems and Control Letters*, vol. 38, no. 4-5, pp. 271–281, Dec. 1999.

[52] ——, "Observer linearization by output-dependent time-scaling transformation," *IEEE Transactions on Automatic Control*, vol. AC-47, no. 10, pp. 1730–1735, Oct. 2002.

[53] H. Hammouri and J. de Leon Morales, "The time varying linearization up to output injection," in *Proceedings of the 28th Conference on Decision and Control*, Tampa, FL, 1989, pp. 1038–1039.

[54] ——, "Observer synthesis for state-affine systems," in *Proceedings of the 29th Conference on Decision and Control*, Honolulu, HI, 1990, pp. 784–785.

[55] H. Hammouri and J. P. Gauthier, "Bilinearization up to output injection," *Systems and Control Letters*, vol. 11, no. 2, pp. 139–149, Aug. 1988.

[56] ——, "Global time-varying linearization up to output injection," *SIAM Journal on Control and Optimization*, vol. 30, no. 6, pp. 1295–1310, Nov. 1992.

[57] H. Hammouri and M. Kinnaert, "A new procedure for time-varying linearization," *Systems and Control Letters*, vol. 28, no. 3, pp. 151–157, Jul. 1996.

[58] C.-C. Hang, "A new form of stable adaptive observer," *IEEE Transactions on Automatic Control*, vol. AC-21, no. 8, pp. 474–484, Aug. 1976.

[59] J. Hauser, S. Sastry, and P. Kokotović, "Nonlinear control via approximate input-output linearization: The ball and beam example," *IEEE Transactions on Automatic Control*, vol. AC-37, no. 3, pp. 392–398, Mar. 1992.

[60] R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. AC-22, no. 5, pp. 728–740, Oct. 1977.

[61] G. Hostetter and J. S. Meditch, "Observing systems with unmeasurable inputs," *IEEE Transactions on Automatic Control*, vol. AC-18, no. 3, pp. 307–308, Jun. 1973.

[62] L. R. Hunt and R. Su, "Linear equivalents of nonlinear time-varying systems," in *International Symposium on the Mathematical Theory of Networks and Systems*, Santa Monica, CA, 1981, pp. 119–123.

[63] P. Ioannou and G. Tao, "Frequency domain conditions for strictly positive real functions," *IEEE Transactions on Automatic Control*, vol. AC-32, no. 1, pp. 53–54, Jan. 1987.

[64] A. Isidori, *Nonlinear Control Systems*, 3rd ed. Berlin: Springer-Verlag, 1995.

[65] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-35, no. 2, pp. 131–140, Feb. 1990.

[66] B. Jakubczyk and W. Respondek, "On linearization of control systems," *Bulletin L'Académie Polonaise des Science, Série des Sciences Mathématiques*, vol. 28, no. 1, pp. 517–522, 1980.

[67] P. Jouan, "Immersion of nonlinear systems into linear systems modulo output injection," *SIAM Journal on Control and Optimization*, vol. 41, no. 6, pp. 1756–1778, 2003.

[68] R. E. Kalman, "Contributions to the theory of optimal control," *Boletin de la Sociedad Mathematica Mexicana*, vol. 5, no. 1, pp. 102–119, 1960.

[69] ——, "A new approach to linear filtering and prediction problems," *Transactions of the ASME, series D, Journal of Basic Engineering*, vol. 82, pp. 33–45, 1960.

[70] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Transactions of the ASME, series D, Journal of Basic Engineering*, vol. 83, pp. 95–108, 1961.

[71] L. Karsenti, F. Lamnbhi-Lagarrigue, and G. Bastin, "Adaptive control of nonlinear systems with nonlinear parameterization," *Systems and Control Letters*, vol. 27, no. 2, pp. 87–97, Feb. 1996.

[72] N. Kazantzis and C. Kravaris, "A nonlinear Luenberger-type observer with application to catalyst activity estimation," in *Proceedings of the 1995 American Control Conference*, Seattle, WA, 1995, pp. 1756–1761.

[73] ——, "Nonlinear observer design using Lyapunov's Auxiliary Theorem," *Systems and Control Letters*, vol. 34, no. 5, pp. 241–247, Jul. 1998.

[74] H. Keller, "Nonlinear observer design by transformation into a generalized observer canonical form," *International Journal of Control*, vol. 46, no. 6, pp. 1915–1930, Dec. 1987.

[75] H. Keller and H. Fritz, "Design of nonlinear observers by a two-step-transformation," in *Algebraic and Geometric Methods in Nonlinear Control Theory*, M. Fliess and M. Hazewinkel, Eds. D. Reidel, 1986, pp. 89–99.

