Development of Sliding Mode Observer Schemes for Infinitely Unobservable Descriptor Systems

Joseph Chan Chang Lun

B. Eng. (Hons) in Mechatronics Engineering

Monash University

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2018

School of Engineering

Contents

COPYRIGHT NOTICE

ABSTRACT

This thesis describes the utilisation of sliding mode observers (SMOs) for the estimation of states and unknown inputs for a class of descriptor systems. In existing SMOs schemes, suitable processing of the equivalent output error injection term yields an accurate estimate of the unknown input. These SMO schemes however require the observed system to be infinitely observable which can limit the applicability of the scheme, or do not consider the effects of external disturbances on the estimate of the unknown input. This thesis therefore aims to further develop SMO theory for *non*-infinitely observable descriptor systems (NIODS), and consider the effects of disturbances affecting the system. The approaches presented first reformulate the system into a form facilitating further analysis, capitalising on the ability of the SMO to estimate unknown inputs. Central to this reformulation is the treatment of certain states as unknown inputs, which results in the NIODS being re-expressed as an infinitely observable reduced-order system. Existing techniques which are applicable to infinitely observable systems are designed based on this reduced-order system to estimate its states and unknown inputs. Linear matrix inequality (LMI) techniques are used to design the observer gains such that the effects of disturbances on the estimates of the unknown input are bounded, thus achieving *robust reconstruction of the unknown input*. The use of cascaded SMOs for estimating states and unknown inputs is explored and compared to existing cascaded observer schemes in the literature. In the final chapter, a more complex reformulation technique (where certain states are re-expressed as a linear combination of other states) is presented, along with considerations for disturbances. The necessary and sufficient conditions for the feasibility of all schemes are shown in terms of the original system matrices, allowing designers to quickly verify if the schemes are applicable. Each scheme is also accompanied by a set of design procedures, and are verified using numerical examples.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed : [redacted]

Print Name : Joseph Chan Chang Lun

Date : 29 March 2019

PUBLICATIONS DURING ENROLMENT

Published Results

1. J. C. L. Chan, C. P. Tan, and H. Trinh, Robust fault reconstruction for a class of infinitely unobservable descriptor systems, *International Journal of Systems Science*, 48(8):1646–1655, 2017.
2. J. C. L. Chan, C. P. Tan, H. Trinh, and M. A. S. Kamal, State and fault estimation for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade, *Journal of the Franklin Institute*, 356(5):3010–3029, 2019.
3. J. C. L. Chan, C. P. Tan, H. Trinh, M. A. S. Kamal, and Y. S. Chiew, Robust fault reconstruction for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade,' *Applied Mathematics and Computation*, 350:78–92, 2019.

Accepted for Publication

1. J. C. L. Chan, C. P. Tan, J. H. T. Ooi, and H. Trinh, New results in robust fault reconstruction for a class of non-infinitely observable descriptor systems, to be presented at the *American Control Conference (ACC) 2019*.

ACKNOWLEDGEMENTS

My Ph.D. journey has been an arduous one, and completing it would have been nigh impossible if not for the following individuals:

My main supervisor, Associate Professor Edwin Tan Chee Pin, who guided me steadfastly throughout my candidature, patiently providing pointers and feedback without tire, and kept my eyes on the prize whenever I started veering off track.

My co-supervisor Associate Professor Hieu Trinh, who I could always rely on for helpful advice when disseminating results, and was supportive of my proposals for future collaboration.

My family - specifically my mother Siew Yee Ha and brother Joshua Chan Chang Zhi. You both made sure I could still function every day by providing for my every need at home. The house has always been a wonderful workplace thanks to your efforts.

The members of the Control Research Group, Dr Kee Chew Yee, Dr Jeremy Ooi Hor Teong, Dr Ng Jiunn Yea, Dr Chua Wen-Shyan, and Loo Junn Yong, who gave their best whenever I asked for their assistance. In particular, I thank Drs. Ooi and Chua for their extended help prior to and throughout my Ph.D. journey.

My fellow postgraduates, Dr Andreas Aditya Hermawan, Gajeshwaren V. Balakrishnan, and Chang Tak Kwin, who reminded me that I am not alone on this perilous journey, and offered alternative views on the problems we all face at one point or another.

My good friends Voon Seng Wai, Ng Jian Wai, and Jonathan Tan, who kept me sane when the going got tough. The fun we had together helped stave off burnout during the Ph.D. grind.

Last but not least, my dearest Dr Janet Leong Yip Cheng, who was always there with me through every step, braving stormy times and tight deadlines, to support and care for me, and provide guidance and assistance wherever she could.

Chapter 1

Motivation and thesis overview

1.1 Introduction

Modern engineering systems are increasingly designed to be operated autonomously. The larger amount of periphery required for the operation and performance of the systems however results in a greater likelihood of unpermitted deviations occurring in properties or parameters of the system from acceptable or standard conditions (hereafter referred to as *faults*) [22, 66]. These faults could lead to undesirable consequences, such as financial losses which are incurred due to lower quality output from the system, or when the system is taken offline for maintenance. Thus, a cost-effective form of preventive maintenance would be to detect and reconstruct these faults immediately when and as they occur, i.e. *fault detection and identification* (FDI). FDI schemes are vital for the operation of any autonomous system, since they provide information such as the shape and magnitude of faults, which are essential for remedial action [103, 148].

While most engineering systems can be represented using the state-space system representation, there exist systems which also contain algebraic equations, which are better represented using the *descriptor system representation*, which has the form

$$E\dot{x} = Ax + Bu + Mf, \tag{1.1}$$

$$y = Cx, \tag{1.2}$$

where $E,\ A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m},\ M \in \mathbb{R}^{n\times q},\ C \in \mathbb{R}^{p\times n}$, and E is generally rank-deficient, i.e. $\mathrm{rank}(E) = r < n$. The vectors $x,\ u,\ f$, and y represent the states, inputs, faults, and outputs, respectively. These algebraic equations in descriptor system represent interactions between subsystems evolving on different time scales [38]. Unlike regular state-space systems which only have one measure of observability, descriptor systems have multiple measures of observability [30]. Amongst these measures of observability is the concept of *infinite observability*, which implies

$$\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n. \tag{1.3}$$

Existing FDI schemes require (1.3) to hold for system (1.1)–(1.2) in order to reconstruct f (i.e. system (1.1)–(1.2) is infinitely observable). This requirement may be restrictive, and thus forms a current limitation on the applicability of FDI schemes in descriptor systems.

Additionally, most FDI schemes rely on a model of the system, which is a mathematical representation derived from physical operation or through simulation of the system under practical situations. These models however may oversimplify the actual system, or fail to take into account unknown signals acting on the system. Parasitic dynamics, uncertain parameters, or unknown external influences (all of which are hereafter referred to collectively as *disturbances*) may be present in the physical system. These disturbances may corrupt the signals reconstructed by the FDI scheme. These erroneous reconstructions may raise false alarms, or even worse, mask the effect of a fault.

Thus this thesis aims to present work on sliding mode observer (SMO) schemes for non-infinitely observable descriptor systems (i.e. systems where (1.3) is not satisfied), and to design SMO schemes such that the effects of disturbances on the fault reconstruction are bounded. The rest of this chapter presents an outline of the structure of the thesis, and the contribution of each chapter to this area of research.

1.2 Thesis structure

Chapter 2 presents a survey of the literature as well as prior work on topics that underlie the thesis. The concept of state and fault estimation using observers is first presented. The sliding mode observer (SMO) is then introduced, and the concepts of *robust fault reconstruction* and *observers in cascade* are explored. The descriptor system representation and observers for descriptor systems are then elaborated upon. Finally, observer schemes for non-infinitely observable descriptor systems (NIODS) are detailed.

Chapter 3 presents the SMO scheme for state and fault estimation in *infinite observable* [49, 160] descriptor systems by Yeu et al. [156, 157]. The error system for the SMO is analysed, and basic sliding mode concepts are reviewed. A simulation example is then used to illustrate the concepts presented. Next, two SMO schemes state and fault estimation in a class of NIODS by Ooi et al. [100, 101] are studied. The schemes re-express NIODS as infinitely observable reduced-order systems by treating certain states as unknown inputs, and in the second scheme, by re-expressing certain states in terms of other states as well. These re-expression techniques would provide the basis of the work presented in the rest of the thesis. A simulation example is used to demonstrate the second scheme by Ooi et al. [101].

Chapter 4 presents a robust fault reconstruction scheme for a class of NIODS which builds on the work in [100]. Firstly, the system is re-expressed by treating certain states as unknown inputs, resulting in an infinitely observable reduced-order system. The SMO by Yeu et al. [157] is then designed based on this reduced-order system. The necessary and sufficient conditions for the feasibility of the observer are investigated and presented in terms of the original system matrices. The observer gains are then designed using the Bounded Real Lemma [128] such that the $\mathcal{L}_2$ gain from the disturbances on the fault reconstruction is minimised. The work in this chapter has been published [17]; its details are as follows: J. C. L. Chan, C. P. Tan, and H. Trinh, Robust fault reconstruction for a class of infinitely unobservable descriptor systems, *International Journal of Systems Science*, 48(8):1646–1655, 2017.

Chapter 5 presents a state and fault estimation scheme for a class of NIODS using two SMOs in cascade. The system is first re-expressed into an infinitely observable reduced-order system by treating certain states as unknown inputs. The SMO by Yeu et al. [157] is designed based on this reduced-order system, and it is found that the switching term of the observer is the output of an analytical regular (non-descriptor) state-space system that treats the remaining fault components as unknown inputs. The Edwards-Spurgeon SMO [6] is then designed based on this analytical system to estimate the remaining faults. The proposed scheme is able to estimate all states and faults in finite time. The necessary and sufficient conditions for the existence of the scheme are studied and shown in terms of the original system matrices, and it is found that the proposed scheme has less stringent conditions compared to existing schemes utilising single SMOs, or observers in cascade. The work in this chapter has been published [18]; its details are as follows: J. C. L. Chan, C. P. Tan, H. Trinh, and M. A. S. Kamal, State and fault estimation for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade, *Journal of the Franklin Institute*, 356(5):3010–3029, 2019.

Chapter 6 extends the findings in Chapter 5 to perform robust fault reconstruction for a class of NIODS using two SMOs in cascade. The approach in Chapter 5 is used to reconstruct the faults. Linear matrix inequality (LMI) techniques and the Bounded Real Lemma [128] are used to design the observers such that the $\mathcal{L}_2$ gain from the disturbances onto the fault reconstruction is minimised. An interesting observation is that the gains of both cascaded observers can be designed using a single LMI pair, as opposed to designing each SMO separately (which would increase the conservativeness of the solution) as in previous works. The existence conditions are investigated and presented in terms of the original system matrices. The work in this chapter has been published [19]; its details are as follows: J. C. L. Chan, C. P. Tan, H. Trinh, M. A. S. Kamal, and Y. S. Chiew, Robust fault reconstruction for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade, *Applied Mathematics and Computation*, 350:78–92, 2019.

Chapter 7 presents a robust fault reconstruction scheme for a class of NIODS which builds on the scheme in [101]. The re-expression in this chapter differs from those of the previous chapter: some states are first re-expressed in terms of other signals, and then certain other states are treated as unknown inputs to form an infinitely observable reduced-order system. The observer by Yeu et al. [157] is then applied onto this reduced-order system. The observer gains are then designed using the Bounded Real Lemma [128] such that the $\mathcal{L}_2$ gain from the disturbances on the fault reconstruction is minimised. The necessary and sufficient conditions for the feasibility of the scheme are then presented in terms of the original system matrices. It is found that this method of re-expression increases the applicability of the scheme, as compared to those in the previous chapters. The work in this chapter has accepted for publication; its details are as follows: J. C. L. Chan, C. P. Tan, J. H. T. Ooi, and H. Trinh, New results in robust fault reconstruction for a class of non-infinitely observable descriptor systems, to be presented at the *American Control Conference (ACC) 2019*.

Chapter 8 then summarises the thesis by drawing conclusions from each chapter, and also suggesting potential future lines of inquiry related to the presented work.

Chapter 2

Literature review

2.1 Introduction

Engineering systems are increasingly designed to operate autonomously, which places greater importance on the controllers maintaining the required performance [84]. These controllers in turn rely on information regarding the current state of the system, which are in turn derived from sensor measurements. For example, sensors are utilised in the petroleum industry to monitor fluids being transported in pipelines and detect leaks [1]. Sensors and their periphery however incur steep costs. Additionally, some states may have no physical meaning or are impossible to measure. Hence states that cannot be measured have to be estimated using a *state observer*, which would calculate these unmeasured states based on information derived from measured inputs and outputs.

The chapter first discusses the concept and development of the state observer into more advanced methods to estimate information regarding the system in §2.2. The sliding mode observer (SMO) is a type of observer that has several advantages over other observers, resulting in the field receiving intense research interest and development, which will also be detailed in §2.3. More advanced techniques utilising the SMO, such as for robust fault reconstruction and SMOs in cascade will also be introduced. Finally, in §2.4, the descriptor system representation, the concept of *infinite observability* as well as observer schemes for non-infinitely observable descriptor systems (NIODS) will be discussed in detail as part of a comprehensive literature survey for the work underlying this thesis.

2.2 State estimation

A popular alternative to utilising hardware-based methods (i.e. sensors) for state measurement is the use of software-based methods. One of the most established software-based methods for state measurement is the use of an observer scheme. The state observer is essentially a simulation of a dynamical model representing the system, from which estimates of the states are derived. The observer was proposed by Luenberger [80], where the unmeasured states are estimated from the measured input and output signals. The system and observer are fed the same input signals, and their outputs are compared and the differences constitute the *output estimation error*, which is fed back into the observer (which is in turn designed such that the state estimation error goes to zero). To illustrate the concept of a Luenberger observer, consider the following linear-time invariant (LTI) system modelled using the state-space representation:

$$\dot{x} = Ax + Bu, \tag{2.1}$$

$$y = Cx, \tag{2.2}$$

where $A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$ are known and constant. The vectors $x,\ u$, and y represent the states, inputs, and outputs, respectively. Only u and y are measurable, and it is desired to estimate x. The Luenberger observer for system (2.1)–(2.2) has the form

$$\dot{\hat{x}} = A\hat{x} + Bu - Le_y, \tag{2.3}$$

where $\hat{x}$ is the measurable estimate of $x,\ L$ is a design matrix, and $e_y = C\hat{x} - y$. The state estimation error is defined to be $e = \hat{x} - x$. By combining (2.1)–(2.3), the following error system (which characterises the performance of the observer) is obtained:

$$\dot{e} = (A - LC)\, e. \tag{2.4}$$

Hence, by choosing L such that the eigenvalues of $A - LC$ (denoted as $\lambda\,(A - LC)$) have negative real parts, then e converges to zero, which implies $\hat{x}$ asymptotically estimates x (i.e. $\hat{x} \to x$ as $t \to \infty$). A schematic of the Luenberger observer (2.3) for system (2.1)–(2.2) is shown in Figure 2.1.

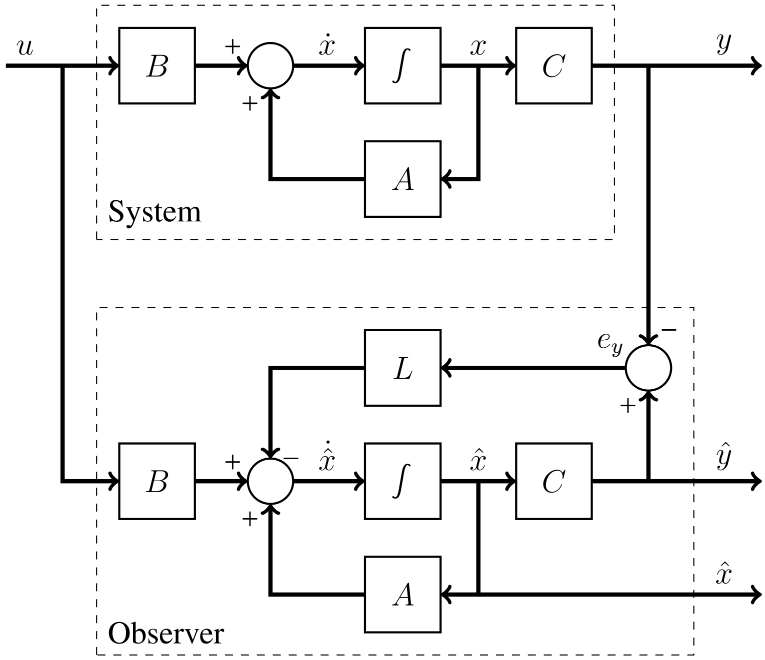

Figure 2.1: Schematic diagram of the Luenberger observer.

The states estimated by the observer can then be used in the controller to regulate the performance of the system. This observer-controller methodology has been widely studied: Muehlebach and Trimpe [88] designed a distributed control system for a multi-agent network, where each agent has its own observer-controller pair to monitor a part of the system and provide feedback to each part. Sadelli et al. [111] presented a 2D observer-based control scheme for a magnetic microbot navigating in a cylindrical blood vessel. Chen et al. [21] studied an observer-based control scheme for multi-input multi-output cascade nonlinear systems. Jafari et al. [68] investigated the use of a Lyapunov observer-controller in haptic interfaces and teleoperation systems. Di Giorgio et al. [37] designed a defence strategy against destabilising cyberphysical systems and power systems consisting of observers monitoring the system and reconstructing attack vectors, and then feeding that information into controllers that provide feedback that mitigates the effects of the attack. Chen et al. [25] proposed an observer-based control system for a class of discrete switched stochastic systems with mixed delays. Gholami and Binazadeh [56] presented an observer-based control scheme for nonlinear one-side Lipschitz systems affected by time delays. The observer-based controller methodology has also been applied onto many other practical systems, such as boost converters [27], multi-link flexible manipulators [146], and DC motors [74].

2.2.1 Robust state estimation

The representation in (2.1)–(2.2) is ideal; practical systems cannot be represented using that representation due to mismatches between the actual system and the mathematical model (upon which the observer is designed). These mismatches could arise due to parasitic dynamics, time-varying parameters, and external influence. A practical system would thus be better represented as:

$$\dot{x} = Ax + Bu + Mf, \tag{2.5}$$

$$y = Cx, \tag{2.6}$$

where $f \in \mathbb{R}^q$ represents the *unknown* input. A well-established method to estimate the states of (2.5)–(2.6) is the *unknown input observer* (UIO), which is a linear observer that has been investigated thoroughly [12, 15, 31, 33, 65]. An UIO by Darouach et al. [36] has the structure:

$$\dot{z} = Nz + Hy + Gu, \tag{2.7}$$

$$\hat{x} = z - Fy, \tag{2.8}$$

where $z, \hat{x} \in \mathbb{R}^n$, and (N, G, H, F) are matrices that can be chosen to satisfy design criteria, and hence constitute design freedom. The design process for these matrices is as follows. Substituting (2.8) into (2.7) yields

$$P\dot{\hat{x}} = (NP + HC)\,\hat{x} + Gu, \tag{2.9}$$

where $P = I_n + FC$. Then define the error $e = P\,(\hat{x} - x)$. Therefore, pre-multiplying (2.5) with P and then subtracting it from (2.9) yields the error system for UIO (2.7)–(2.8) (which describes the performance of the UIO):

$$\dot{e} = Ne + (NP + HC - PA)x + (G - PB)u - PMf, \tag{2.10}$$

For error e to converge independently of f (and therefore $\hat{x} \to x$), P needs to be chosen such that f does not influence error (2.10), and N must be designed to be stable.