[76] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[77] H. K. Khalil and F. Esfandiari, "Semiglobal stabilization of a class of nonlinear systems using output feedback," *IEEE Transactions on Automatic Control*, vol. AC-38, no. 9, pp. 1412–1415, Sep. 1993.

[78] A. Kojić and A. M. Annaswamy, "Adaptive control of nonlinearly parameterized systems with a triangular structure," *Automatica*, vol. 38, no. 1, pp. 115–123, Jan. 2002.

[79] A. Kojić, A. M. Annaswamy, A.-P. Loh, and R. Lozano, "Adaptive control of a class of nonlinear systems with convex/concave parameterization," *Systems and Control Letters*, vol. 37, no. 5, pp. 267–274, Aug. 1999.

[80] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Transactions on Automatic Control*, vol. AC-22, no. 1, pp. 2–8, Feb. 1977.

[81] ——, "The generation of adaptive law structures for globally convergent adaptive observers," *IEEE Transactions on Automatic Control*, vol. AC-24, no. 3, pp. 510–512, Jun. 1979.

[82] A. Krener and M. Xiao, "Nonlinear observer design in the Siegel domain," *SIAM Journal on Control and Optimization*, vol. 41, no. 3, pp. 932–953, 2002.

[83] ——, "Erratum: Nonlinear observer design in the Siegel domain," *SIAM Journal on Control and Optimization*, vol. 43, no. 1, pp. 377–378, 2004.

[84] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Systems and Control Letters*, vol. 3, no. 1, pp. 47–52, Jun. 1983.

[85] A. J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *SIAM Journal on Control and Optimization*, vol. 23, no. 2, pp. 197–216, 1985.

[86] A. J. Krener, "Conditioned invariant and locally conditioned invariant distributions," *Systems and Control Letters*, vol. 8, no. 1, pp. 69–74, Oct. 1986.

[87] M. Krstić, I. Kanellakopoulos, and P. Kokotović, *Nonlinear and Adaptive Control Design.* New York, NY: John Wiley and Sons, 1995.

[88] P. Kudva and K. S. Narendra, "Synthesis of an adaptive observer using Lyapunov's direct method," *International Journal of Control*, vol. 18, no. 6, pp. 1201–1210, Dec. 1973.

[89] W. Lee and K. Nam, "Observer design for autonomous discrete-time nonlinear systems," *Systems and Control Letters*, vol. 17, no. 1, pp. 49–58, Jul. 1991.

[90] J. Levine and R. Marino, "Nonlinear system immersion. Observers and finite-dimensional filters," *Systems and Control Letters*, vol. 7, no. 2, pp. 133–142, Apr. 1986.

[91] F. L. Lewis, *Optimal Estimation.* New York, NY: John Wiley and Sons, 1986.

[92] C. W. Li and L. W. Tao, "Observing non-linear time-variable systems through a canonical form observer," *International Journal of Control*, vol. 44, no. 6, pp. 1703–1713, Dec. 1986.

[93] Y. Liang and H. J. Marquez, "Gain scheduling synchronization method for quadratic chaotic systems," *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, vol. 55, no. 4, pp. 1097–1107, May 2008.

[94] W. Lin and C. Qian, "Adaptive regulation of cascade systems with nonlinear parameterization," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 12, pp. 1093–1108, Oct. 2002.

[95] A.-P. Loh, A. M. Annaswamy, and F. P. Skantze, "Adaptation in the presence of a general nonlinear parameterization: An error model approach," *IEEE Transactions on Automatic Control*, vol. AC-44, no. 9, pp. 1634–1652, Sep. 1999.

[96] G. Lüders and K. S. Narendra, "A new canonical form for an adaptive observer," *IEEE Transactions on Automatic Control*, vol. AC-19, pp. 117–119, Apr. 1974.

[97] G. Lüders and K. S. Narendra, "An adaptive observer and identifier for a linear system," *IEEE Transactions on Automatic Control*, vol. AC-18, no. 5, pp. 496–499, Oct. 1973.

[98] A. F. Lynch and S. A. Bortoff, "Least squares approximate error linearization," in *Proceedings of the 1997 American Control Conference*, Albuquerque, NM, 1997, pp. 176–181.

[99] ——, "Nonlinear observer design by approximate error linearization," *Systems and Control Letters*, vol. 32, no. 3, pp. 161–172, Nov. 1997.

[100] A. F. Lynch, S. A. Bortoff, and K. Röbenack, "Nonlinear tension observers for web machines," *Automatica*, vol. 40, no. 9, pp. 1517–1524, Sep. 2004.

[101] L. Marconi, L. Praly, and A. Isidori, "Output stabilization via nonlinear Luenberger observers," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2277–2298, 2006.

[102] R. Marino and P. Tomei, "Adaptive observers for a class of multi-output nonlinear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 6, no. 4, pp. 353–365, Jul. 1992.

[103] ——, "Global adaptive observers for nonlinear systems via filtered transformations," *IEEE Transactions on Automatic Control*, vol. AC-37, no. 8, pp. 1239–1245, Aug. 1992.

[104] ——, "Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-40, no. 7, pp. 1300–1304, Jul. 1995.

[105] R. Marino, "Adaptive observers for single output nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-35, no. 9, pp. 1054–1058, Sep. 1990.

[106] R. Marino and P. Tomei, "Dynamic output feedback linearization and global stabilization," *Systems and Control Letters*, vol. 17, no. 2, pp. 115–121, Aug. 1991.

[107] ——, "Global adaptive output-feedback control of nonlinear systems, part I: Linear parameterization," *IEEE Transactions on Automatic Control*, vol. AC-38, no. 1, pp. 17–32, Jan. 1993.

[108] ——, "Global adaptive output-feedback control of nonlinear systems, part II: Nonlinear parameterization," *IEEE Transactions on Automatic Control*, vol. AC-38, no. 1, pp. 33–48, Jan. 1993.

[109] ——, *Nonlinear Control Design: Geometric, Adaptive, and Robust.* Hertfordshire, UK: Prentice-Hall, 1995.

[110] F. Matsumura and T. Yoshimoto, "System modeling and control design of a horizontal-shaft magnetic-bearing system," *IEEE Transactions on Magnetics*, vol. 22, no. 3, pp. 196–203, May 1986.

[111] K. Meyer, "On the existence of Lyapunov functions for the problem on Lur'e," *SIAM Journal on Control and Optimization*, vol. 3, pp. 373–383, 1965.

[112] P. E. Moraal and J. W. Grizzle, "Observer design for nonlinear systems with discrete-time measurements," *IEEE Transactions on Automatic Control*, vol. AC-40, no. 3, pp. 395–404, Mar. 1995.

[113] K. Nam, "An approximate nonlinear observer with polynomial coordinate transformation maps," *IEEE Transactions on Automatic Control*, vol. AC-42, no. 4, pp. 522–527, Apr. 1997.

[114] K. Narendra and P. Kudva, "Stable adaptive schemes for system identification and control–parts I & II," *IEEE Transactions on Systems, Man and Cybernetcis*, vol. 4, no. 6, pp. 542–560, Nov. 1974.

[115] K. Narendra and L. Valavani, "Stable adaptive observers and controllers," in *Proceedings of the IEEE*, vol. 64, Aug. 1976, pp. 1198–1208.

[116] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1989.

[117] S. Nicosia, A. Tornambè, and P. Valigi, "Use of observers for the inversion of nonlinear maps," *Systems and Control Letters*, vol. 16, no. 6, pp. 447–455, Jun. 1991.

[118] H. Nijmeijer, "Observability of a class of nonlinear systems: A geometric approach," *Ricerche Di Automatica*, vol. 12, no. 1, pp. 1107–1130, 1981.

[119] ——, "An observer looks at synchronization," *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, vol. 44, no. 10, pp. 882–890, Oct. 1997, special Issue on Chaos Synchronization and Control: Theory and Applications.

[120] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems.* New York, NY: Springer-Verlag, 1990.

[121] D. Noh, N. Jo, and J. Seo, "Nonlinear observer design by dynamic observer error linearization," *IEEE Transactions on Automatic Control*, vol. AC-49, no. 10, pp. 1746–1753, Oct. 2004.

[122] J. O'Reilly, *Observers for Linear Systems.* New York, NY: Academic Press, 1983.

[123] W. Perruquetti, T. Floquet, and E. Moulav, "Finite-time observers: application to secure communication," *IEEE Transactions on Automatic Control*, vol. AC-53, no. 2, pp. 356–360, Feb. 2008.

[124] A. R. Phelps, "A simplification of nonlinear observer theory," Ph.D. dissertation, University of California, Berkeley, CA, 1987.

[125] ——, "On constructing nonlinear observers," *SIAM Journal on Control and Optimization*, vol. 29, no. 3, pp. 516–534, 1991.

[126] M. Poincarré, "Sur les Propriétés des fonctions Définies par les équations aux différences partielles," Ph.D. dissertation, Faculté des Sciences de Paris, Paris, France, 1879.