To achieve this, let N and P be given by the following:

$$N = PA - KC, \; PM = 0, \tag{2.11}$$

where $F \in \mathbb{R}^{n\times p}$ is a design matrix that is chosen such that the eigenvalues of N have negative real parts, i.e. $\lambda(N) \in \mathbb{C}^{-}$. Then, to remove the influence of x, u on e, the remaining terms in (2.10) are calculated as:

$$H = K(I_p + CF) - PAF, \tag{2.12}$$

$$F = -M(CM)^{\dagger} + Y(I_p - (CM)(CM)^{\dagger}), \tag{2.13}$$

$$G = PB, \tag{2.14}$$

where $(CM)^{\dagger}$ is the Moore-Penrose inverse for CM, i.e.

$$(CM)^{\dagger}(CM) = I_q, \tag{2.15}$$

and Y is an arbitrary matrix of appropriate dimension. Since $PM = 0$ from (2.11), the error equation (2.10) is no longer affected by the unknown input f, thus guaranteeing robust state estimation. Thus the design process can be summarised as: choose (P, H, F, G) such that (2.11)–(2.14) are satisfied. It is also found that (P, H, F, G) can be designed to satisfy (2.11)–(2.14), and K can be chosen such that N in (2.11) is stable, if and only if the following necessary and sufficient conditions hold:

- $\mathrm{rank}(CM) = \mathrm{rank}(M)$,
- all invariant zeros of (A, M, C) (if any) must be stable.

Besides UIOs, robust state estimation has been carried out using other methods as well, such through secure state estimation [64, 115] and Kalman filters [57, 69, 76].

2.3 The sliding mode observer

The application of sliding mode techniques onto observer design is an ongoing research effort with widespread attention. Sliding mode techniques first arose out of the variable structure control system (VSCS) methodology [134]. Unlike other observers where the output estimation error is fed back only linearly, the sliding mode observer (SMO) also feeds the output estimation error back via a non-linear switching function. This difference results in the SMO being able to force its output estimation error to zero in finite time (unlike in linear observers where the output estimation error converges asymptotically). While the dynamics of the system are constrained onto this *sliding manifold*, the induced reduced-order motion has been shown to be insensitive to uncertainties implicit in the input channel [46]. This robustness has been attributed to the non-linear switching function [121]. The SMO was first introduced by Utkin in 1977 [133], and since then has garnered intense research interest. Slotine et al. [118, 119] later fed the output estimation error in both a linear and a discontinuous manner, with the aim of maximising the *sliding patch* [14], which is the region in which the SMO can attain sliding motion. Walcott and Zak [139, 140] later identified the conditions for the existence of the SMO, as well as proposed a set of design procedures. Edwards and Spurgeon [44] improved on these works by proposing a simpler design method (where the location of the poles of the observer can be chosen) as well as investigating the necessary and sufficient conditions for robust state estimation in terms of the original system matrices. Tan and Edwards [127] then developed a systematic pole placement methodology using Linear Matrix Inequalities (LMIs), where the poles of the SMO can be assigned based on required specifications. Yan and Edwards [149, 150] extended these findings into Lipschitz non-linear systems. Fridman [51, 116] investigated the use of higher-order sliding modes to expand the applicability of SMOs into systems with relative degree higher than one with respect to the faults (that is, $\dot{y}$ does not explicitly contain f [67]). In addition to being a very active area of research, SMOs have also been applied onto various practical systems, such as monitoring the battery state of charge for electric vehicles [23], flight control of a quadrotor [9], state estimation for synchronous generators [108] and piezoeletric actuator systems [2], fault-tolerant control schemes for over-actuated aircraft [105], hypersonic missiles in the terminal phase [162], and for a network of thermal and hydroelectric power plants [107].

There has been extensive work into comparing the performance of SMOs with other types of observer. Walcott et al. [138] reviewed four types of non-linear state observation techniques: a Lie-algebraic approach to observer canonical form, an extended linearisation approach, a Thau observer, and a SMO. It was found that the SMO gave the best performance from the comparative example considered. Chen and Dunnigan [20] compared the performance of the SMO with the standard and extended Kalman filters. The SMO was found to perform better than the Kalman filters on several performance metrics: it was more robust towards uncertainties, simpler to implement, the dynamic performance can be altered, and no knowledge of the noise statistics was required. Edwards and Tan [42, 47] compared the SMO and UIO; it was found that the UIO had more restrictive existence conditions when compared to the SMO. Furthermore, the SMO was found to have performed better, and has a well-defined design methodology allowing the optimal choice of observer design parameters. Zhang [166] investigated the performance of the Luenberger observer, SMO, and extended Kalman filter on both simulated and actual sets of sensorless induction machine drives. It was found that the SMO was more practically applicable, and performed the best in most aspects out of the three schemes.

2.3.1 Sliding mode observers for unknown input estimation

Aside from being robust towards uncertainties, the SMO is capable of reconstructing (i.e. providing an estimate of) the unknown input. This is especially useful when detecting incipient signals or slow drifts that may not be immediately obvious. By using the principle of the equivalent output injection signal, Edwards et al. [42] developed a scheme that uses the SMO for unknown input estimation. This ability is useful in reconstructing unpermitted deviations (which constitute unknown inputs) occurring in properties or parameters of the system from acceptable or standard conditions (hereafter referred to as *faults*) [66]. Thus SMOs are widely used for fault detection, identification, and reconstruction, where faulty behaviour can be diagnosed which is vital for corrective action to be taken in a timely manner [103]. Comparing this to conventional fault detection methods using filters, where a residual is generated from the difference between measured values of the variables and their estimates [11] to raise alarms above a certain threshold is reached (and a fault is said to have occurred) [72], it is obvious that SMOs provide more information regarding the fault, allowing for more precise corrective action to be taken. It was also found that the SMO performs better than the UIO for fault reconstruction, and requires less restrictive conditions [47].

Research into utilising SMOs for fault reconstruction is very active across many fields: for example, fault reconstruction in non-linear systems [147, 149, 150], in terminal sliding mode schemes [86, 87, 130], and in SMOs with higher order sliding modes [51, 116].

2.3.2 Robust fault reconstruction

Recall that the fault reconstruction scheme is designed based on the mathematical model of the system in (2.5)–(2.6). In addition to the fault signal f (which we want to reconstruct), there may also be other modelling uncertainties [75], parasitic dynamics, parameter variations [154], or external influences (hereafter collectively referred to as *disturbances*). The system would be therefore better represented as

$$\dot{x} = Ax + Bu + Mf + Q\xi, \tag{2.16}$$

$$y = Cx, \tag{2.17}$$

where $\xi \in \mathbb{R}^h$ represents the disturbances, which are unknown. These disturbances may corrupt the fault reconstruction, resulting in false alarms being raised, or even worse, the effects of the fault being masked so the fault remains undetected [3]. Tan and Edwards [128] proposed a method using linear matrix inequalities (LMIs) and the Bounded Real Lemma [26, 52] to design the gains of the observer such that the effect of ξ on the reconstruction of f is bounded by a positive scalar, thus achieving *robust fault reconstruction*. This method has proven popular and has been applied for fault reconstruction in aircraft and aerospace systems [7, 82, 105].

2.3.3 Sliding mode observers in cascade

A major limitation of the scheme by Tan and Edwards [128] was that the SMO scheme required the system to have relative degree of one with respect to the faults (that is, $\dot{y}$ explicitly contains f [67]). This is due to the design of the SMO: the dynamics of f must be fully captured by the switching term. Components of f which are not captured by the sliding term (i.e. faults which affect only non-output states) cannot be reconstructed with a single SMO.

To circumvent this problem, Tan and Edwards [126] presented two methods based on the SMO scheme by [44], where two SMOs in cascade are used to reconstruct faults. It was found that signals from the first observer formed the output of an analytical system that treats the faults in the original system as unknown inputs. Thus the second SMO was used to reconstruct the faults. Ng et al. [91, 92] (shown in Figure 2.2) used two SMOs in cascade for robust fault reconstruction in systems with relative degree higher than one with respect to f (that is, $\dot{y}$ does not explicitly contain f [67]).

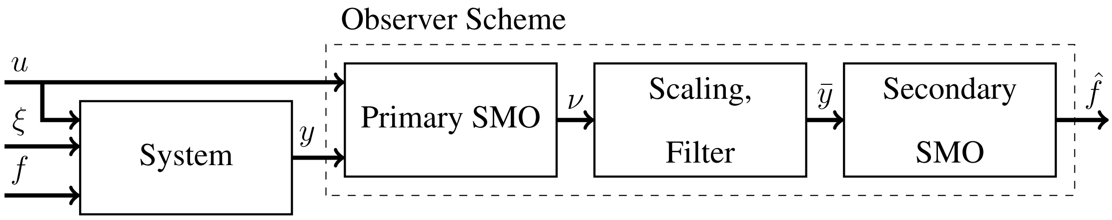

Figure 2.2: Schematic diagram of the scheme proposed by Ng et al. [91].

Thus the scheme using SMOs in cascade is applicable to more systems than schemes using only one SMO (which require the system to be relative degree one with respect to f) [125]. Tan et al. [6, 93, 94, 129] later extended the findings to include an arbitrary number of SMOs. It was found that the existence conditions of the scheme could be relaxed by increasing the number of SMOs in cascade (up until a certain number which is less than or equal to the number of states in the original system). Edwards et al. [43] perform fault reconstruction for a civil aircarft model in the presence of uncertainties (arising due to aerodynamic derivatives) utilising three SMOs in cascade. Aside from fault reconstruction, SMOs in cascade have also been investigated for functional state estimation [71]. Zhang et al. [164] used two SMOs in cascade to estimate the load torque and its derivative for feedback into a proportional-integral sliding mode controller for the servo of a permanent magnet synchronous motor system.

2.4 Descriptor systems

While the state-space representation is a simple way to represent systems containing only differential equations, there are systems that also contain algebraic equations representing links between subsystems of differential equations [38, 41]. Both types of equations often appear in systems where variables are interrelated but evolve based on different time domains [81, 90]. These systems are better represented by the *descriptor system representation* (also known as singular systems, generalised systems, or differential algebraic systems [38]), with the structure

$$E\dot{x} = Ax + Bu, \tag{2.18}$$

$$y = Cx, \tag{2.19}$$

where $E \in \mathbb{R}^{n\times n}$ is generally rank-deficient, i.e. $\operatorname{rank}(E) < n$. In the case where $E = I_n$, it can be seen that (2.18)–(2.19) reduces to the regular state-space system representation in (2.1)–(2.2), thus showing that descriptor systems are more general than the standard state-space [81]. Descriptor systems exist across many different domains. Duan [41] showed several examples where constrained mechanical systems, electric circuits, and a robot arm system are represented as descriptor systems. Chan et al. [17] modelled a chemical mixing tank as a descriptor system, where the volumetric flow rates and concentrations are related via mass balances. Yang et al. [153] used the descriptor system representation to model the charge-flux description of a RLC circuit containing a capacitor with non-linear characteristics. In the field of population ecology, descriptor systems have been used to represent the Leslie population growth model [40]. Ding et al. [38] investigated the uniform exponential admissibility problem for switched descriptor systems with time-delays, which is prevalent in electrical networks. Cantó et al. [13] studied the relationship between the positive $N-$periodic descriptor system in discrete time and its associated invariant systems, which could be used to model impulsive behaviour in biological phenomena. In economics, Gandolfo [53] describes the Leontief dynamic model of multi-sector economy using the descriptor system representation. Descriptor systems have also been used extensively in the field of fuzzy logic to represent Takagi-Sugeno fuzzy descriptor systems [165]. Kawai and Hori [70] presented a method to convert regular continuous-time descriptor systems with initial conditions consistent with its input into the discrete-time domain using a general mapping method.

2.4.1 Observers and observability for descriptor systems

The concepts of observability and detectability are central for observer designs because these concepts dictate the existence of the observer scheme. Unlike state-space systems, descriptor systems have multiple measures of observability [30]. Apart from the finite dynamic modes that also exist in a regular state-space system, descriptor systems also contain *infinite dynamic modes* as well as infinite non-dynamic modes [49]. These infinite dynamic modes may generate undesired impulse behaviour, which have direct implications on the observability of the descriptor system. These impulsive behaviours were first studied by Verghese et al. [137] and Yip and Sincovec [159], and these findings were extended into the time domain by Cobb [29] using distribution theory. The effect of impulsive behaviours on the observability of descriptor systems received much research attention [63, 89, 114, 123, 141].

The observability of a descriptor system is generally governed by the observability of its finite and infinite dynamic modes [62], while the infinite non-dynamic modes do not have known implications [49]. An established criterion for finite observability is

$$\operatorname{rank}\begin{bmatrix} sE - A \\ C \end{bmatrix} = n, \tag{2.20}$$

while the criterion for infinite observability is

$$\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n. \tag{2.21}$$

Finite observability is necessary (but not sufficient) for the existence of a Luenberger observer [61], while satisfying both criteria implies that the system is *completely* (or globally) *observable* [24]. This assumption is quite common: for example, global observability is assumed in work on adaptive observers for a class of multi-input multi-output linear descriptor systems [5], for sensor fault reconstruction in non-linear singularly perturbed systems [32], in designing a non-linear Luenberger-like observer for non-linear descriptor systems [168], for fault diagnosis [58] and estimation [144] in linear parameter varying descriptor systems affected by time delays, in an active fault tolerant control scheme based on a proportional-derivative extended state observer for linear parameter varying descriptor systems [113], and in a Takagi-Sugeno observer scheme for discrete-time non-linear descriptor systems [48]. This assumption is also made in fault reconstruction schemes for state-space systems that re-express the system as a descriptor system by augmenting the original state and fault vectors. These schemes make assumptions that result in the re-expressed descriptor system being completely observable. This method has been applied in many schemes, such as for state and fault reconstruction in a class of Itô stochastic systems [78], in Lipschitz non-linear systems affected by actuator and sensor faults and uncertainties [16, 120], and in discrete-time systems [39, 59, 79].

A SMO to estimate states and faults for descriptor systems was first developed by Yeu et al. [156, 157]. Yu and Liu [160] then investigated the conditions required for the feasibility of the observer by Yeu et al. [157]. It was found that the observer by Yeu et al. requires (2.21) to be satisfied, i.e. the system needs to be *infinitely observable*.

2.4.2 Non-infinitely observable descriptor systems

While only finite observability is generally necessary for the existence of an observer [61], infinite observability is also usually assumed to satisfied (i.e. (2.21) holds). This additional requirement may be restrictive, as it stipulates that the output distribution matrix C needs to have a certain rank and structure (to satisfy (2.21)). This physically translates to the system requiring a certain number of sensors or measured outputs; which increases the cost and complexity of the system. Even though infinite observability is a restrictive assumption, it is still assumed in many different work in the literature [5, 34, 39, 54, 55, 58, 77, 78, 104, 143, 153, 156, 157, 160, 167, 168]. Thus, infinite observability and non-infinitely observable descriptor systems (NIODS) form an important frontier in current observer research [122], and has garnered considerable attention recently. Yang et al. [151] described the design for non-linear generalised proportional-integral (PI) and non-linear generalised proportional-integral-derivative (PID) observers for a class of non-linear NIODS. Zerrougui et al. [163] designed $\mathcal{H}_\infty$ observers for a class of Lipschitz continuous non-linear descriptor systems. Yang et al. [152] proposed a robust observation scheme for different classes of descriptor systems such that the non-linearities are slope-restricted using $\mathcal{H}_\infty$ observers designed using LMIs. Darouach et al. [35] proposed a functional observer for NIODS designed using LMIs in both continuous time and discrete time. These schemes however are not capable of actuator fault reconstruction. Nikoukhah et al. [97] investigated the design of observers for general descriptor systems where an augmented descriptor system treating faults as states was formulated, so the problem becomes a standard observer problem. The feasibility of the scheme was analysed in terms of the augmented system matrices, but no analysis was carried out in terms of the original system matrices. Darouach et al. [4, 102] have studied various forms of observer design for NIODS, but these schemes all require the fault to be constant. Gao and Ding [55] presented a scheme to estimate states and faults in Lipschitz non-linear NIODS. Ying et al. [158] investigated fault reconstruction for proportional-derivative (PD) observer design for descriptor systems affected by faults and disturbances. Using a disturbance observer Yao et al. [155] proposed an anti-disturbance control scheme for non-infinitely observable descriptor Markovian jump systems. These schemes however assume that the fault is piecewise continuous. It can therefore be established that existing observer schemes for NIODS utilising linear observers either do not consider fault reconstruction, or place restrictive assumptions on the dynamics of the faults that can be reconstructed.

The SMO for NIODS has also attracted much research interest: Ooi et al. [100] first proposed a technique to re-express certain states of a NIODS as unknown inputs, thereby reformulating the NIODS as an infinitely observable reduced-order system compatible with the observer by Yeu et al. [157]. Ooi et al. [101] later built on this technique that also re-expresses some states in terms of other signals to reduce the number of states treated as unknown inputs, which reduces the number of required measured outputs and increasing the applicability of the scheme. Ooi et al. [99] and Yu et al. [161] later developed schemes utilising SMOs in cascade to relax the conditions required for the schemes to be feasible. Mellucci et al. [85] developed a second-order multi-variable super-twisting SMO to reconstruct a class of load-altering faults in a (non-infinitely observable) power network. These five schemes however did not consider the effects of disturbances on the fault reconstruction, which could result in false alarms being raised, or even mask the effects of the fault. Thus one direction underlying the work in the thesis is the design of the gains (which constitute design freedom) in the SMO such that the fault reconstruction is robust towards disturbances.

2.5 Conclusion

The concept of an observer, which underlies the entirety of the work presented in this thesis, has been detailed. State estimation is an economical method to derive information regarding unmeasured states. Robust state estimation was then established. A type of observer, namely the sliding mode observer, was introduced. It has been shown (from the literature) that the sliding mode observer outperforms linear observers, and are robust towards uncertainties. This leads the sliding mode observer being utilised to not only estimate states, but to also reconstruct unknown inputs. The sliding mode observer has then been developed for use in robust fault reconstruction and cascaded observer schemes. Next, the descriptor system (which is a more general representation compared to the state-space representation) and its multiple measures of observability were introduced. Many observer schemes assume that the system is infinitely observable, which may be a restrictive assumption as they potentially require more sensors to be feasible. Existing sliding mode observer schemes for non-infinitely observable descriptor systems are however scarce, are not robust towards disturbances, and/or assume that the fault has certain dynamics. This thesis thus aims to contribute in this area by developing sliding mode observer schemes capable of robust fault reconstruction for a class of non-infinitely observable descriptor systems with more relaxed feasibility conditions than those found in the literature.

Chapter 3

The sliding mode observer for descriptor systems

3.1 Introduction

The sliding mode observer (SMO) [8] is a type of observer that utilises two forms of feedback (both of which are dependent on the output estimation error): the first form of feedback is linear, while the second is a nonlinear discontinuous switching term [136]. The SMO is able to estimate states, even if the model of the system (which is used to design the observer) is inaccurate, or if the system is influenced by unknown inputs or disturbances [42, 116]. The aim of this chapter is to describe existing SMO schemes for descriptor systems, which will form the basis for the remainder of the thesis.

This chapter first presents the Yeu-Kim-Kawaji SMO (otherwise known as the observer by Yeu et al.) [157], which is a SMO scheme for estimating states and faults in a descriptor system, subject to the system being *infinitely observable* in §3.2. Two SMO schemes which circumvent this requirement (and therefore estimate the states and faults for non-infinitely observable descriptor systems (NIODS)) are then analysed in detail in §3.3. Finally, a simulation example to demonstrate the scheme to estimate states and faults in a NIODS is included in §3.4.

3.2 SMO for descriptor systems

The proofs in this subsection are original work which underlie the work by Yeu et al. [157] presented in §3.2.1. Consider the following descriptor system:

$$E\dot{x} = Ax + Bu + Mf, \tag{3.1}$$

$$y = Cx, \tag{3.2}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $M \in \mathbb{R}^{n\times q}$, and $C \in \mathbb{R}^{p\times n}$. The vectors x, u, y, and f represent the states, measurable inputs, outputs, and faults, respectively. Only u and y are measurable. Assume without loss of generality that E is rank-deficient, i.e. $\mathrm{rank}(E) = r < n$, and that M and C are full-column rank and full-row rank, respectively, i.e. $\mathrm{rank}(M) = q$, $\mathrm{rank}(C) = p$. The fault signal f represents abnormal behaviour in the system that needs to be estimated to facilitate remedial action. The following lemma introduces the descriptor system equivalent of similarity transformations, which would be used to transform system (3.1)–(3.2) into a form facilitating further analysis in Proposition 3.1.

Lemma 3.1 *Let state equation* (3.1) *be pre-multiplied by* T_1*, state* x *be transformed such that* $x \mapsto T_2 x$*, and fault* f *be transformed such that* $f \mapsto T_3 f$*, where* T_1*,* T_2*, and* T_3 *are non-singular matrices (referred to hereafter as a* state equation transformation*, a* state transformation*, and a* fault transformation*, respectively). System* (3.1)–(3.2) *will be transformed into the form*

$$(T_1 E T_2^{-1})(T_2 \dot{x}) = (T_1 A T_2^{-1})(T_2 x) + (T_1 B)u + (T_1 M T_3^{-1})(T_3 f), \tag{3.3}$$

$$y = (C T_2^{-1})(T_2 x). \tag{3.4}$$

Note that the matrices (E, A, B, M, C) *have been transformed. These transformations are used to change these matrices into forms that are more convenient for analysis and design, but preserve the dynamic properties of the overall system* (3.1)–(3.2)*. It can be seen that the inputs and outputs of the system remain unchanged. The transfer function* $G(s)$ *of a descriptor system is given by* $G(s) = C(sE - A)^{-1}B$*. Therefore, if* T_1*,* T_2*, and* T_3 *are non-singular, the transfer functions of systems* (3.1)–(3.2) *and* (3.3)–(3.4) *are identical. Thus, the dynamic properties such as the inputs, outputs, finite poles (the finite values of* $s \in \mathbb{C}$ *satisfying* $|sE - A| = 0$*), and dynamical order (the ranks of* E *and* $T_1 E T_2^{-1}$*) of both systems are the same, and systems* (3.1)–(3.2) *and* (3.3)–(3.4) *are said to be* equivalent. ♯

Introduce a non-singular matrix $T_a = \begin{bmatrix} N_C^T \\ C \end{bmatrix}$, where N_C spans the null-space of C, i.e.

$CN_C = 0$. Hence

$$E \mapsto ET_a^{-1} = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix}, \; C \mapsto CT_a^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}. \tag{3.5}$$

Proposition 3.1 *Assume that*

$$\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n, \tag{3.6}$$

that is, system (3.1)–(3.2) *is* infinitely observable *(the satisfaction of* (3.6) *will be addressed later in Proposition 3.3). Assumption* (3.6) *implies* $\operatorname{rank}(E_{a1}) = n - p$ *(i.e.* E_{a1} *is full-column rank) and hence there exists a non-singular matrix* T_b *such that*

$$T_b E_{a1} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}, \; T_b M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \tag{3.7}$$

where $M_2 \in \mathbb{R}^{p \times q}$*. Assume also that*

$$\operatorname{rank}(M_2) = q, \tag{3.8}$$

that is, M_2 *is full-column rank (the satisfaction of* (3.8) *will also be addressed later in Proposition 3.3). Then there exist transformations introduced in Lemma 3.1 for system* (3.1)–(3.2) *such that system* (3.1)–(3.2) *would have the following structure:*

$$\begin{bmatrix} I_{n-p} & \bar{E}_1 \\ 0 & \bar{E}_2 \end{bmatrix} \dot{x} = \begin{bmatrix} \bar{A}_1 & \bar{A}_2 \\ \bar{A}_3 & \bar{A}_4 \end{bmatrix} x + \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix} u + \begin{bmatrix} 0 \\ \bar{M}_2 \end{bmatrix} f, \tag{3.9}$$

$$y = \begin{bmatrix} 0 & I_p \end{bmatrix} x, \tag{3.10}$$

$$\bar{E}_2 = \begin{bmatrix} E_{21} \\ E_{22} \end{bmatrix}, \; \bar{A}_3 = \begin{bmatrix} A_{31} \\ A_{32} \end{bmatrix}, \; \bar{M}_2 = \begin{array}{c} \begin{bmatrix} I_q \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; q \\ \updownarrow \; p-q \end{array}, \; x = \begin{array}{c} \begin{bmatrix} x_1 \\ y \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; n-p \\ \updownarrow \; p \end{array}, \tag{3.11}$$

where $\bar{A}_1 = A_1 - M_1 A_{31}$*, the partitions of* $\bar{E}_2, \bar{A}_3, \bar{M}_2$ *have the same row dimensions and the dimensions of* (E, A) *are conformable to the partitions of* x*.* ♯

Proof Using (3.5) and (3.7), the matrices E and A are transformed into

$$T_b E T_a^{-1} = \begin{bmatrix} I_{n-p} & E_1 \\ 0 & E_2 \end{bmatrix}, \; T_b A T_a^{-1} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \tag{3.12}$$

where $A_1 \in \mathbb{R}^{(n-p)\times(n-p)}$ and $A_4 \in \mathbb{R}^{p \times p}$. Recall from (3.8) that M_2 is full-column rank - hence apply QR decomposition on M_2 such that

$$T_{c1} M_2 = \begin{bmatrix} I_q \\ 0 \end{bmatrix}, \tag{3.13}$$

where T_{c1} is non-singular. Finally, define the non-singular matrix

$$T_c = \left[\begin{array}{c|cc} I_{n-p} & -M_1 & 0 \\ \hline 0 & \multicolumn{2}{c}{I_p} \end{array}\right]\begin{bmatrix} I_{n-p} & 0 \\ 0 & T_{c1} \end{bmatrix}. \tag{3.14}$$

Therefore, apply the state equation transformation T_cT_b and the state transformation T_a onto system (3.1)–(3.2) such that

$$\begin{gathered} E \mapsto (T_cT_b)\, ET_a^{-1},\ A \mapsto (T_cT_b)\, AT_a^{-1},\ B \mapsto (T_cT_b)\, B, \\ M \mapsto (T_cT_b)M,\ C \mapsto CT_a^{-1}. \end{gathered} \tag{3.15}$$

It can be seen from the structures of E in (3.12) and M_2 from (3.13) that the system in the coordinates of (3.15) is identical to the structures given in (3.9)–(3.11), thus completing the proof. ■

The SMO by Yeu et al. [157] is now designed based on (E, A, M, C) and driven by u and y to estimate x and f.

3.2.1 The SMO by Yeu et al.

The work in this subsection has been presented in [157]. The SMO by Yeu et al. [157] to estimate the states and faults for descriptor systems has the following structure:

$$\dot{z} = (RA - G_lC)\, z - RBu - (G_l\,(I_p - CV) + RAV)\, y - G_n\nu, \tag{3.16}$$

$$\hat{x} = Vy - z, \tag{3.17}$$

$$\nu = -\rho\frac{e_y}{\|e_y\|},\ e_y = C\hat{x} - y \tag{3.18}$$

where $R \in \mathbb{R}^{n\times n}$ is invertible, $V \in \mathbb{R}^{n\times p}$, and $\rho \in \mathbb{R}^+$. The gain matrices $G_l, G_n \in \mathbb{R}^{n\times p}$ are chosen to have the structures

$$G_l = \begin{bmatrix} A_2 + R_2A_4 \\ H + R_4A_4 \end{bmatrix},\ G_n = \begin{array}{c l} \begin{bmatrix} 0 \\ G_{n2} \end{bmatrix} & \begin{array}{l} \updownarrow\ n-p \\ \updownarrow\ p \end{array} \end{array}, \tag{3.19}$$

where $H \in \mathbb{R}^{p\times p} > 0$ and $G_{n2} = G_{n2}^T > 0$ is non-singular. Pre-multiply (3.9) with R and add $V\dot{y}$ to both sides to obtain

$$RE\dot{x} + V\dot{y} = (RE + VC)\,\dot{x} = RAx + RBu + RMf + V\dot{y}. \tag{3.20}$$

Suppose $RE + VC = I_n$. Equation (3.20) becomes

$$\dot{x} = RAx + RBu + RMf + VC\dot{x}. \tag{3.21}$$

Corollary 3.1 *If* R, V *from* (3.16)–(3.17) *are chosen such that* $RE + VC = I_n$*, then* R, V *will have the following structures:*

$$R = \begin{bmatrix} I_{n-p} & R_2 \\ 0 & R_4 \end{bmatrix}, \; V = \begin{bmatrix} -\left(\bar{E}_1 + R_2\bar{E}_2\right) \\ I_p - R_4\bar{E}_2 \end{bmatrix}, \tag{3.22}$$

where $|R_4| \neq 0$. ♯

Proof Since $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n$ (from (3.6)), R and V can be chosen such that

$$\begin{bmatrix} R & V \end{bmatrix}\begin{bmatrix} E \\ C \end{bmatrix} = I_n, \tag{3.23}$$

that is, $\begin{bmatrix} R & V \end{bmatrix}$ is chosen to be the Moore-Penrose inverse of $\begin{bmatrix} E \\ C \end{bmatrix}$. Partition the matrices R and V generally as follows:

$$\begin{bmatrix} R & V \end{bmatrix} = \left[\begin{array}{cc|c} R_1 & R_2 & V_1 \\ R_3 & R_4 & V_2 \end{array}\right], \tag{3.24}$$

where $R_1 \in \mathbb{R}^{(n-p)\times(n-p)}$ and $R_4 \in \mathbb{R}^{p\times p}$. By substituting the structures of E and C from (3.9)–(3.11), and R and V from (3.24) into (3.23), it can be seen R and V would take the forms given in (3.22). ■

Then substituting z from (3.17) into (3.16) yields

$$\left(V\dot{y} - \dot{\hat{x}}\right) = \left(RA - G_lC\right)\left(Vy - \hat{x}\right) - RBu - \left(G_l\left(I_p - CV\right) + RAV\right)y - G_n\nu. \tag{3.25}$$

Since $RE + VC = I_n$, (3.25) becomes

$$\dot{\hat{x}} = \left(RA - G_lC\right)\hat{x} + RBu + G_lCx + G_n\nu + VC\dot{x}. \tag{3.26}$$

Define the state estimation error as follows:

$$e = \hat{x} - x = \begin{bmatrix} e_1 \\ e_y \end{bmatrix} \begin{array}{l} \updownarrow \; n-p \\ \updownarrow \; p \end{array}, \tag{3.27}$$

Therefore, subtracting (3.21) from (3.26) yields the following error system (which characterises the performance of the observer):

$$\dot{e} = (RA - G_l C)\, e - RMf + G_n \nu. \tag{3.28}$$

Furthermore, let R_2 in (3.22) have the structure

$$R_2 = \begin{bmatrix} 0 & L \end{bmatrix}, \tag{3.29}$$

where $L \in \mathbb{R}^{(n-p)\times(p-q)}$. Substitute for G_l and G_n from (3.19) and R_2 from (3.29) into (3.28) and partition according to (3.9)–(3.10) to obtain

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_y \end{bmatrix} = \begin{bmatrix} \bar{A}_1 + LA_{32} & 0 \\ R_4\bar{A}_3 & -H \end{bmatrix} \begin{bmatrix} e_1 \\ e_y \end{bmatrix} - \begin{bmatrix} 0 \\ R_4\bar{M}_2 \end{bmatrix} f + \begin{bmatrix} 0 \\ G_{n2} \end{bmatrix} \nu. \tag{3.30}$$

Proposition 3.2 *Suppose there exists a matrix $P = P^T > 0$ that satisfies*

$$P(RA - G_l C) + (RA - G_l C)^T P < 0, \tag{3.31}$$

where $P = \begin{bmatrix} P_1 & 0 \\ 0 & G_{n2}^{-1} \end{bmatrix}$, $P_1 \in \mathbb{R}^{(n-p)\times(n-p)}$. *If ρ in* (3.18) *is chosen as follows:*

$$\rho \geq \|G_{n2}^{-1}RM\|\alpha + \zeta, \;\; \alpha \geq \|f\|_{\max}, \;\; \zeta > 0, \tag{3.32}$$

then an ideal sliding motion $(e_y, \dot{e}_y = 0)$ takes place in finite time. ♯

Proof The proof of convergence will consist of two parts: the first part aims to show that the state estimation error e is quadratically stable. Define a Lyapunov candidate function $W = e^T P e > 0$. Differentiating W with respect to time yields

$$\dot{W} = e^T \left(P\,(RA - G_l C) + (RA - G_l C)^T\, P \right) e - 2e^T PRMf + 2e^T PG_n \nu. \tag{3.33}$$

Using the structure of P and (3.30), it can be seen that

$$PRM = C^T G_{n2}^{-1} RM, \;\; PG_n = C^T. \tag{3.34}$$

By using (3.31) and (3.34), and by setting ρ to satisfy (3.32), equation (3.33) becomes

$$\dot{W} \leq -2\|e_y\| \left(\rho - \|G_{n2}^{-1}RM\|\alpha \right) \leq -2\zeta\|e_y\| < 0 \text{ for } e \neq 0, \tag{3.35}$$

which therefore shows e is quadratically stable. The next (and remaining) portion of the proof aims to show how sliding motion $(e_y, \dot{e}_y = 0)$ is induced. Define another Lyapunov candidate function $W_y = e_y^T G_{n2}^{-1} e_y > 0$, where

$$G_{n2}^{-1} H + H^T G_{n2}^{-1} > 0. \tag{3.36}$$

Differentiating W_y with respect to time yields

$$\begin{aligned}\dot{W}_y &= -e_y^T \left(G_{n2}^{-1} H + H^T G_{n2}^{-1} \right) e_y + 2 e_y^T G_{n2} R_4 (\bar{A}_3 e_1 - \bar{M}_2 f) + 2 e_y^T \nu \\ &\leq 2 \|e_y\| \|G_{n2}^{-1} R_4 \bar{A}_3 e_1\| - 2 \|e_y\| \left(\rho - \|G_{n2}^{-1} R_4 \bar{M}_2\| \alpha \right) \\ &\leq -2 \| \left(\zeta - \|G_{n2}^{-1} R_4 \bar{A}_3 e_1\| \right). \end{aligned} \tag{3.37}$$

Since e is quadratically stable, $\zeta > \|G_{n2}^{-1} R_4 \bar{A}_3 e_1\|$ in finite time, and (3.37) becomes

$$\dot{W}_y \leq -2 \zeta_0 \|e_y\|, \tag{3.38}$$

where $\zeta_0 = \zeta - \|G_{n2}^{-1} R_4 \bar{A}_3 e_1\| > 0$. Equation (3.38) is the so-called *reachability condition* which implies that e_y will converge to zero in finite time (as will be shown after this). Notice also that

$$\|e_y\|^2 = \left(\sqrt{G_{n2}^{-1}} e_y \right)^T G_{n2} \left(\sqrt{G_{n2}^{-1}} e_y \right) \geq \lambda_{\min} \left(G_{n2} \right) \| \sqrt{G_{n2}^{-1}} e_y \|^2 = \lambda_{\min} \left(G_{n2} \right) W_y. \tag{3.39}$$

which implies

$$\begin{aligned} \dot{W}_y = \frac{dW_y}{dt} &\leq -2 \zeta_0 \sqrt{\lambda_{\min} \left(G_{n2} \right)} \sqrt{W_y} \\ \Rightarrow \frac{1}{\sqrt{W_y}} \, dW_y &\leq -2 \zeta_0 \sqrt{\lambda_{\min} \left(G_{n2}^{-1} \right)} \, dt \\ \int_{W_y(t_0)}^{0} \frac{1}{\sqrt{W_y}} \, dW_y &\leq \int_{t_0}^{t_f} -2 \zeta_0 \sqrt{\lambda_{\min} \left(G_{n2}^{-1} \right)} \, dt \\ t_f &\leq \frac{1}{\zeta_0} \sqrt{\frac{W_y(t_0)}{\lambda_{\min} \left(G_{n2} \right)}} + t_0, \end{aligned} \tag{3.40}$$

where t_0 is the time when $\zeta > \|G_{n2}^{-1} R_4 \bar{A}_3 e_1\|$ is first achieved, and t_f is the time when sliding motion is attained. Therefore, if the reachability condition (3.38) is satisfied then e_y converges and remains at zero (i.e. $e_y, \dot{e}_y = 0$), and a *sliding motion* is deemed to have occurred. Since $t_f < \infty$, sliding motion will be achieved in finite time, thus proving the proposition. ∎

Remark 3.1 *Proposition 3.2 implies that $\bar{A}_1 + LA_{32}$ can be made stable. The condition required to satisfy this assumption would be addressed later in Proposition 3.4.* ♯

3.2.1.1 Reconstructing the fault

After sliding motion $(e_y, \dot{e}_y = 0)$ occurs, error system (3.30) reduces into

$$\dot{e}_1 = \left(\bar{A}_1 + LA_{32}\right) e_1, \tag{3.41}$$

$$0 = R_4 \left(\bar{A}_3 e_1 - \bar{M}_2 f\right) + G_{n2}\nu_{eq}, \tag{3.42}$$

where ν_{eq} is the *equivalent output error injection* required to maintain the sliding motion.

Remark 3.2 *When sliding motion (i.e. $(e_y, \dot{e}_y = 0)$) occurs, the switching term ν switches at high frequency between the values -1 and $+1$ [116] to maintain the sliding motion. Its low-frequency component ν_{eq} (which is the so-called equivalent output error injection) is an analytical continuous signal satisfying* (3.42) *while also maintaining $(e_y, \dot{e}_y = 0)$. It has been shown that a good approximation of ν_{eq} can be obtained by passing ν through a low-pass filter [133, 142].* ♯

Assume that $\bar{A}_1 + LA_{32}$ is stable, which results in $e_1 \to 0$ (the satisfaction of this assumption will be addressed in Proposition 3.4). Therefore (3.42) becomes

$$R_4^{-1} G_{n2}\nu_{eq} \to M_2 f. \tag{3.43}$$

Hence define a measurable signal

$$\hat{f} = \underbrace{\begin{bmatrix} I_q & 0 \end{bmatrix} T_{c1} R_4^{-1} G_{n2}}_{G_f} \nu_{eq}. \tag{3.44}$$

Recall from (3.13) that $T_{c1}M_2 = \begin{bmatrix} I_q \\ 0 \end{bmatrix}$. Therefore, by combining (3.43)–(3.44), it can be seen that

$$\hat{f} \to f, \tag{3.45}$$

that is, the scheme is able to asymptotically estimate the faults. Furthermore, since $e_1 \to 0$ (and $e_y, \dot{e}_y = 0$) in finite time, the scheme is also able to asymptotically estimate the states (and estimate the outputs in finite time).

3.2.2 Existence conditions

The theorem in this portion has been stated by Yu and Liu in [160], but the proofs underlying it are original work. The subsection is presented in this manner to ease comparison with schemes in the later portions of the thesis.

Theorem 3.1 *The SMO by Yeu et al. can estimate x and f for system* (3.1)–(3.2) *if and only if the following conditions hold:*

A1. $\operatorname{rank}\begin{bmatrix} E & M \\ C & 0 \end{bmatrix} = n + q,$

A2. $\operatorname{rank}\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix} = n + q \;\forall\; s \in \mathbb{C}^{+}.$ ♯

Proof The remainder of this subsection forms the constructive proof for Theorem 3.1. ■

In the prior subsection, the following assumptions were made during the formulation of the SMO scheme:

B1. $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n$ in (3.6), so T_b in (3.7) exists such that $T_b E_{a1} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}$,

B2. $\operatorname{rank}(M_2) = q$ in (3.8) such that T_{c1} in (3.13) exists where $T_{c1} M_2 = \begin{bmatrix} I_q \\ 0 \end{bmatrix}$,

B3. $\bar{A}_1 + LA_{32}$ is stable, so error system (3.41)–(3.42) can be reduced into (3.43).

It is therefore of interest to re-express these conditions in terms of the original system matrices so that designers can know from the outset whether the scheme is applicable.