[127] F. D. Priscoli, L. Marconi, and A. Isidori, "Adaptive observers as nonlinear internal models," *Systems and Control Letters*, vol. 55, no. 8, pp. 640–649, Aug. 2006.

[128] C. Qian and W. Lin, "Recursive observer design, homogeneous approximation, and nonsmooth output feedback stabilization of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. AC-51, no. 9, pp. 1457–1471, Sep. 2006.

[129] R. Rajamani, "Observers for nonlinear systems, with application to active automotive suspensions," Ph.D. dissertation, University of California, Berkeley, CA, 1993.

[130] K. Reif and R. Unbehauen, "Linearization along trajectories and the Extended Kalman Filter," in *Proceedings of the IFAC 13th World Congress*, vol. H, 1996, pp. 509–514.

[131] W. Respondek, "Orbital feedback linearization of single-input nonlinear control systems," *Proceedings of IFAC NOLCOS'98*, pp. 499–504, 1998.

[132] W. Respondek, A. Pogromsky, and H. Nijmeijer, "Time scaling for observer design with linearizable error dynamics," *Automatica*, vol. 40, no. 2, pp. 277–285, Feb. 2004.

[133] E. Rimon and K. S. Narendra, "A new adaptive estimator for linear systems," *IEEE Transactions on Automatic Control*, vol. AC-37, no. 3, pp. 410–412, Mar. 1992.

[134] J. Rudolph and M. Zeitz, "Block triangular nonlinear observer normal form," *Systems and Control Letters*, vol. 23, no. 1, pp. 1–8, Jul. 1994.

[135] J. Rudolph, "Kanonische formen nichtlinearer systeme," Inst. f. Systemdynamik u. Regelungstechnik, Univ. Stuttgart, Stuttgart, Germany, Diplomarbeit, 1989.

[136] W. Rugh, *Linear System Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[137] M. Sampei and K. Furuta, "On time scaling for nonlinear systems: Applications to linearization," *IEEE Transactions on Automatic Control*, vol. AC-31, no. 5, pp. 459–462, May. 1986.

[138] S. S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness.* Englewood Cliffs, NJ: Prentice-Hall, 1989.

[139] J. Schaffner and M. Zeitz, "Decentral nonlinear observer design using a block-triangular form," *International Journal of System Science*, vol. 30, no. 10, pp. 1131–1142, Oct. 1999.

[140] ——, "Variants of nonlinear normal form observer design," in *New Direction in Nonlinear Observer Design*, ser. Lecture Notes in Control and Inform. Sci., H. Hijmeijer and T. I. Fossen, Eds. UK: Springer-Verlag, 1999, vol. 244, pp. 161–180.

[141] H. Shim, Y. I. Son, and J. H. Seo, "Semi-global observer for multi-output nonlinear systems," *Systems and Control Letters*, vol. 42, no. 3, pp. 233–244, Mar. 2001.

[142] F. P. Skantze, A. Kojić, A.-P. Loh, and A. M. Annaswamy, "Adaptive estimation of discrete-time systems with nonlinear parameterization," *Automatica*, vol. 36, no. 12, pp. 1879–1887, Dec. 2000.

[143] G. L. Smith, S. Schmidt, and L. A. McGee, "Application of statistical filtering to the optimal estimation of position and velocity on board a circumlunar vehicle," NASA Ames Research Center Report, Tech. Rep. TND-1205, 1963.

[144] Y. Song and J. W. Grizzle, "The Extended Kalman Filter as a local asymptotic observer for nonlinear discrete-time systems," in *Proceedings of the 1992 American Control Conference*, Chicago, IL, 1992, pp. 3365–3369.

[145] Z. Soons, J. Shi, J. Stigter, L. van der Pol, G. van Straten, and A. van Boxtel, "Observer design and tuning for biomass growth and K_La using online and offline measurements," *Journal of Process Control*, vol. 18, no. 7-8, pp. 621–631, Aug.-Sep. 2008.

[146] H. J. Sussmann, "Single-input observability of continuous-time systems," *Mathematical Systems Theory*, vol. 12, pp. 371–393, 1979.

[147] J. H. Taylor, "Strictly positive-real functions and the Lefschetz-Kalman-Yakubovich (LKY) lemma," *IEEE Transactions on Circuits and Systems*, vol. 21, no. 2, pp. 310–311, Feb. 1974.

[148] A. Tornambè, "Use of asymptotic observers having high-gains in the state and parameter estimation," in *Proceedings of the 28th Conference on Decision and Control*, Tampa, FL, 1989, pp. 1774–1779.