Proposition 3.3 *Condition A1 is necessary and sufficient to satisfy B1–B2.* ♯

Proof Using the structures in (3.5), the left-hand side (LHS) of A1 can be simplified as:

$$\operatorname{rank}\begin{bmatrix} E & M \\ C & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} E_{a1} & M \end{bmatrix}}_{\bar{E}_m} + p, \tag{3.46}$$

which (together with the right-hand side (RHS) of A1) implies

$$\operatorname{rank}\left(\bar{E}_m\right) = n - p + q. \tag{3.47}$$

To show the necessity of A1, assume that it does not hold, i.e.

$$\operatorname{rank}\left(\bar{E}_m\right) < n - p + q. \tag{3.48}$$

Since $E_{a1} \in \mathbb{R}^{n\times(n-p)}$, (3.48) would imply that it is impossible to attain both $\operatorname{rank}\left(E_{a1}\right) = n - p$ and $\operatorname{rank}(M) = q$ (it is possible either condition is met, but never both). This would in turn imply that it is impossible for both B1–B2 to be satisfied (it is possible either B1 or B2 is still satisfied, but never both), thus proving the necessity of A1.

On the other hand, if A1 is satisfied, since $\bar{E}_m \in \mathbb{R}^{n\times(n-p+q)}$, $E_{a1} \in \mathbb{R}^{n\times(n-p)}$ and $p \geq q$, (3.47) implies $\operatorname{rank}\left(E_{a1}\right) = n - p$ (and therefore B1 is satisfied). Hence T_b (which is non-singular) in (3.7) exists, and the LHS of (3.47) can be simplified into

$$\operatorname{rank}\begin{bmatrix} E_{a1} & M \end{bmatrix} = \operatorname{rank}\left(T_b\begin{bmatrix} E_{a1} & M \end{bmatrix}\right) = \operatorname{rank}\begin{bmatrix} I_{n-p} & M_1 \\ 0 & M_2 \end{bmatrix} = \operatorname{rank}\left(M_2\right) + n - p, \tag{3.49}$$

which (together with the RHS of (3.47)) implies $\operatorname{rank}\left(M_2\right) = q$, which shows B2 is also satisfied. Thus the sufficiency of A1 is shown. ■

Proposition 3.4 *For B3 to be satisfied, A2 is necessary and sufficient.* ♯

Proof Expand the LHS of A2 using the structures in (3.9)–(3.11) to obtain

$$\operatorname{rank}\underbrace{\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix}}_{R(s)} = \operatorname{rank}\begin{bmatrix} sI_{n-p} - \bar{A}_1 & s\bar{E}_1 - \bar{A}_2 & 0 \\ -A_{31} & sE_{21} - A_{41} & I_q \\ -A_{32} & sE_{22} - A_{42} & 0 \\ 0 & I_p & 0 \end{bmatrix}, \tag{3.50}$$

where $R(s)$ is the Rosenbrock matrix of (E, A, M, C), and the values of s that make $R(s)$ lose rank are the zeros of the system (E, A, M, C) [109]. Further expanding $R(s)$ yields

$$\operatorname{rank}\left(R(s)\right) = \operatorname{rank}\underbrace{\begin{bmatrix} sI_{n-p} - \bar{A}_1 \\ A_{32} \end{bmatrix}}_{R_2(s)} + p + q. \tag{3.51}$$

From the Popov-Hautus-Rosenbrock (PHR) rank test [60], if the values of s that make $R_2(s)$ (i.e. the unobservable modes of $\bar{A}_1$) are stable, the pair $\left(\bar{A}_1, A_{32}\right)$ is said to be *detectable* - hence A2 is recast as: $\left(\bar{A}_1, A_{32}\right)$ is *detectable*.

Proof of Necessity:

Condition B3 states that $\bar{A}_1 + LA_{32}$ needs to be stable, which implies $\lambda\left(\bar{A}_1 + LA_{32}\right) < 0$, i.e. $\left(\bar{A}_1, LA_{32}\right)$ is *detectable*. Notice that the detectability of $\left(\bar{A}_1, LA_{32}\right)$ depends on L, which constitutes design freedom. Hence, the requirement is recast as: matrix L exists such that $\left(\bar{A}_1, LA_{32}\right)$ is detectable. From the PHR rank test, if the values of s that make the following matrix $R_3(s)$ lose rank are stable, then $\left(\bar{A}_1, LA_{32}\right)$ is said to be *detectable*, whereby

$$R_3(s) = \begin{bmatrix} sI_{n-p} - \bar{A}_1 \\ LA_{32} \end{bmatrix} = \begin{bmatrix} I_{n-p} & 0 \\ 0 & L \end{bmatrix} \underbrace{\begin{bmatrix} sI_{n-p} - \bar{A}_1 \\ A_{32} \end{bmatrix}}_{R_2(s)}. \tag{3.52}$$

From (3.52), it follows that $\operatorname{rank}(R_3(s)) \leq \operatorname{rank}(R_2(s))$. Therefore if a value of s makes $R_2(s)$ lose rank, it will also make $R_3(s)$ lose rank, and hence the zeros of (E, A, M, C) are also the unobservable modes of $\left(\bar{A}_1, LA_{32}\right)$. This therefore shows A2 is necessary for $\left(\bar{A}_1, LA_{32}\right)$ to be detectable.

Proof of Sufficiency:

Let U be a matrix containing the generalised right-eigenvectors of $\bar{A}_1$; therefore $U^{-1}\bar{A}_1U$ is a matrix in the *Jordan canonical form*, where the diagonal elements are the real parts of the eigenvalues of $\bar{A}_1$ [98]. Pre-multiply $R_3(s)$ with $\begin{bmatrix} U^{-1} & 0 \\ 0 & I_{p-q} \end{bmatrix}$ and post-multiply with U, i.e.

$$\begin{bmatrix} U^{-1} & 0 \\ 0 & I_{p-q} \end{bmatrix} \begin{bmatrix} sI_{n-p} - \bar{A}_1 \\ LA_{32} \end{bmatrix} U = \begin{bmatrix} I_{n-p} & 0 \\ 0 & L \end{bmatrix} \begin{bmatrix} U^{-1}\left(sI_{n-p} - \bar{A}_1\right)U \\ A_{32}U \end{bmatrix}. \tag{3.53}$$

A zero of (E, A, M, C), which is an unobservable mode of $\left(U^{-1}\bar{A}_1U, LA_{32}U\right)$, will therefore appear as an element of $U^{-1}\bar{A}_1U$ where its corresponding column in $LA_{32}U$ is zero. If A2 is satisfied however, the columns within $A_{32}U$ corresponding to positive diagonal elements (which indicate unstable eigenvalues) within $U^{-1}\bar{A}_1U$ will be non-zero. Recall that L is design freedom; thus a single row within L can be chosen such that $LA_{32}U$ has non-zero elements at the columns corresponding to the diagonals of $U^{-1}\bar{A}_1U$ that are unstable in order to guarantee that the unstable modes are observable (and therefore are not a zero of (E, A, M, C)).

When $p = q$, L and A_{32} do not exist, and error system (3.41)–(3.42) becomes

$$\dot{e}_1 = \bar{A}_1 e_1, \tag{3.54}$$

$$0 = R_4 \left(\bar{A}_3 e_1 - \bar{M}_2 f \right) + G_{n2} \nu_{eq}. \tag{3.55}$$

Notice from (3.51) that

$$\operatorname{rank}(R(s)) = \operatorname{rank}\left(sI_{n-p} - \bar{A}_1 \right) + p + q, \tag{3.56}$$

which implies that if D2 is satisfied, $\bar{A}_1$ would have stable eigenvalues, and error system (3.54)–(3.55) will reduce into (3.43). Thus, the sufficiency of A2 is shown for both cases where $p > q$ and $p = q$. ■

Therefore, Propositions 3.3–3.4 have been proven: if A1–A2 are satisfied, the observer by Yeu et al. [157] can estimate x and f for system (3.1)–(3.2), thus proving Theorem 3.1. □

3.2.3 Design procedure

The design procedure for the scheme in this section is summarised as follows:

1. Check that A1–A2 hold for system (3.1)–(3.2). If not, the scheme in this subsection is not applicable.

2. Obtain the state equation transformation $T_c T_b$ and state transformation T_a from (3.5), (3.7), and (3.14).

3. Choose the matrices R and V to satisfy (3.23), where R and V have the structures given in (3.22), and R_2 has the structure in (3.29).

4. Calculate the gains G_l and G_n from (3.19).

5. Set a value for ρ in (3.32).

6. Calculate G_f and estimate f using (3.44).

3.2.4 Simulation example

The following subsection is original work. The concepts covered in the previous subsections are illustrated using the following example. Consider the two-loop electric circuit given in [41], which is described by the following dynamical model:

$$2\dot{V}_{C1} = I_1, \tag{3.57}$$

$$3\dot{V}_{C2} = I_2, \tag{3.58}$$

$$-7\dot{I}_2 = -V_{C1} + V_{C2}, \tag{3.59}$$

$$0 = V_{C1} + 4I_1 + 4I_2 - V_e, \tag{3.60}$$

where V_{C1} and V_{C2} are the voltages across the capacitors C_1, and C_2, respectively, I_1 and I_2 are the currents across the first and second loops, respectively, and V_e is the potential difference supplied by the power source. The actuator signal for V_e is generated by a faulty first-order device:

$$\dot{V}_S = -0.2V_e + 0.2V_{Sc} + f, \tag{3.61}$$

where V_{Sc} is the reference signal, and f is the fault.

3.2.4.1 System formulation

The system matrices (E, A, B, M) in the framework of (3.1)–(3.2) are

$$E = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & -7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 4 & 4 & -1 \\ 0 & 0 & 0 & 0 & -0.2 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0.2 \end{bmatrix}, \; M = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{3.62}$$

for the system variables

$$x = \begin{bmatrix} \text{voltage across capacitor 1, } V_{C1} \text{ (V)} \\ \text{voltage across capacitor 2, } V_{C2} \text{ (V)} \\ \text{current across loop 2, } I_2 \text{ (A)} \\ \text{current across loop 1, } I_1 \text{ (A)} \\ \text{supplied voltage, } V_e \text{ (V)} \end{bmatrix}, \; f \text{ (V)}, \tag{3.63}$$

$$u = \begin{bmatrix} \text{supplied voltage reference, } V_{Sc} \text{ (V)} \end{bmatrix}.$$

Assume that measurements are only available for I_1 and V_e, so C has the form

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.64}$$

The observer by Yeu et al. [157] will now be designed.

3.2.4.2 Observer design

To ease readability, the steps in §3.2.3 will be referred to in the following design of the observer scheme.

Step 1: It can quickly be verified that A1–A2 hold, so the existence of the observer by Yeu et al. [157] is guaranteed.

Step 2: The state equation transformation T_cT_b and state transformation T_a are calculated to be

$$T_cT_b = \begin{bmatrix} 0 & 0 & -0.1429 & 0 & 0 \\ -0.5 & 0 & 0 & 0 & 0 \\ 0 & -0.3333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad T_a = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.65}$$

Step 3: The matrices R and V are chosen to be

$$R = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1255 \\ 0 & 1 & 0 & 0 & 5040 \\ 0 & 0 & 1 & 0 & 4833 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right], \quad V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}. \tag{3.66}$$

Step 4: The poles of the observer were chosen as $\{-5, -5, -5, -6, -8\}$, while the poles of the sliding motion are $\{-5, -6, -8\}$. The gains G_l and G_n are therefore calculated as

$$G_l = \begin{bmatrix} 5021 & -125.5 \\ 20160 & -504 \\ 19334 & -4834 \\ 5 & -0.2 \\ 4 & 4 \end{bmatrix}, \quad G_n = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.67}$$

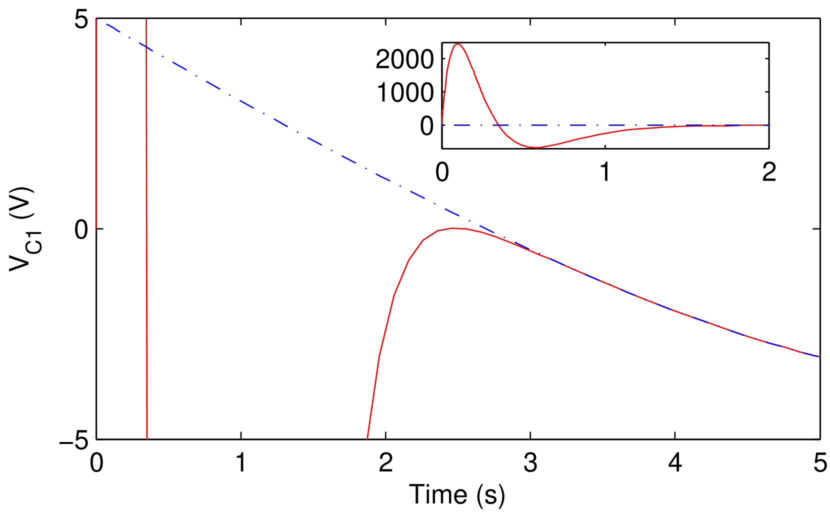

Figure 3.1: The first state (dash-dotted) and its estimate (solid).

Steps 5–6: The parameter ρ and the matrix in (3.44) are set as

$$\rho = 5, \; G_f = \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{3.68}$$

3.2.4.3 Simulation results

The initial conditions of the system was set at $x(0) = \{5, -5, 2, -4, -3\}$, while the initial condition of the observer was set at zero. The input V_{Sc} was set as a step input with a magnitude of 2. The fault signal was simulated as

$$f = 2 \sin\left(t + \frac{\pi}{3}\right) + 1. \tag{3.69}$$

Figures 3.1–3.6 show the evolution of the system states and fault, and their estimates - it can be seen that about $t = 3$ s that asymptotic tracking of the states takes place. It can be seen that the observer estimates track the states and faults well. Notice however from Figures 3.4–3.5 that the estimates for the fourth and fifth states converge onto the real states in finite time (since they are the outputs). This is consistent with Figure 3.7, which shows the output estimation error e_y against time. At approximately $t = 0.22$ s, e_y goes to and remains at zero, indicating that a *sliding motion* on the surface $\mathcal{S} = \{e : Ce = 0\}$ has taken place. Notice also from Figures 3.1–3.3 that the state estimates experience an abrupt change in dynamics at the same time: this is because the dynamics of the state estimation errors e are now governed by the *reduced order motion* in (3.41). These features are

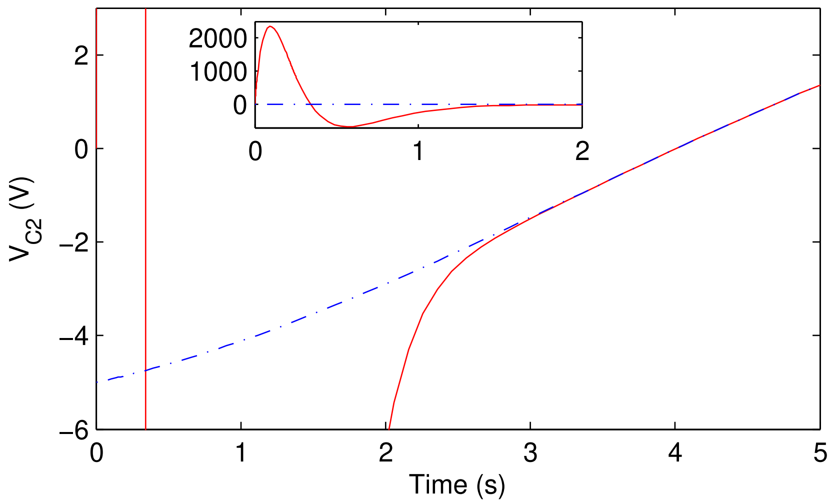

Figure 3.2: The second state (dash-dotted) and its estimate (solid).

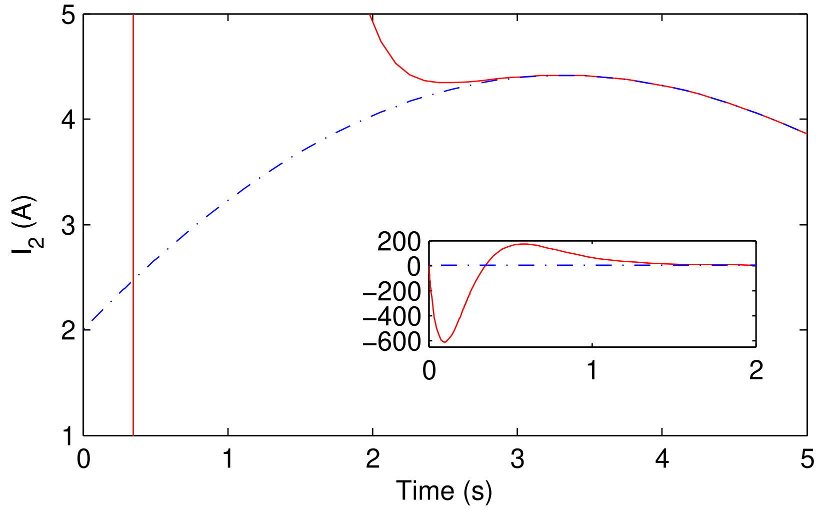

Figure 3.3: The third state (dash-dotted) and its estimate (solid).

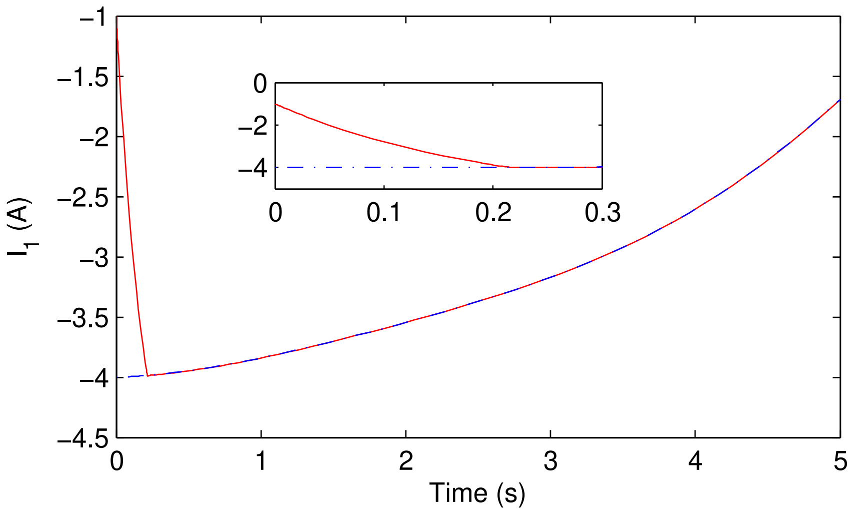

Figure 3.4: The fourth state (dash-dotted) and its estimate (solid).

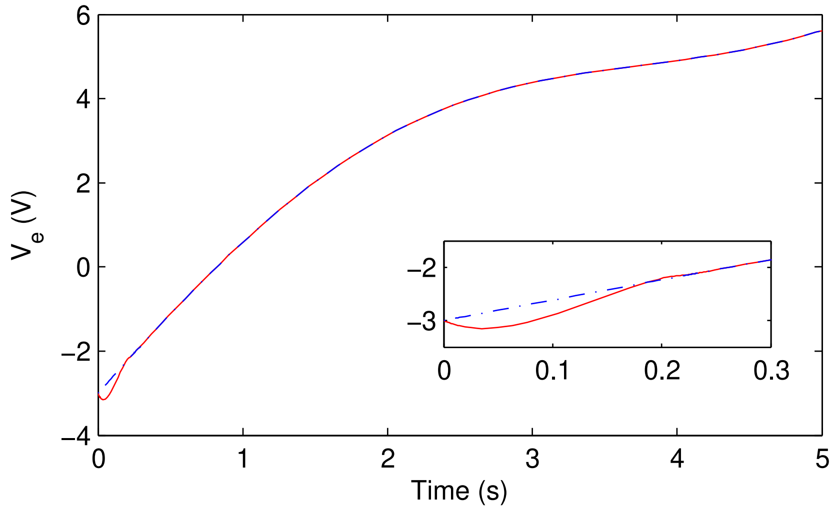

Figure 3.5: The fifth state (dash-dotted) and its estimate (solid).