[149] J. Tsinias, "Global stabilization by output feeback for triangular systems," *Automatica*, vol. 35, no. 1, pp. 183–186, Jan. 1999.

[150] ——, "Backstepping design for time-varying nonlinear systems with unknown parameters," *Systems and Control Letters*, vol. 39, no. 4, pp. 219–227, Apr. 2000.

[151] ——, "Further results on the observer design problem," *Systems and Control Letters*, vol. 14, no. 5, pp. 411–449, Jun. 1990.

[152] A. J. van der Schaft, "On nonlinear observers," *IEEE Transactions on Automatic Control*, vol. AC-30, no. 12, pp. 1254–1256, Dec. 1985.

[153] S. H. Wang, E. J. Davison, and P. Dorato, "Observing the states of systems with unmeasurable disturbances," *IEEE Transactions on Automatic Control*, vol. AC-20, no. 5, pp. 716–717, Oct. 1975.

[154] Y. Wang and A. F. Lynch, "A block triangular form for nonlinear observer design," *IEEE Transactions on Automatic Control*, vol. AC-51, no. 11, pp. 1803–1808, Nov. 2006.

[155] ——, "A block triangular observer forms for nonlinear observer design," *International Journal of Control*, vol. 81, no. 2, pp. 177–188, Feb. 2008.

[156] ——, "Observer design using a generalized time scaled block triangular observer form," *Systems and Control Letters*, 2008, submitted.

[157] ——, "Observer design using a time scaled block triangular observer form," in *Proceedings of the 2006 American Control Conference*, Minneapolis, MN, Jun. 2006, pp. 799–804.

[158] ——, "On the existence of a block triangular form," in *Proceedings of the 2007 American Control Conference*, New York, NY, Jul. 2007, pp. 2346–2351.

[159] ——, "Further discussion on a block triangular observer form," in *Proceedings of the 2008 Chinese Control Conference*, Kunming, China, Jul. 2008, pp. 332–336.

[160] ——, "Time scaling of a multi-output observer form," in *Proceedings of the 2008 American Control Conference*, Seattle, WA, Jun. 2008, pp. 5242–5247.

[161] ——, "Time scaling of a multi-output observer form," *IEEE Transactions on Automatic Control*, 2008, submitted.

[162] R. Wasserman, *Tensors and Manifold: With Application to Applications*, 2nd ed. Oxford, England: Oxfold University Press, 2004.

[163] W. A. Wolovich, "On state estimation of observable systems," in *Proceedings of the 9th Joint Automatic Control Conference*, Jun. 1968, pp. 210–220.

[164] F. F. Wu and C. A. Desoer, "Global inverse function theorem," *IEEE Transactions on Circuits Theory*, vol. 19, no. 2, pp. 199–201, Mar. 1972.

[165] X.-H. Xia and W.-B. Gao, "Non-linear observer design by observer canonical form," *International Journal of Control*, vol. 47, no. 4, pp. 1081–1100, Apr. 1988.

[166] X. Xia and W. Gao, "On exponential observers for nonlinear systems," *Systems and Control Letters*, vol. 11, pp. 319–325, Oct. 1988.

[167] ——, "Nonlinear observer design by observer error linearization," *SIAM Journal on Control and Optimization*, vol. 27, no. 1, pp. 199–216, 1989.

[168] X. Xia, "Estimation of HIV/AIDS parameters," *Automatica*, vol. 39, no. 11, pp. 1983–1988, Nov. 2003.

[169] A. Xu and Q. Zhang, "Nonlinear system fault diagnosis based on adaptive estimation," *Automatica*, vol. 40, no. 6, pp. 1181–1193, Jun. 2004.

[170] M. Zeitz, "Comments on "Comparative Study of Non-Linear State-Observation Techniques"," *International Journal of Control*, vol. 46, no. 5, pp. 1861–1863, Nov. 1987.

[171] Q. Zhang, "Adaptive observer for Multiple-Input-Multiple-Output (MIMO) linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. AC-47, no. 3, pp. 525–529, Mar. 2002.

[172] T. Zhang and M. Guay, "Adaptive nonliner observers of microbial growth processes," *Journal of Process Control*, vol. 12, no. 5, pp. 633–643, Aug. 2002.

[173] G. Zheng, D. Boutat, and J. Barbot, "Multi-output dependent observability normal form," *Journal of Nonlinear Analysis A: Theorey, Methods & Application*, to appear.

[174] ——, "Single output dependent observability normal form," *SIAM Journal on Control and Optimization*, vol. 46, no. 6, pp. 2242–2255, 2007.