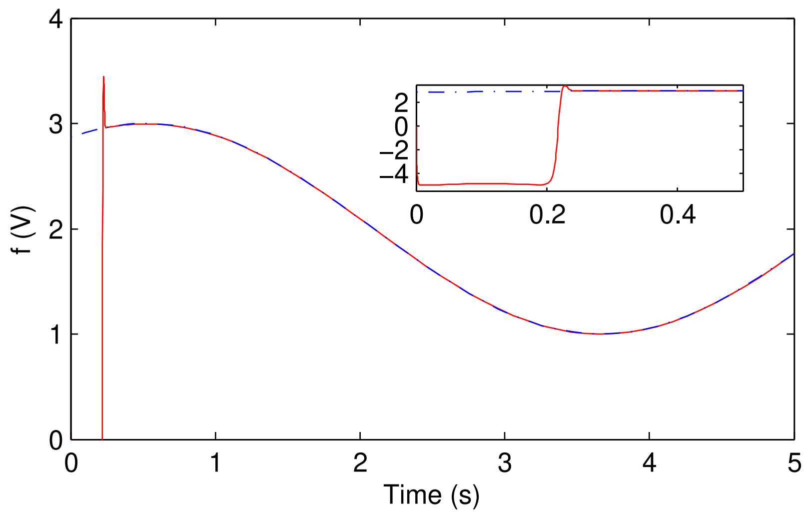

Figure 3.6: The fault (dash-dotted) and its estimate (solid).

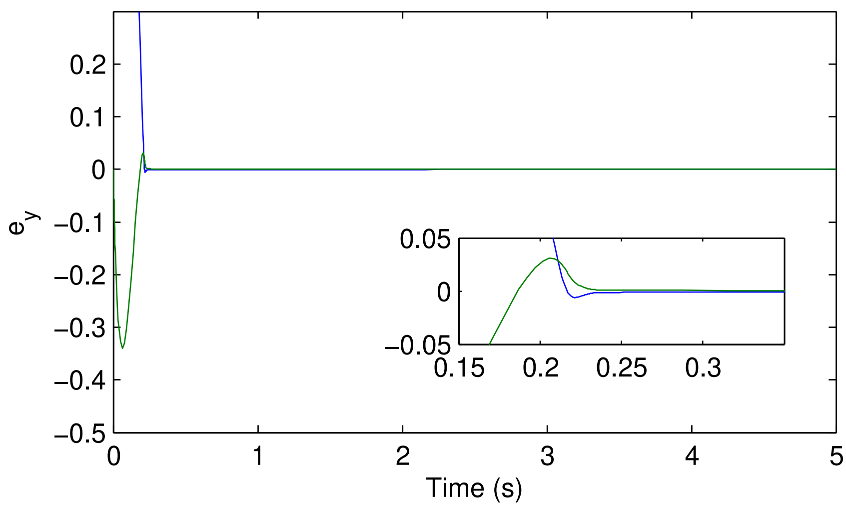

Figure 3.7: Output estimation error $e_y = \hat{y} - y$.

typical of sliding mode systems.

3.3 Non-infinitely observable descriptor systems (NIODS)

In the previous section, it has been shown that the observer by Yeu et al. [157] can estimate states and faults for a descriptor system if and only if A1–A2 are satisfied. Conditions A1–A2 however imply that the system is *infinitely observable* [49]. This requirement limits the applicability of the SMO schemes as it forms a stringent requirement on the number of states that need to be measured. Ooi et al. [100, 101] circumvented this requirement by treating certain states as unknown inputs, thus increasing the applicability of the SMO schemes as compared to the work in [157]. This section outlines the two methods devised by Ooi et al. for state and fault estimation for non-infinitely observable descriptor systems (NIODS).

3.3.1 State and fault estimation for NIODS

The work in this subsection has been presented in [100]. Consider the following descriptor system:

$$E\dot{x} = Ax + Mf, \tag{3.70}$$

$$y = Cx. \tag{3.71}$$

where $E, A \in \mathbb{R}^{n\times n}$, $M \in \mathbb{R}^{n\times q}$, and $C \in \mathbb{R}^{p\times n}$. The vectors x, y, and f represent the states, outputs, and faults, respectively. Only y is measurable. Assume generally that E is rank-deficient, i.e. $\mathrm{rank}(E) = r < n$, and that M and C are full-column rank and full-row rank, respectively, i.e. $\mathrm{rank}(M) = q$, $\mathrm{rank}(C) = p$. Furthermore, it is assumed without loss of generality that system (3.70)–(3.71) is *not infinitely observable*, i.e.

$$\mathrm{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n} < n. \tag{3.72}$$

The fault signal f represents anomalous behaviour in the system that needs to be reconstructed for corrective action to be taken. The SMO by Yeu et al. [157] presented in the previous section is however not applicable, as it requires system (3.70)–(3.71) to be infinitely observable [49, 160] $\left(\text{i.e. } \operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n\right)$, which is clearly not the case. Thus, Ooi et al. [100] proposed reformulating system (3.70)–(3.71) such that certain states are treated as unknown inputs.

Remark 3.3 *It can be seen that the term Bu, which is present in system* (3.1)–(3.2) *in the prior subsection, is missing from system* (3.70)–(3.71). *Notice however that the known input u does not appear in the error system of the observer by Yeu et al. [157] in* (3.28), *and therefore does not affect its design. As will be shown, the method by Ooi et al. [100] merely re-expresses system* (3.70)–(3.71) *(without manipulating any known inputs) and then applies the observer by Yeu et al. onto the re-expressed system, the term Bu does not affect the error analysis for the succeeding methods and is therefore omitted for brevity.* ♯

The reformulation by Ooi et al. [100] is described in the following proposition.

Proposition 3.5 *There exist transformations such that (E, A, M, C) and x can be written as follows:*

$$E = \begin{bmatrix} 0 & 0 & E_{21} \\ 0 & I_{\bar{n}-p} & 0 \\ 0 & 0 & E_4 \end{bmatrix}, \; A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \; M = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix} \begin{array}{l} \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; p \end{array}, \qquad (3.73)$$
$$C = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix}, \; x = \begin{bmatrix} x_{11} \\ x_{12} \\ y \end{bmatrix} \begin{array}{l} \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; p \end{array},$$

where the partitions of E, A, M have the same row dimensions, and the column partitions of E, A, C are conformable to the partitions of x. ♯

Proof Introduce the state transformation $T_a = \begin{bmatrix} N_C^T \\ C \end{bmatrix}$, where N_C spans the null-space of C, i.e. $N_C C = 0$, and $x \mapsto T_a x = x_a$. The following structures are obtained:

$$E \mapsto E T_a^{-1} = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix}, \; C \mapsto C T_a^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}. \qquad (3.74)$$

Since $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n}$, it follows from (3.74) that $\operatorname{rank}(E_{a1}) = \bar{n} - p$. Therefore perform singular-value decomposition (SVD) on E_{a1} such that

$$X_1^{-1} E_{a1} X_2^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & S_1 \end{bmatrix}, \tag{3.75}$$

where X_1, X_2 are orthogonal and $S_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$ is invertible. Then there would also exist two non-singular matrices X_3, X_4 with the structures

$$X_3 = \begin{bmatrix} I_{n-\bar{n}} & 0 & 0 \\ 0 & S^{-1} & 0 \\ 0 & 0 & I_p \end{bmatrix}, \; X_4 = \begin{bmatrix} I_{n-\bar{n}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\bar{n}-p} \\ 0 & 0 & I_{2p-\bar{n}} & 0 \\ 0 & I_{\bar{n}-p} & 0 & 0 \end{bmatrix}, \tag{3.76}$$

such that

$$X_3 X_4 X_1^{-1} E_{a1} X_2^{-1} = \begin{array}{c} \begin{bmatrix} 0 & 0 \\ 0 & I_{\bar{n}-p} \\ 0 & 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; p \end{array} . \tag{3.77}$$

Define two non-singular matrices $T_b = X_3 X_4 X_1^{-1}$ and $T_c = \begin{bmatrix} X_2 & 0 \\ 0 & I_p \end{bmatrix}$. Apply the state equation transformation T_b and state transformation T_c onto the system in the coordinates of (3.74) to obtain

$$E \mapsto T_b E T_c^{-1} = \begin{bmatrix} 0 & 0 & E_{c21} \\ 0 & I_{\bar{n}-p} & E_{c22} \\ 0 & 0 & E_{c4} \end{bmatrix}, \; C \mapsto C T_c^{-1} = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix}, \; x_a \mapsto T_c x_a = x_c. \tag{3.78}$$

Define a non-singular matrix $T_d = \begin{bmatrix} I_{n-\bar{n}} & 0 & 0 \\ 0 & I_{\bar{n}-p} & E_{c22} \\ 0 & 0 & I_p \end{bmatrix}$. Finally, apply the state equation transformation T_b and state transformation $T_a T_c T_d$ on system (3.70)–(3.71):

$$\begin{gathered} E \mapsto T_b E \left(T_a T_c T_d\right)^{-1}, \; A \mapsto T_b A \left(T_a T_c T_d\right)^{-1}, \; M \mapsto T_b M, \\ C \mapsto C \left(T_a T_c T_d\right)^{-1}, \; x \mapsto T_a T_c T_d x_c. \end{gathered} \tag{3.79}$$

From the structures of E and C in (3.78), it can be seen that the system in the coordinates of (3.79) is identical to the structures stipulated in (3.73). Thus, the proof is complete. ∎

The system in the coordinates of (3.73) is then reformulated as follows: treat x_{11} as an unknown input, and define

$$\bar{x} = \begin{bmatrix} x_{12} \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} f \\ x_{11} \end{bmatrix}, \tag{3.80}$$

which respectively are the state and fault of the following reduced-order system (which has been re-expressed from (3.73)):

$$\underbrace{\begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & E_4 \end{bmatrix}}_{\bar{E}} \dot{\bar{x}} = \underbrace{\begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix}}_{\bar{A}} \bar{x} + \underbrace{\begin{bmatrix} M_2 & A_{21} \\ M_3 & A_{31} \end{bmatrix}}_{\bar{M}} \bar{f}, \tag{3.81}$$

$$y = \underbrace{\begin{bmatrix} 0 & I_p \end{bmatrix}}_{\bar{C}} \bar{x}. \tag{3.82}$$

Notice that $\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$ (full-column rank), i.e. the reduced-order system (3.81)–(3.82) is infinitely observable, and thus the observer by Yeu et al. [157] can be designed based on $\left(\bar{E}, \bar{A}, \bar{M}, \bar{C}\right)$ and driven by u and y.

3.3.2 Improved state and fault estimation for NIODS

The technique used by Ooi et al. [100] in the prior subsection to re-express NIODS as infinitely observable reduced-order systems relies on certain states being treated as unknown inputs. This increases the number of unknown inputs in the reduced-order system, which must not exceed the number of states in the reduced-order system. Ooi et al. [101] built on this technique by also re-expressing some states in terms of other states before treating certain other states as unknown inputs. This reduces the number of states treated as unknown inputs, which increases the applicability of the SMO scheme. This subsection details the scheme as it was presented in [101] (with the exception of Remarks 3.5 and 3.6, which are original work added in for clarification).

3.3.2.1 Preliminary transformations

Consider the following descriptor system:

$$E\dot{x} = Ax + Mf, \tag{3.83}$$

$$y = Cx. \tag{3.84}$$

where $E, A \in \mathbb{R}^{n\times n},\ M \in \mathbb{R}^{n\times q}$, and $C \in \mathbb{R}^{p\times n}$. The vectors $x,\ y$, and f represent the states, outputs, and faults, respectively. Only y is measurable. Assume generally that E is rank-deficient, i.e. $\mathrm{rank}(E) = k < n$, and that M and C are full-column rank and full-row rank, respectively, i.e. $\mathrm{rank}(M) = q,\ \mathrm{rank}(C) = p$. Furthermore, it is assumed without loss of generality that system (3.83)–(3.84) is *not infinitely observable*, i.e.

$$\mathrm{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n} < n. \tag{3.85}$$

The fault signal f represents irregular behaviour in the system that needs to be reconstructed for remedial purposes. The SMO by Yeu et al. [157] is again not applicable, as it requires system (3.83)–(3.84) to be infinitely observable [49, 160], which is clearly not the case. Thus, Ooi et al. proposed reformulating system (3.83)–(3.84) such that certain states are treated as unknown inputs. As per Remark 3.3, the known inputs into the system do not affect the error analysis for the observer and are therefore omitted for brevity. The reformulation by Ooi et al. [101] is described in the following proposition.

Proposition 3.6 *There exist transformations such that (E, A, M, C) and x would have the structures*

$$E = \begin{bmatrix} 0 & 0 & \mathcal{E}_2 \\ 0 & I_{\bar{n}-p} & 0 \end{bmatrix},\ \mathcal{E}_2 = \begin{array}{c} \begin{bmatrix} 0 \\ \mathcal{E}_{22} \end{bmatrix} \end{array} \begin{array}{l} \updownarrow\ n-k \\ \updownarrow\ k+p-\bar{n} \end{array},\ C = \begin{bmatrix} 0 & C_3 \end{bmatrix}, \tag{3.86}$$

$$A = \begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \mathcal{A}_3 \\ \mathcal{A}_4 & \mathcal{A}_5 & \mathcal{A}_6 \\ \mathcal{A}_7 & \mathcal{A}_8 & \mathcal{A}_9 \end{bmatrix} = \begin{array}{c} \left[\begin{array}{cc|c|c} 0 & 0 & A_3 & A_4 \\ 0 & I_j & A_7 & A_8 \\ \hline A_9 & A_{10} & A_{11} & A_{12} \\ A_{13} & A_{14} & A_{15} & A_{16} \\ \hline A_{17} & A_{18} & A_{19} & A_{20} \end{array}\right] \end{array} \begin{array}{l} \updownarrow\ n-\bar{n}-j \\ \updownarrow\ j \\ \updownarrow\ \bar{n}-k \\ \updownarrow\ k+p-\bar{n} \\ \updownarrow\ \bar{n}-p \end{array}, \tag{3.87}$$

$$M = \begin{bmatrix} \mathcal{M}_1 \\ \mathcal{M}_2 \\ \mathcal{M}_3 \end{bmatrix} = \left[\begin{array}{c} M_1 \\ M_2 \\ \hline M_3 \\ M_4 \\ \hline M_5 \end{array}\right], \quad x = \left[\begin{array}{c} x_{11} \\ x_{12} \\ \hline x_2 \\ \hline y \end{array}\right] \begin{array}{l} \updownarrow\ n-\bar{n}-j \\ \updownarrow\ j \\ \updownarrow\ \bar{n}-p \\ \updownarrow\ p \end{array}, \tag{3.88}$$

where $|C_3| \neq 0,\ \operatorname{rank}(\mathcal{E}_{22}) = k+p-\bar{n}$*, the the partitions of* A, M *have the same row dimensions, and the column partitions of* E, A *are conformable to the partitions of* x. ♯

Proof Introduce the state transformation $T_a = \begin{bmatrix} N_C^T \\ C \end{bmatrix}$, where N_C spans the null-space of C, i.e. $N_C C = 0$, and $x \mapsto T_a x = x_a$. The following structures are obtained:

$$E \mapsto E T_a^{-1} = E_a = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix}, \; C \mapsto C T_a^{-1} = C_a = \begin{bmatrix} 0 & I_p \end{bmatrix}. \tag{3.89}$$

Since $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n}$, it follows from (3.89) that $\operatorname{rank}(E_{a1}) = \bar{n}-p$. Therefore perform SVD on E_{a1} such that

$$X_1 E_{a1} X_2 = \begin{bmatrix} 0 & 0 \\ 0 & I_{\bar{n}-p} \end{bmatrix}, \tag{3.90}$$

where $X_1 \in \mathbb{R}^{n\times n}$ and $X_2 \in \mathbb{R}^{(n-p)\times(n-p)}$ are non-singular. Partition X_1 as follows:

$$X_1 = \begin{bmatrix} X_{11} \\ X_{12} \end{bmatrix} \begin{array}{l} \updownarrow\ n-\bar{n}+p \\ \updownarrow\ \bar{n}-p \end{array}, \tag{3.91}$$

and therefore let

$$\bar{E}_{a2} = X_{11} E_{a2}, \; \bar{E}_{a4} = X_{12} E_{a2}. \tag{3.92}$$

By using (3.90)–(3.92), the following can be obtained:

$$X_1 \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix} \begin{bmatrix} X_2 & 0 \\ 0 & I_p \end{bmatrix} = \left[\begin{array}{cc|c} 0 & 0 & \bar{E}_{a2} \\ 0 & I_{\bar{n}-p} & \bar{E}_{a4} \end{array}\right]. \tag{3.93}$$

Since $\mathrm{rank}(E) = k$, it can be seen that $\mathrm{rank}\left(\bar{E}_{a2}\right) = k - \bar{n} + p$. Therefore, apply QR decomposition onto $\bar{E}_{a2}$ such that $X_3\bar{E}_{a2} = \mathcal{E}_2$, where $X_3 \in \mathbb{R}^{(n-\bar{n}+p)\times(n-\bar{n}+p)}$. Then define two non-singular matrices with the following structures

$$T_b = \begin{bmatrix} X_3 & 0 \\ 0 & I_{\bar{n}-p} \end{bmatrix} X_1,\ T_c = \left[\begin{array}{cc|c} I_{n-\bar{n}} & 0 & 0 \\ 0 & I_{\bar{n}-p} & \bar{E}_{a4} \\ \hline 0 & 0 & I_p \end{array}\right] \begin{bmatrix} X_2^{-1} & 0 \\ 0 & I_p \end{bmatrix}, \tag{3.94}$$

such that

$$\begin{gathered} E_a \mapsto T_b E_a T_c^{-1} = E_c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \mathcal{E}_{22} \\ 0 & I_{\bar{n}-p} & 0 \end{bmatrix},\ C_a \mapsto C_a T_c^{-1} = C_c = \begin{bmatrix} 0 & 0 & C_3 \end{bmatrix}, \\ A \mapsto T_b A (T_a T_c)^{-1} = A_c = \begin{bmatrix} A_{b1} & A_{b2} & A_{b3} \\ A_{b4} & A_{b5} & A_{b6} \\ A_{b7} & A_{b8} & A_{b9} \end{bmatrix},\ M \mapsto T_b M = M_c = \begin{bmatrix} M_{b1} \\ M_{b2} \\ M_{b3} \end{bmatrix}, \end{gathered} \tag{3.95}$$

where $A_{b1} \in \mathbb{R}^{(n-k)\times(n-\bar{n})}$, $A_{b5} \in \mathbb{R}^{(k+p-\bar{n})\times(\bar{n}-p)}$, $A_{b9} \in \mathbb{R}^{(\bar{n}-p)\times p}$, $|C_3| \neq 0$, the partitions of E_c, A_c, M_c have the same row dimensions, and the partitions of E_c, A_c, C_c have conformable column dimensions. Next, pre-multiply A_c with a non-singular matrix $T_d = \begin{bmatrix} T_o & 0 \\ 0 & I_k \end{bmatrix}$ such that $A_d = T_d A_c$, where $T_o \in \mathbb{R}^{(n-k)\times(n-k)}$ is non-singular and will be defined later in (3.115). Let A_{d1} be the top-left $(n-\bar{n}) \times (n-\bar{n})$ block of A_d, and $\mathrm{rank}\left(A_{d1}\right) = j$. Hence apply SVD on A_{d1} to obtain

$$X_5 A_{d1} X_6 = \begin{bmatrix} 0 & 0 \\ 0 & I_j \end{bmatrix}, \tag{3.96}$$

where $X_5, X_6 \in \mathbb{R}^{(n-\bar{n})\times(n-\bar{n})}$ are non-singular. Hence define two non-singular matrices with the structures

$$T_e = \begin{bmatrix} X_5 & 0 \\ 0 & I_{\bar{n}} \end{bmatrix},\ T_f = \begin{bmatrix} X_6^{-1} & 0 \\ 0 & I_{\bar{n}} \end{bmatrix}. \tag{3.97}$$

Therefore apply the state equation transformation $T_e T_d T_b$ and the state transformation $T_a T_c T_f$ such that

$$\begin{gathered} E \mapsto T_e T_d T_b E \left(T_a T_c T_f\right)^{-1},\ A \mapsto T_e T_d T_b A \left(T_a T_c T_f\right)^{-1},\ M \mapsto T_e T_d T_b M, \\ C \mapsto CE \mapsto C\left(T_a T_c T_f\right)^{-1},\ x \mapsto \left(T_a T_c T_f\right) x. \end{gathered} \tag{3.98}$$

From the structures of E_c, C_b in (3.95) and A_{d1} in (3.96), it can be seen that the structures of the systems in (3.86)–(3.88) and in (3.98) are identical, thus completing the proof. ■

Define two vectors

$$\bar{x} = \begin{bmatrix} x_2 \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} x_{11} \\ f \end{bmatrix}. \tag{3.99}$$

From the structures of $E,\ A$, and x in (3.86)–(3.88), it can be seen that x_{12} can be re-expressed as

$$x_{12} = -\begin{bmatrix} A_7 & A_8 \end{bmatrix}\bar{x} - \begin{bmatrix} 0 & M_2 \end{bmatrix}\bar{f}. \tag{3.100}$$

Thus, by treating x_{11} as an unknown input, and using re-expressing x_{12} in terms of $\bar{x}$ and $\bar{f}$, the following reduced-order system of order $\bar{n}$ (which treats $\bar{x}$ and $\bar{f}$ as its state and fault respectively):

$$\bar{E}\dot{\bar{x}} = \bar{A}\bar{x} + \bar{M}\bar{f},\ y = \bar{C}\bar{x}, \tag{3.101}$$

$$\bar{E} = \begin{bmatrix} 0 & 0 \\ 0 & \mathcal{E}_{22} \\ I_{\bar{n}-p} & 0 \end{bmatrix}, \ \bar{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{15} & A_{16} \\ A_{19} & A_{20} \end{bmatrix} - \begin{bmatrix} A_{10} \\ A_{14} \\ A_{18} \end{bmatrix}\begin{bmatrix} A_7 & A_8 \end{bmatrix}\bar{x}, \tag{3.102}$$

$$\bar{M} = \begin{bmatrix} A_9 & M_3 \\ A_{13} & M_4 \\ A_{17} & M_5 \end{bmatrix} - \begin{bmatrix} A_{10} \\ A_{14} \\ A_{18} \end{bmatrix}\begin{bmatrix} 0 & M_2 \end{bmatrix}, \ \bar{C} = \begin{bmatrix} 0 & C_3 \end{bmatrix}. \tag{3.103}$$

Notice that $\mathrm{rank}\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$; this implies that the reduced-order system (3.101) is *infinitely observable* [49], and the observer by Yeu et al. [157] can be applied onto system (3.101) to estimate $\bar{x}$ and $\bar{f}$, thereby estimating x and f.

Remark 3.4 *In the previous scheme by Ooi et al. [100] in §3.3.1,* $\begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix}$ *is treated as an unknown input, while in this scheme only* x_{11} *is treated as an unknown input. Recall the condition* $q \le p \le n$*; since fewer states are treated as unknown inputs, the condition is less likely to be violated, and so the scheme can work even with fewer output measurements. Therefore the scheme presented in this subsection is applicable to a wider class of systems compared to the scheme in §3.3.1.* ♯

3.3.2.2 The state and fault estimation scheme

Since $\operatorname{rank}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = \bar{n}$, there exist two matrices $R \in \mathbb{R}^{\bar{n}\times\bar{n}}$ (which is non-singular) and $V \in \mathbb{R}^{\bar{n}\times p}$ which can be chosen such that

$$R\bar{E} + V\bar{C} = I_{\bar{n}}. \tag{3.104}$$

Pre-multiply the state equation in (3.101) with R and add $V\dot{y}$ to both sides. Using the output equation in (3.101), the following can be obtained:

$$R\bar{E}\dot{\bar{x}} + V\dot{y} = \left(R\bar{E} + V\bar{C}\right)\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{M}\bar{f} + V\dot{y}. \tag{3.105}$$

The SMO by Yeu et al. [157] for system (3.101) has the following structure:

$$\dot{z} = \left(R\bar{A} - G_l\bar{C}\right)z - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu, \tag{3.106}$$

$$\hat{\bar{x}} = Vy - z, \tag{3.107}$$

$$\nu = -\rho\frac{e_y}{\|e_y\|}, \tag{3.108}$$

where $G_l, G_n \in \mathbb{R}^{\bar{n}\times p}$, $e_y = \bar{C}\hat{\bar{x}} - y$, and $\rho \in \mathbb{R}^{+}$. Substituting for z from (3.107) into (3.106) yields

$$\dot{\hat{\bar{x}}} = R\bar{A}\hat{x} - G_l e_y + G_n\nu + V\dot{y}. \tag{3.109}$$

Define the state estimation error $e = \hat{x} - x$. Deducting (3.105) from (3.109) yields the following error system (which characterises the performance of the observer):

$$\dot{e} = \left(R\bar{A} - G_l\bar{C}\right)e - R\bar{M}\bar{f} + G_n\nu. \tag{3.110}$$

Proposition 3.7 *If G_l and G_n are designed appropriately, and if ρ is chosen to satisfy*

$$\rho > \|\left(\bar{C}G_n\right)^{-1}\bar{C}R\bar{M}\|\left(\alpha_x + \alpha_f\right), \tag{3.111}$$

where $\alpha_x, \alpha_f \in \mathbb{R}^{+}$ are chosen to be greater than the upper bounds of x and f respectively, then sliding motion $(e_y, \dot{e}_y = 0)$ is attained in finite time, and x and f can be estimated by the observer. ♯

Proof Error equation (3.110) has the same form as the error equation for the Edwards-Spurgeon SMO for a system represented by the triple $\left(R\bar{A}, R\bar{M}, \bar{C}\right)$ [8]. Additionally, (3.111) implies

$$\rho > \| \left(\bar{C}G_n\right)^{-1} \bar{C}R\bar{M} \| \| \bar{f} \|. \tag{3.112}$$

Therefore, the observer by Yeu et al. in (3.106)–(3.108) can be designed in the same way as the Edwards-Spurgeon SMO (for example, using the methods presented in [44, 47]. Sliding motion $(e_y, \dot{e}_y = 0)$ can then be induced in finite time, and $\hat{x}_2 \to x_2$. Furthermore, f and x_{11} can be estimated $\left(\text{i.e.} \begin{bmatrix} \hat{f} \\ \hat{x}_{11} \end{bmatrix} \to \begin{bmatrix} f \\ x_{11} \end{bmatrix}\right)$ from the following measurable signal:

$$\begin{bmatrix} \hat{f} \\ \hat{x}_{11} \end{bmatrix} = \underbrace{\left(\bar{C}R\bar{M}\right)^{\dagger} \left(\bar{C}G_n\right)}_{G_f} \nu. \tag{3.113}$$

Next, define

$$\hat{x}_{12} = -A_7 \hat{x}_2 - \hat{A}_8 \hat{y} - M_2 \hat{f}. \tag{3.114}$$

After the estimates for x_2, $\hat{y}$, and $\hat{f}$ are obtained, it follows from (3.86)–(3.88) and (3.100) that $\hat{x}_{12} \to x_{12}$. Thus x and f are estimated, completing the proof. ■

Remark 3.5 *The method to design G_l and G_n in Proposition 3.7 require x to be bounded, and the bound to be known. These bounds can be obtained through various ways: such as (but not limited to) from experience with or knowledge about the physical operation of the system, or by simulating the system operating under practical conditions. It is also reasonable to assume that x is bounded as a controller would be used to stabilise the system - the design of the controller is however not considered as does not have any effect on the fault estimation since u does not appear in the error equation from* (3.110) *onwards. Note that the exact value of the bounds do not need to be known; a conservative estimate could be used instead, since sliding motion is guaranteed to occur as long as ρ satisfies* (3.111). ♯

Remark 3.6 *Note that the requirement for x to be bounded reduces the applicability of the scheme since it does not cater to cases where x is unbounded. This is however not a stringent restriction: practical systems operate around a certain operating point, and by* *finding reasonable bounds on the states of the system as per Remark 3.5, the scheme can still be applied across a wide range of systems.* ♯

3.3.2.3 Existence conditions

Yu and Liu [160] investigated the necessary and sufficient conditions for Proposition 3.7 to be satisfied, which are given as follows:

C1. $\text{rank}\begin{bmatrix} \bar{E} & \bar{M} \\ \bar{C} & 0 \end{bmatrix} = \text{rank}\left(\bar{M}\right) + \bar{n},$

C2. $\text{rank}\left(\bar{M}\right) = n - \bar{n} - j + q,$

C3. $\text{rank}\begin{bmatrix} s\bar{E} - \bar{A} & \bar{M} \\ \bar{C} & 0 \end{bmatrix} = \text{rank}\left(\bar{M}\right) + \bar{n} \ \forall \ s \in \mathbb{C}^{+}.$

Conditions C1–C3 however depend on the transformation T_o, which is used to transform the system in the coordinates of (3.95). Notice that after the structures in (3.95) have been obtained, the top $n - k$ states undergo the transformation

$$\Phi = \begin{bmatrix} X_5 & 0 \\ 0 & I_{\bar{n}-k} \end{bmatrix} \underbrace{\begin{bmatrix} T_{o1} \\ T_{o2} \end{bmatrix}}_{T_o} \equiv \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ T_{o2} \end{bmatrix} \begin{array}{l} \updownarrow \ n - \bar{n} - j \\ \updownarrow \ j \\ \updownarrow \ \bar{n} - k \end{array} . \tag{3.115}$$

Hence the following partitions of A and M in (3.87)–(3.88) are obtained:

$$\begin{gathered} \Phi A_{b1} = 0, \ \Phi A_{b1} X_6 = \begin{bmatrix} 0 & I_j \end{bmatrix}, \ T_{o2} A_{b1} X_6 = \begin{bmatrix} A_9 & A_{10} \end{bmatrix}, \\ M_2 = \Phi_2 M_{b1}, \ M_3 = T_{o2} M_{b1}, \ M_4 = M_{b2}. \end{gathered} \tag{3.116}$$

Theorem 3.2 *The transformation Φ exists to satisfy C1–C2 if and only if the following conditions hold:*

D1. $\text{rank}\begin{bmatrix} E & A & M \\ C & 0 & 0 \\ 0 & E & 0 \\ 0 & C & 0 \end{bmatrix} = n + \bar{n} + q,$

D2. $\text{rank}\begin{bmatrix} E & A & M \\ 0 & E & 0 \\ 0 & C & 0 \end{bmatrix} - \text{rank}\begin{bmatrix} E & A \\ 0 & E \\ 0 & C \end{bmatrix} \leq \bar{n} - k.$ ♯

Proof Substituting the structures in reduced-order system (3.101)–(3.103) into C1 yields

$$\operatorname{rank}\begin{bmatrix} A_9 & M_3 - A_{10}M_2 \\ A_{13} & M_4 - A_{14}M_2 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A_9 & M_3 - A_{10}M_2 \\ A_{13} & M_4 - A_{14}M_2 \\ A_{17} & M_5 - A_{18}M_2 \end{bmatrix}. \tag{3.117}$$

The left-hand side (LHS) of (3.117) can be re-expressed as follows:

$$\begin{bmatrix} A_9 & M_3 - A_{10}M_2 \\ A_{13} & M_4 - A_{14}M_2 \end{bmatrix} = \begin{bmatrix} A_9 & M_3 \\ A_{13} & M_4 \end{bmatrix} - \begin{bmatrix} A_{10} \\ A_{14} \end{bmatrix}\begin{bmatrix} 0 & M_2 \end{bmatrix}. \tag{3.118}$$

Therefore, by using (3.118) and the Schur Complement on the LHS of (3.117) the following is obtained:

$$\begin{aligned} \operatorname{rank}\begin{bmatrix} A_9 & M_3 - A_{10}M_2 \\ A_{13} & M_4 - A_{14}M_2 \end{bmatrix} &= \operatorname{rank}\left(\begin{bmatrix} A_9 & M_3 \\ A_{13} & M_4 \end{bmatrix} - \begin{bmatrix} A_{10} \\ A_{14} \end{bmatrix}\begin{bmatrix} 0 & M_2 \end{bmatrix}\right) \\ &= \operatorname{rank}\begin{bmatrix} A_9 & M_3 & A_{10} \\ A_{13} & M_4 & A_{14} \\ 0 & M_2 & I_j \end{bmatrix} - j. \end{aligned} \tag{3.119}$$

Repeating this procedure for the right-hand side (RHS) of (3.117) (which also represents $\operatorname{rank}(\bar{M})$) yields

$$\operatorname{rank}\begin{bmatrix} A_9 & M_3 - A_{10}M_2 \\ A_{13} & M_4 - A_{14}M_2 \\ A_{17} & M_5 - A_{18}M_2 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A_9 & M_3 & A_{10} \\ A_{13} & M_4 & A_{14} \\ A_{17} & M_5 & A_{18} \\ 0 & M_2 & I_j \end{bmatrix} - j. \tag{3.120}$$

Substituting the RHS of (3.119)–(3.120) into (3.117) and then rearranging the rows and columns yields

$$\begin{aligned} &\operatorname{rank}\begin{bmatrix} A_9 & M_3 & A_{10} \\ A_{13} & M_4 & A_{14} \\ 0 & M_2 & I_j \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A_9 & M_3 & A_{10} \\ A_{13} & M_4 & A_{14} \\ A_{17} & M_5 & A_{18} \\ 0 & M_2 & I_j \end{bmatrix} \\ \Rightarrow\ &\operatorname{rank}\left[\begin{array}{cc|c} 0 & I_j & M_2 \\ \hline A_9 & A_{10} & M_3 \\ A_{13} & A_{14} & M_4 \end{array}\right] = \operatorname{rank}\left[\begin{array}{cc|c} 0 & I_j & M_2 \\ \hline A_9 & A_{10} & M_3 \\ A_{13} & A_{14} & M_4 \\ \hline A_{17} & A_{18} & M_5 \end{array}\right]. \end{aligned} \tag{3.121}$$

Further partition $\mathcal{A}_1$ from (3.87) as follows:

$$\mathcal{A}_1 = \begin{array}{c} \begin{bmatrix} 0 \\ \mathcal{A}_{12} \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \;\; n - \bar{n} - j \\ \updownarrow \;\; j \end{array} . \tag{3.122}$$

Hence by comparing (3.120) with the structures from (3.87)–(3.88) and (3.122), it can be seen that (3.120) (and hence C1) is equivalent to

$$\operatorname{rank} \underbrace{\begin{bmatrix} \mathcal{A}_{12} & M_2 \\ \mathcal{A}_4 & \mathcal{M}_2 \end{bmatrix}}_{\Xi_1} = \operatorname{rank} \underbrace{\begin{bmatrix} \mathcal{A}_{12} & M_2 \\ \mathcal{A}_4 & \mathcal{M}_2 \\ \mathcal{A}_7 & \mathcal{M}_3 \end{bmatrix}}_{\Xi_2} . \tag{3.123}$$

Substituting the structures from reduced-order system (3.101)–(3.103) and the expression for $\operatorname{rank}\left(\bar{M}\right)$ from (3.120) into C3 results in

$$\operatorname{rank} \begin{bmatrix} \mathcal{A}_{12} & M_2 \\ \mathcal{A}_4 & \mathcal{M}_2 \\ \mathcal{A}_7 & \mathcal{M}_3 \end{bmatrix} = n - \bar{n} + q \Rightarrow \;\operatorname{rank}\left(\Xi_2\right) = n - \bar{n} + q. \tag{3.124}$$

Therefore, combining (3.123)–(3.124) yields the following, which is equivalent to C1 and C3:

$$\operatorname{rank}\left(\Xi_1\right) = n - \bar{n} + q. \tag{3.125}$$

Then by substituting the structures from (3.87)–(3.88) and (3.115)–(3.116), Ξ_1 can be expressed as follows:

$$\Xi_1 = \begin{bmatrix} \Phi_2 & 0 \\ T_{o2} & 0 \\ 0 & I_{k+p-\bar{n}} \end{bmatrix} \underbrace{\begin{bmatrix} A_{b1} & M_{b1} \\ A_{b4} & M_{b2} \end{bmatrix}}_{\Omega} \begin{bmatrix} X_6 & 0 \\ 0 & I_q \end{bmatrix} . \tag{3.126}$$

Furthermore, substituting the structures from (3.86)–(3.88) into D1 results in

$$\operatorname{rank} \underbrace{\begin{bmatrix} 0 & 0 & M_1 \\ 0 & I_j & M_2 \\ A_9 & A_{10} & M_3 \\ A_{13} & A_{14} & M_4 \end{bmatrix}}_{\Xi_3} = n - \bar{n} + q. \tag{3.127}$$

From the non-singular transformations $T_d,\ T_e$, and T_f defined in (3.97), Ξ_3 can be further expanded as

$$\begin{aligned}\operatorname{rank}(\Xi_3) &= \operatorname{rank}\begin{bmatrix} X_5 & 0 \\ 0 & I_{\bar{n}-p}\end{bmatrix}\begin{bmatrix} T_o & 0 \\ 0 & I_{k-p}\end{bmatrix}\begin{bmatrix} A_{b1} & M_{b1} \\ A_{b4} & M_{b2}\end{bmatrix}\begin{bmatrix} X_6 & 0 \\ 0 & I_q\end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} A_{b1} & M_{b1} \\ A_{b4} & M_{b2}\end{bmatrix}. \end{aligned} \tag{3.128}$$

Equations (3.126)–(3.128) imply $\operatorname{rank}(\Omega) = \operatorname{rank}(\Xi_3) = n - \bar{n} + q$, i.e. D1 is equivalent to $\operatorname{rank}(\Omega) = n - \bar{n} + q$. To show the necessity of D1, suppose it is not satisfied, i.e. $\operatorname{rank}(\Omega) < n - \bar{n} + q$. From (3.126) this would imply $\operatorname{rank}(\Xi_1) < n - \bar{n} + q$. This in turn implies (from (3.125)) C1 and C3 are not satisfied, thus showing the necessity of D1.

Next, partition Ω as follows: let $\Omega_1 = \begin{bmatrix} A_{b1} & M_{b1}\end{bmatrix}$ and $\Omega_2 = \begin{bmatrix} A_{b4} & M_{b2}\end{bmatrix}$. The following transformations are performed on the system in the coordinates of (3.95) to facilitate the calculation of $\Phi_1,\ \Phi_2$, and T_{o2}. Let $\operatorname{rank}(A_{b1}) = \phi$, and apply SVD on it as follows:

$$W_1 A_{b1} W_2 = \begin{bmatrix} 0 & 0 \\ 0 & \mathfrak{A}\end{bmatrix}, \tag{3.129}$$

where $\mathfrak{A} \in \mathbb{R}^{\phi\times\phi}$ is non-singular. It can then be shown that

$$W_1\begin{bmatrix} A_{b1} & M_{b1}\end{bmatrix}\begin{bmatrix} W_2 & 0 \\ 0 & I_q\end{bmatrix} = \left[\begin{array}{cc|c} 0 & 0 & \tilde{M}_1 \\ 0 & \mathfrak{A} & \tilde{M}_2\end{array}\right]. \tag{3.130}$$

Let $\operatorname{rank}(\tilde{M}_1) = r,\ \operatorname{rank}(\tilde{M}_2) = v$, and further decompose $\tilde{M}_1, \tilde{M}_2$ using the non-singular matrices $W_3 \in \mathbb{R}^{(n-k-\phi)\times(n-k-\phi)},\ W_4 \in \mathbb{R}^{\phi\times\phi}$, and $W_5 \in \mathbb{R}^{q\times q}$ as follows:

$$\begin{bmatrix} W_3 & 0 \\ 0 & W_4\end{bmatrix}\begin{bmatrix} \tilde{M}_1 \\ \tilde{M}_2\end{bmatrix} W_5 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & I_r & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & I_v\end{array}\right]. \tag{3.131}$$

Then define the invertible matrices

$$\tilde{T}_a = \begin{bmatrix} W_3 & 0 \\ 0 & W_4\end{bmatrix} W_1,\ \tilde{T}_b = \begin{bmatrix} W_2 & 0 \\ 0 & I_q\end{bmatrix}\left[\begin{array}{cc|c} I_{n-\bar{n}-\phi} & 0 & 0 \\ 0 & W_6 & 0 \\ \hline 0 & 0 & W_5\end{array}\right], \tag{3.132}$$

where $W_6 = (W_3\mathfrak{A})^{-1}$, and it can be shown that

$$\Omega_1 \mapsto \tilde{T}_a\Omega_1\tilde{T}_b = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_r & 0 \\ 0 & I_{\phi-v} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_v & 0 & 0 & I_v \end{array}\right]. \tag{3.133}$$

Suppose D1 holds (since it is necessary); this implies that Ω is full rank. Then due to the structure of Ω_1 in (3.133), there exists an invertible matrix $\tilde{T}_c = \begin{bmatrix} I_{n-k} & 0 \\ \tilde{T}_{c3} & \tilde{T}_{c4} \end{bmatrix}$ that when pre-multiplied with Ω results in

$$\left(\tilde{T}_{c3}\Omega_1 + \tilde{T}_{c4}\Omega_2\right)\tilde{T}_b = \left[\begin{array}{ccc|ccc} I_{n-\bar{n}-\phi} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{A}_{42} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{q-r-v} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{c} \\ \\ \\ \updownarrow\ v, \end{array} \tag{3.134}$$

where $v = k + p - n + \phi - q + r$, and $\tilde{A}_{42} \in \mathbb{R}^{n\times n}$ is invertible. Next, substitute the structures in (3.86)–(3.88) into D2 to obtain

$$\text{rank}\begin{bmatrix} 0 & 0 & M_1 \\ 0 & I_j & M_2 \\ A_9 & A_{10} & M_3 \end{bmatrix} - \text{rank}\begin{bmatrix} 0 & 0 \\ 0 & I_j \\ A_9 & A_{10} \end{bmatrix} \leq \bar{n} - k. \tag{3.135}$$

Then recall $\Omega_1 = \begin{bmatrix} A_{b1} & M_{b1} \end{bmatrix}$. From (3.128) it can be seen that

$$\text{rank}\left(\Omega_1\right) - \text{rank}\left(A_{b1}\right) \leq \bar{n} - k. \tag{3.136}$$

Using (3.133), (3.136) can be further simplified into

$$r \leq \bar{n} - k. \tag{3.137}$$

Choose $j = \phi$. Then an appropriate choice for Φ and T_{o2} to satisfy (3.116) would be

$$\left[\begin{array}{c} \Phi_2 \\ \hline T_{o2} \end{array}\right] = \left[\begin{array}{ccc} \bar{T}_{11} & \bar{T}_{12} & I_j \\ \hline \bar{T}_{21} & \bar{T}_{22} & \bar{T}_{23} \end{array}\right] = \left[\begin{array}{c|c|cc} \bar{T}_{111} & \bar{T}_{121} & I_{j-v} & 0 \\ \bar{T}_{112} & \bar{T}_{122} & 0 & I_v \\ \hline \bar{T}_{21} & \bar{T}_{22} & \bar{T}_{231} & \bar{T}_{232} \end{array}\right], \tag{3.138}$$

where $\bar{T}_{11} \in \mathbb{R}^{j \times (n-k-r-v)}$, $T_{22} \in \mathbb{R}^{(\bar{n}-k) \times r}$. Thus by using (3.133)–(3.134) and (3.138), Ξ_1 in (3.126) can be expanded into

$$\Xi_1 = \begin{bmatrix} 0 & I_{j-v} & 0 & 0 & \bar{T}_{121} & 0 \\ 0 & 0 & I_v & 0 & T_{122} & I_v \\ 0 & T_{231} & T_{232} & 0 & T_{22} & T_{232} \\ I_{n-\bar{n}-j} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{A}_{42} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{q-r-v} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.139}$$

It can be seen from (3.139) that Ξ_1 loses rank if and only if Ξ_{11} loses rank, where

$$\Xi_{11} = \begin{bmatrix} I_{j-v} & \bar{T}_{121} & 0 \\ 0 & \bar{T}_{122} & I_v \\ \bar{T}_{231} & \bar{T}_{22} & \bar{T}_{232} \end{bmatrix}. \tag{3.140}$$

Recall that $T_{22} \in \mathbb{R}^{(\bar{n}-k) \times r}$. Hence, for Ξ_1 to be full rank (which implies from (3.126) that C1 and C3 hold), Ξ_{11} (and therefore $\bar{T}_{22}$) would also need to be full rank. This implies that $r \leq \bar{n} - k$. Recall also from (3.137) that D2 implies $r \leq \bar{n} - k$. Thus, it is shown that D2 is necessary for C1 and C3 to hold. Next, let $\bar{T}_{231}, \bar{T}_{232} = 0$. It can be seen that Ξ_{11} (and therefore Ξ_1) is full-rank, therefore proving the sufficiency of D2. Thus the proof for Theorem 3.2 is complete. ∎

Theorem 3.3 *To satisfy C2, the following condition is necessary:*

D3. (E, A, M, C) *is minimum-phase.*

If $\bar{n} \geq k + r$*, then D3 is also sufficient; otherwise (for* $\bar{n} = k + r$*) a sufficient condition is*

D4. (A_x, A_{b52}) *is detectable, where*

$$A_x = A_{b8} - \mathcal{W}_3 \mathcal{W}_2^{-1} \mathcal{W}_1, \tag{3.141}$$

and the components of A_x *will be defined in* (3.144). ♯

Proof Using (3.133), partition

$$\begin{bmatrix} A_{b1} & A_{b2} & M_{b1} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 0 & A_{b21} & 0 \\ 0 & A_{b22} & M_{b12} \\ A_{b13} & A_{b23} & M_{b13} \end{bmatrix} \end{array} \begin{array}{l} \\ \updownarrow\ r \\ \updownarrow\ j \end{array} . \tag{3.142}$$

Next, using (3.134), partition

$$\begin{bmatrix} A_{b5} & A_{b4} & M_{b2} \end{bmatrix} = \begin{bmatrix} A_{b51} & A_{b41} & M_{b21} \\ A_{b52} & 0 & 0 \end{bmatrix} \begin{array}{l} \\ \updownarrow\ v \end{array} . \tag{3.143}$$

Next, introduce

$$\mathcal{W}_1 = \begin{bmatrix} A_{b22} \\ A_{b23} \\ A_{b51} \end{bmatrix}, \ \mathcal{W}_2 = \begin{bmatrix} 0 & M_{b12} \\ A_{b13} & M_{b13} \\ A_{b41} & M_{b21} \end{bmatrix}, \ \mathcal{W}_3 = \begin{bmatrix} A_{b7} & M_{b3} \end{bmatrix}. \tag{3.144}$$

Denote $R(E, A, M, C) = \begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix}$ as the Rosenbrock matrix of (E, A, M, C), and any zero of (E, A, M, C) will cause it to lose rank. It is clear that the invariant zeros of (E, A, M, C) are equivalent to the invariant zeros of (E_b, A_b, M_b, C_b). Hence, the invariant zeros of (E, A, M, C) are the values of s that make the following matrix lose rank:

$$R_b(s) = \left[\begin{array}{ccc|c} -A_{b1} & -A_{b2} & -A_{b3} & M_{b1} \\ -A_{b4} & -A_{b5} & -A_{b6} & M_{b2} \\ -A_{b7} & sI_{\bar{n}-p} - A_{b8} & -A_{b9} & M_{b3} \\ \hline 0 & 0 & C_3 & 0 \end{array}\right]. \tag{3.145}$$

Using the partitions in (3.142)–(3.143), it is clear that $R_b(s)$ loses rank if the following matrix loses rank:

$$R_{b1}(s) = \begin{bmatrix} 0 & -A_{b21} & 0 \\ 0 & -A_{b22} & M_{b21} \\ -A_{b13} & -A_{b23} & M_{b13} \\ -A_{b41} & -A_{b51} & M_{b21} \\ 0 & -A_{b52} & 0 \\ -A_{b7} & sI_{\bar{n}-p} - A_{b8} & M_{b3} \end{bmatrix} \tag{3.146}$$

Equation (3.146) can be further simplified as follows:

$$R_{b1(s)} = \left[\begin{array}{c|cc} -A_{b21} & 0 & 0 \\ \hline -A_{b22} & 0 & M_{b12} \\ -A_{b23} & A_{b13} & M_{b13} \\ -A_{b51} & A_{b41} & M_{b21} \\ \hline -A_{b52} & 0 & 0 \\ \hline sI_{\bar{n}-p} - A_{b8} & A_{b7} & M_{b3} \end{array}\right] \begin{bmatrix} 0 & -I_{n-\bar{n}} & 0 \\ I_{\bar{n}-p} & 0 & 0 \\ 0 & 0 & I_p \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} -A_{b21} & 0 \\ -\mathcal{W}_1 & \mathcal{W}_2 \\ -A_{b52} & 0 \\ sI_{\bar{n}-p} - A_{b8} & \mathcal{W}_3 \end{bmatrix}}_{R_{b2}(s)}. \tag{3.147}$$

Then by using the structure of the reduced-order system (3.101)–(3.103), the Rosenbrock matrix $R(\bar{E}, \bar{A}, \bar{M}, \bar{C})$ is given by

$$\bar{R}(s) = \begin{bmatrix} -A_{11} + A_{10}A_7 & -A_{12} + A_{10}A_8 & A_9 & M_3 - A_{10}M_2 \\ -A_{15} + A_{14}A_7 & s\mathcal{E}_{22} - A_{16} + A_{14}A_8 & A_{13} & M_4 - A_{14}M_2 \\ sI_{\bar{n}-p} - A_{19} + A_{18}A_7 & -A_{20} + A_{18}A_8 & A_{17} & M_5 - A_{18}M_2 \\ 0 & C_3 & 0 & 0 \end{bmatrix}. \tag{3.148}$$

It can be seen that $\bar{R}(s)$ loses rank if the following matrix loses rank:

$$\bar{R}_1(s) = \begin{bmatrix} -A_{11} + A_{10}A_7 & A_9 & M_3 - A_{10}M_2 \\ -A_{15} + A_{14}A_7 & A_{13} & M_4 - A_{14}M_2 \\ sI_{\bar{n}-p} - A_{19} + A_{18}A_7 & A_{17} & M_5 - A_{18}M_2 \end{bmatrix}$$
$$= \begin{bmatrix} -A_{11} & A_9 & M_3 \\ -A_{15} & A_{13} & M_4 \\ sI_{\bar{n}-p} - A_{19} & A_{17} & M_5 \end{bmatrix} - \begin{bmatrix} A_{10} \\ A_{14} \\ A_{18} \end{bmatrix} \begin{bmatrix} -A_7 & 0 & M_2 \end{bmatrix}. \tag{3.149}$$

By using the Schur complement on (3.149), $\bar{R}_1(s)$ loses rank if and only if the following matrix loses rank:

$$\bar{R}_2(s) = \left[\begin{array}{ccc|c} -A_{11} & A_9 & M_3 & A_{10} \\ -A_{15} & A_{13} & M_4 & A_{14} \\ sI_{\bar{n}-p} - A_{19} & A_{17} & M_5 & A_{18} \\ \hline -A_7 & 0 & M_2 & I_j \end{array}\right]$$

$$= \begin{bmatrix} 0 & I_{\bar{n}} \\ I_j & 0 \end{bmatrix} \underbrace{\begin{bmatrix} -A_7 & I_j & 0 & M_2 \\ -A_{11} & A_{10} & A_9 & M_3 \\ -A_{15} & A_{14} & A_{13} & M_4 \\ sI_{\bar{n}-p} - A_{19} & A_{18} & A_{17} & M_5 \end{bmatrix}}_{\bar{R}_3(s)} \begin{bmatrix} I_{\bar{n}-p} & 0 & 0 \\ 0 & 0 & I_j \\ 0 & I_{n-\bar{n}-j+q} & 0 \end{bmatrix}. \quad (3.150)$$

Notice that

$$\bar{R}_3(s) = \begin{bmatrix} \Phi_2 & 0 \\ T_{o2} & 0 \\ 0 & I_k \end{bmatrix} R_{b2}(s), \quad (3.151)$$

which shows that $\operatorname{rank}\left(R(\bar{E}, \bar{A}, \bar{M}, \bar{C})\right) \leq \operatorname{rank}\left(R(E, A, M, C)\right)$, thus proving the necessity of D3. Next, set $\begin{bmatrix} \Phi_2 \\ T_{o2} \end{bmatrix}$ to have the following structure:

$$\left[\begin{array}{c} \Phi_2 \\ \hline T_{o2} \end{array}\right] = \left[\begin{array}{ccc} 0 & 0 & I_j \\ \hline \bar{T}_{21} & 0 & 0 \\ 0 & \bar{T}_{25} & 0 \end{array}\right], \quad (3.152)$$

where $\bar{T}_{21}$ and $\begin{bmatrix} \Phi_2 \\ T_{o2} \end{bmatrix}$ are chosen such that they are both full-row rank. Substituting R_{b2} from (3.147) and $\begin{bmatrix} \Phi_2 \\ T_{o2} \end{bmatrix}$ from (3.152) into (3.151) yields

$$\bar{R}_3(s) = \begin{bmatrix} 0 & 0 & I_j & 0 \\ I_{\bar{n}-r-k} & 0 & 0 & 0 \\ 0 & \bar{T}_{25} & 0 & 0 \\ 0 & 0 & 0 & I_k \end{bmatrix} \begin{bmatrix} -\bar{T}_{21} A_{b12} & 0 \\ -\mathcal{W}_1 & \mathcal{W}_2 \\ -A_{b52} & 0 \\ sI_{\bar{n}-p} - A_{b8} & \mathcal{W}_3 \end{bmatrix}. \quad (3.153)$$

Recall that $\operatorname{rank}(\Omega) = n - \bar{n} + q$, and that $\Omega = \begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix}$. From the structures in (3.133)–(3.134) and the partitions in (3.142)–(3.143), it can be seen that $\mathcal{W}_2 \in \mathbb{R}^{(n-\bar{n}+q)\times(n-\bar{n}+q)}$ is non-singular. Therefore, define

$$\mathcal{I} = \begin{bmatrix} I_{n-\bar{n}+q} & 0 & 0 \\ 0 & I_{\bar{n}-r-k} & 0 \\ \mathcal{W}_3\mathcal{W}_2^{-1} & 0 & I_{\bar{n}-p} \end{bmatrix}. \tag{3.154}$$

It can then be shown that

$$\bar{R}_3(s) = \begin{bmatrix} I_{\bar{n}-r-k} & 0 \\ 0 & \mathcal{I} \end{bmatrix} \underbrace{\begin{bmatrix} -\bar{T}_{21}A_{b12} & 0 \\ -\mathcal{W}_1 & \mathcal{W}_2 \\ -A_{b52} & 0 \\ sI_{\bar{n}-p} - A_x & 0 \end{bmatrix}}_{\bar{R}_4(s)}. \tag{3.155}$$

Since $\mathcal{W}_2$ is non-singular, $\bar{R}_4(s)$ loses rank if and only if the following matrix loses rank:

$$\bar{R}_5(s) = \begin{bmatrix} -\Sigma \\ sI_{\bar{n}-p} - A_x \end{bmatrix}, \ \Sigma = \begin{bmatrix} \bar{T}_{21}A_{b21} \\ A_{b52} \end{bmatrix}, \tag{3.156}$$

where the unobservable modes of (Σ, A_x) are the zeros of $\left(\bar{E}, \bar{A}, \bar{M}, \bar{C}\right)$. Let H be a matrix containing the generalised right-eigenvectors of A_x. Hence $H^{-1}A_xH$ is a matrix in the *Jordan canonical form*, where the diagonal elements are the real parts of the eigenvalues of A_x [98]. Pre-multiply $\bar{R}_5(s)$ with $\begin{bmatrix} I_{\bar{n}-r-k+\upsilon} & 0 \\ 0 & H^{-1} \end{bmatrix}$ and post-multiply with H to obtain

$$\begin{bmatrix} I_{\bar{n}-r-k+\upsilon} & 0 \\ 0 & H^{-1} \end{bmatrix} \begin{bmatrix} -\Sigma \\ sI_{\bar{n}-p} - A_x \end{bmatrix} H = \begin{bmatrix} \bar{T}_{21} & 0 \\ 0 & I_{\upsilon+\bar{n}-p} \end{bmatrix} \underbrace{\left[\begin{array}{c} A_{b21}H \\ \hline A_{b52}H \\ H^{-1}\left(sI_{\bar{n}-p} - A_x\right)H \end{array}\right]}_{\bar{R}_6(s)}. \tag{3.157}$$

A zero of $\left(\bar{E}, \bar{A}, \bar{M}, \bar{C}\right)$ which corresponds to an unobservable mode of $(\Sigma H, H^{-1}A_xH)$ will appear as an element of $H^{-1}A_xH$ where its corresponding column in ΣH is zero. Next, recall from (3.149) that the zeros of (E, A, M, C) are the values of s that make

$R_{b2}(s)$ lose rank. Using the structure of $\mathcal{I}$ in (3.154), it can be shown that

$$R_{b2}(s) = \begin{bmatrix} I_{n-k-r-j} & 0 \\ 0 & \mathcal{I} \end{bmatrix} \underbrace{\begin{bmatrix} -A_{b21} & 0 \\ -\mathcal{W}_1 & \mathcal{W}_2 \\ -A_{b52} & 0 \\ sI_{\bar{n}-p} - A_x & 0 \end{bmatrix}}_{R_{b3}(s)}. \tag{3.158}$$

Since $\mathcal{W}_2$ is non-singular, $R_{b3}(s)$ loses rank if and only if the following matrix loses rank:

$$R_{b4}(s) = \begin{bmatrix} A_{b21}H \\ A_{b52}H \\ H^{-1}\left(sI_{\bar{n}-p} - A_x\right)H \end{bmatrix}, \tag{3.159}$$

which is identical to $\bar{R}_6(s)$ from (3.157). Therefore, if D3 holds, the corresponding columns in $\begin{bmatrix} A_{b21} \\ A_{b52} \end{bmatrix} H$ corresponding to positive diagonal elements (which indicate unstable eigenvalues) within $H^{-1}A_xH$ will be non-zero. Hence if $\bar{n} > r + k$, then $\bar{T}_{21}$ (which constitutes design freedom) exists, and can be chosen such that ΣH has non-zero elements at columns corresponding to the diagonals of $H^{-1}A_xH$ that are unstable, guaranteeing that the unstable modes are observable (and are therefore not zeros of (E, A, M, C)). Thus the sufficiency of D3 for the case when $\bar{n} > r + k$ is proven.

For the case $\bar{n} = r + k,\ A_{b21}$ and $\bar{T}_{21}$ do not exist, and therefore $R_{b2}(s)$ from (3.147) becomes

$$R_{b2}(s) = \mathcal{I} \underbrace{\begin{bmatrix} -\mathcal{W}_1 & \mathcal{W}_2 \\ -A_{b52} & 0 \\ sI_{\bar{n}-p} - A_x & 0 \end{bmatrix}}_{R_{b21}(s)}. \tag{3.160}$$

Since $\mathcal{W}_2$ is non-singular, $R_{b21}(s)$ loses rank if and only if the following matrix loses rank:

$$R_{b22}(s) = \begin{bmatrix} A_{b52} \\ sI_{\bar{n}-p} - A_x \end{bmatrix}. \tag{3.161}$$

By comparing (3.156) and (3.161), it can be seen that $R_{b22}(s)$ is identical to $\bar{R}_5(s)$ (since $\bar{T}_{21}$ and A_{b21} do not exist). Thus, the values of s that make $R_{b22}(s)$ lose rank are the unobservable modes of (A_x, A_{b52}), and the sufficiency of D4 is proven. ■

3.3.2.4 Design procedure

The design procedure for the scheme in this subsection is summarised as follows:

1. Check that D1–D3 (and D4 if necessary) hold for system (3.83)–(3.84). If not, the scheme in this subsection is not applicable.

2. Determine the matrices $T_a,\ T_b$, and T_c from (3.89) and (3.94).

3. Design Φ to have the following structure:

$$\Phi = \begin{bmatrix} T_{x1} & T_{x2} & 0 \\ 0 & 0 & I_j \\ \bar{T}_{21} & 0 & 0 \\ 0 & \bar{T}_{25} & 0 \end{bmatrix}, \tag{3.162}$$

   where $T_{x1} \in \mathbb{R}^{(n-\bar{n}-j)\times(n-k-r-j)}$ and $\bar{T}_{25} \in \mathbb{R}^{r\times r}$ are chosen such that $\begin{bmatrix} T_{x1} \\ \bar{T}_{21} \end{bmatrix}$ and Φ are non-singular.

4. Calculate the matrices $T_d,\ T_e$, and T_f using (3.96)–(3.97).

5. Apply the state equation transformation $T_eT_dT_b$ and the state transformation $T_aT_cT_f$.

6. Derive the reduced-order system (3.101)–(3.103).

7. Calculate the matrices R and V from (3.104).

8. Design the SMO by Yeu et al. [157] (3.106)–(3.108) for the reduced-order system using Proposition 3.7.

9. Calculate G_f from (3.113), and estimate x_2 according to (3.114).

3.4 Simulation example

The concepts covered in the previous section on SMOs for NIODS are illustrated using the following example. Consider a chemical mixing tank [157] described by the following dynamical model:

$$\dot{c}_3 = -0.375c_3 - 0.0667q_3 + 0.1q_1, \tag{3.163}$$

$$0 = -q_3 + q_1, \tag{3.164}$$

$$\dot{c}_5 = 0.3c_3 + 0.0533q_3 - 0.5c_3 - 0.04q_5 + 0.02q_4, \tag{3.165}$$

$$0 = q_3 - q_5 + q_4, \tag{3.166}$$

where q_1 is the flow rate of the influent into the first tank, c_3 and q_3 are the concentration and flow rate of the influent from the first tank into the second tank, respectively, q_4 is the flow rate of influent from another pipe into the second tank, and c_5 and q_5 are the concentration and flow rate of the effluent from the second tank, respectively. In this example, the influents from external sources are treated as faults, i.e.

$$f_a = q_1, \;\; f_b = q_4, \tag{3.167}$$

3.4.1 System formulation

The system matrices (E, A, M) in the framework of (3.83)–(3.84) are

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; A = \begin{bmatrix} -0.375 & -0.0667 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0.3 & 0.0533 & -0.5 & -0.04 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \; M = \begin{bmatrix} 0.1 & 0 \\ 1 & 0 \\ 0 & 0.02 \\ 0 & 1 \end{bmatrix}, \tag{3.168}$$

for the system variables

$$x = \begin{bmatrix} \text{concentration, } c_3 \text{ (mol/l)} \\ \text{flow rate, } q_3 \text{ (l/s)} \\ \text{concentration, } c_5 \text{ (mol/l)} \\ \text{flow rate, } q_5 \text{ (l/s)} \end{bmatrix}, \; f = \begin{bmatrix} f_a \text{ (l/s)} \\ f_b \text{ (l/s)} \end{bmatrix}. \tag{3.169}$$

Assume that only c_3 and q_3 are measurable, and therefore the matrix C is

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.170}$$

It can be seen that $n = 4$ and $\bar{n} = 3$, which (from (3.72)) indicates that system (3.168)–(3.170) is not infinitely observable. Therefore, the observer by Yeu et al. [157] outlined in §3.2 cannot be used for this system. We now design the observer scheme for NIODS presented in §3.3.2.

3.4.2 Observer design

For readability, the steps outlined in §3.3.2.4 will be referred to in the following design of the observer scheme.

Steps 1–2: Conditions D1–D3 can be quickly verified to hold for system (3.168)–(3.170), and so the transformations introduced in Proposition 3.6 exist for system (3.168)–(3.170). Since $\bar{n} = k + r$ however, D4 needs to be verified. Apply the state equation transformation transformation T_cT_b and T_a as follows:

$$T_cT_b = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; T_a = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.171}$$

The value of A_x was found to be

$$A_x = -0.675 < 0, \tag{3.172}$$

thus showing D4 is also satisfied, thereby guaranteeing the existence of the observer scheme proposed in §3.3.2. The proposed observer scheme is now designed.

Steps 3–4: The following partitions were obtained:

$$A_{b1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; \tilde{M}_1 = \begin{bmatrix} 0 & -1 \end{bmatrix}, \tag{3.173}$$

which (from their definitions in (3.96) and (3.131)) imply $j = r = 1$. Thus since $n - \bar{n} - j = n - k - r - j = 0,\ T_{x1},\ T_{x2}$ and $\bar{T}_{21}$ do not exist. Hence a suitable choice for $\Phi,\ T_d,\ T_e,\ T_f$ would be

$$\Phi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; T_d = T_e = T_f = I_4. \tag{3.174}$$

Step 6: Since $n - \bar{n} - j = 0,\ x_{11}$ does not exist. Then using (3.100), x_{12} is re-expressed as

$$x_{12} = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix} \underbrace{\begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\bar{x}} + \begin{bmatrix} 0 & 1 \end{bmatrix} f. \tag{3.175}$$

The reduced-order system (3.101)–(3.103) therefore has the structure

$$\underbrace{\left[\begin{array}{c|cc} 0 & 0 & 0 \\ 0 & -1 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]}_{\bar{E}}\dot{\bar{x}} = \underbrace{\left[\begin{array}{c|cc} 0 & 0 & -1 \\ -0.3 & 0.5 & -0.0133 \\ \hline -0.375 & 0 & -0.0667 \end{array}\right]}_{\bar{A}}\bar{x} + \underbrace{\left[\begin{array}{cc} 1 & 1 \\ 0 & 0.0333 \\ \hline 0.1 & 0.0667 \end{array}\right]}_{\bar{M}} f, \tag{3.176}$$

$$y = \underbrace{\left[\begin{array}{c|cc} 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]}_{\bar{C}}\bar{x}. \tag{3.177}$$

It can be seen that $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n} = 3$, i.e. the reduced-order system (3.176)–(3.177) is *infinitely observable* [49], and hence the SMO by Yeu et al. [157] is applicable onto the system to estimate $\bar{x}$ and f.

Step 7: From the structures of $\bar{E}$ and $\bar{C}$ in (3.176)–(3.177), the matrices R and V are calculated to be

$$R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -0.5 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \; V = \begin{bmatrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.178}$$

Steps 8–9: The poles of the observer and sliding motion were chosen as $\{-0.5, -13, -15\}$ and $\{-0.5\}$, respectively. Using the design method outlined in [44] yields

$$G_l = \begin{bmatrix} 25.25 & 1.395 \\ 13.05 & 0.0216 \\ 0 & 14 \end{bmatrix}. \; G_n = \begin{bmatrix} 2 & 0.1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.179}$$

Furthermore, ρ is set as $\rho = 20$, and G_f and $\hat{x}_{12}$ are calculated as

$$G_f = \begin{bmatrix} 60.06 & 1 \\ -60.06 & 0 \end{bmatrix}, \; \hat{x}_{12} = -\hat{x}_4 + \hat{f}_2. \tag{3.180}$$

3.4.3 Simulation results

The initial condition of the system is set as $x(0) = \{4, 2.866, 3, 4.116\}$, while the initial condition of the observer was set at zero. The fault signals were simulated as

$$f_1 = \sin\left(t + \frac{\pi}{3}\right) + 2, \; f_2 = 0.5\sin\left(1.8t + \frac{5\pi}{6}\right) + 1. \tag{3.181}$$

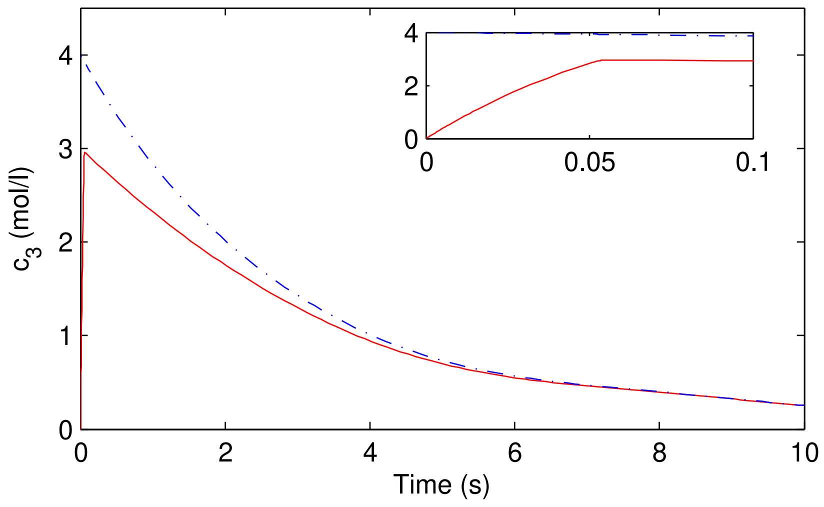

Figure 3.8: The first state (dash-dotted) and its estimate (solid).

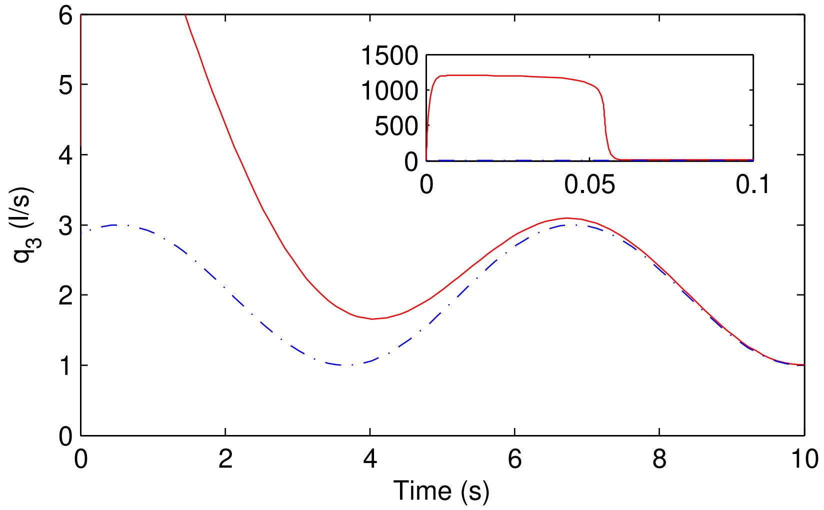

Figure 3.9: The second state (dash-dotted) and its estimate (solid).

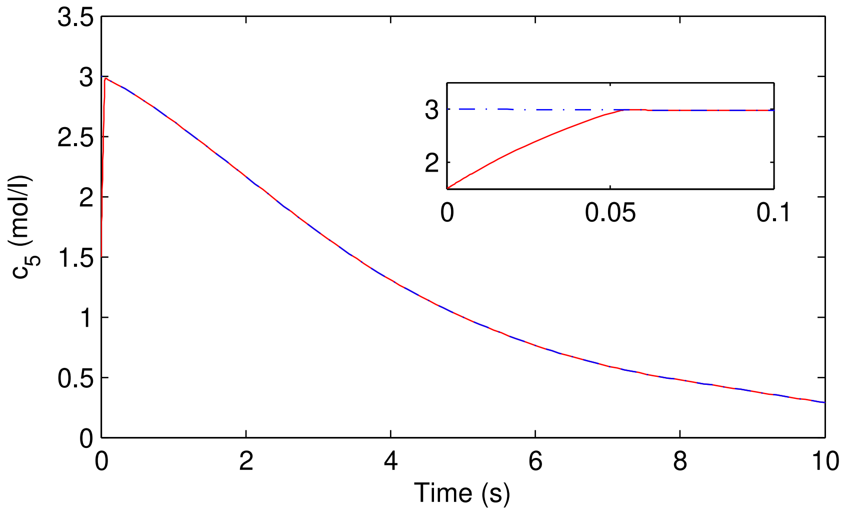

Figure 3.10: The third state (dash-dotted) and its estimate (solid).

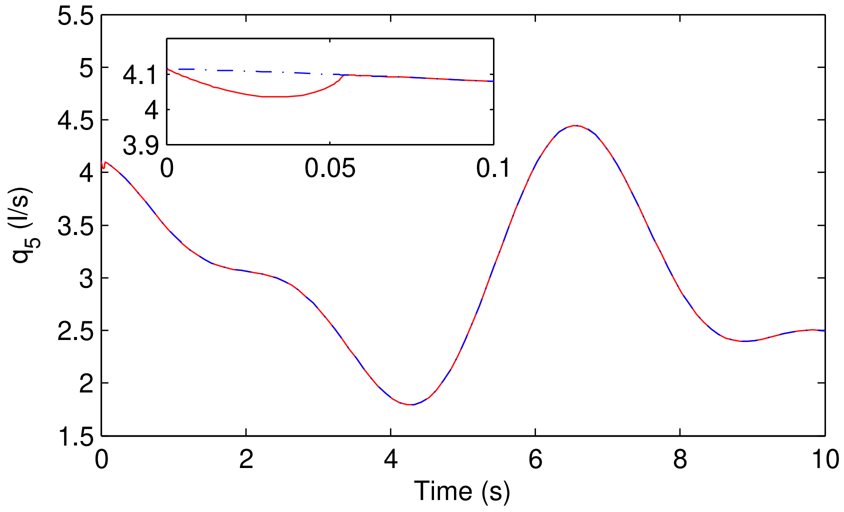

Figure 3.11: The fourth state (dash-dotted) and its estimate (solid).

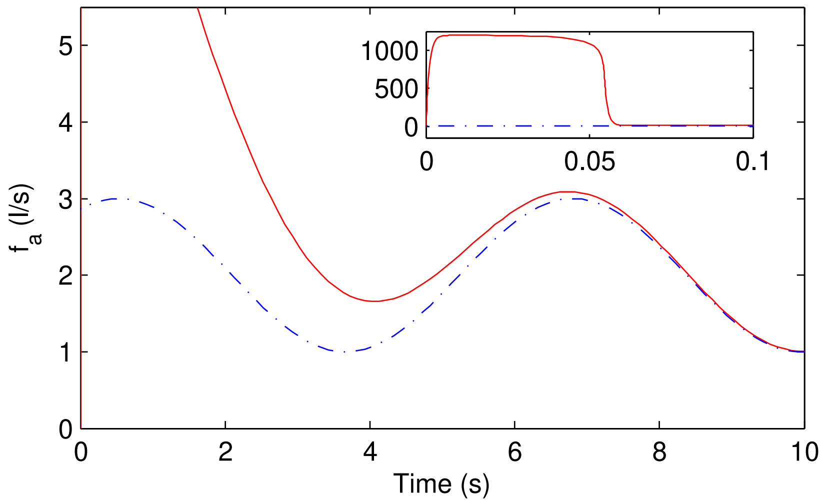

Figure 3.12: The first fault (dash-dotted) and its estimate (solid).

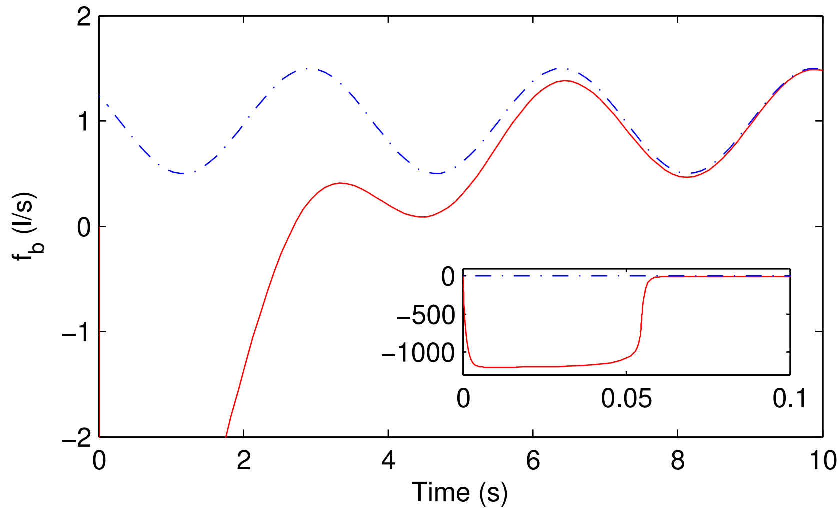

Figure 3.13: The second fault (dash-dotted) and its estimate (solid).

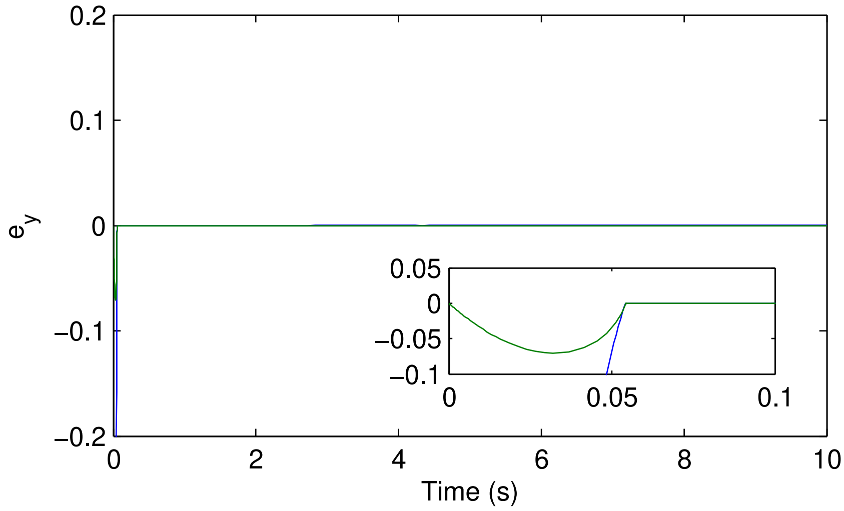

Figure 3.14: Output estimation error $e_y = \hat{y} - y$.

Figures 3.8–3.13 show the evolution of the system states and faults, and their estimates. At about $t = 8$ s, asymptotic tracking of the states takes place - the slow convergence is due to the small pole of the reduced-order error dynamics at $s = -0.675$. The observer estimates track the states and faults faithfully after convergence. Figure 3.14 shows the output estimation error e_y against time. At approximately $t = 0.05$ s, e_y goes to and remains at zero, indicating that *sliding motion* on the surface $\mathcal{S} = \{e : Ce = 0\}$ has taken place (which is reflected in Figures 3.10–3.11, where the estimates of the third and fourth states (which are outputs) converge on the true values in finite time). Notice also from Figures 3.8–3.11 that the state estimates experience an abrupt change in dynamics at the same time.

3.5 Conclusion

This chapter presents existing sliding mode observer (SMO) schemes for descriptor systems in the literature (with some original proofs and remarks for clarity). The sliding mode observer to estimate states and faults for descriptor systems by Yeu et al. [157] was presented in this chapter. The existence conditions for the observer by Yeu et al. were then investigated, and it was found that the system needs to be infinitely observable for the observer to be feasible. Infinite observability thus presents a limitation for SMO schemes in descriptor systems. To overcome this, Ooi et al. presented two techniques to re-express non-infinitely observable descriptor systems as infinitely observable reduced-order descriptor systems which are compatible with the observer by Yeu et al. The first method [100] re-expresses the original system by treating certain states as unknown inputs. The second method [101] also re-expresses certain states in terms of other states to reduce the number of states treated as unknown inputs, thereby increasing the applicability of the system. Finally, a numerical example demonstrates the efficacy of the second method.

Chapter 4

Robust fault reconstruction for NIODS

4.1 Introduction

In the previous chapter (in particular §3.2.2), a necessary condition for the sliding mode observer (SMO) by Yeu et al. [157] to estimate the states and reconstruct the faults of a descriptor system is that the system is infinitely observable [49]. As described in §3.3.1, Ooi et al. [100] later extended these findings such that the observer scheme is also applicable to a class of *non-infinitely observable descriptor systems* (NIODS); their scheme however did not fully exploit the design freedom available, and the robustness of the fault reconstruction against disturbances was not considered as well. This chapter thus aims to extend these results by presenting a SMO scheme for a class of NIODS where the observer parameters are designed using linear matrix inequalities (LMIs) such that the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction is minimised, thus achieving *robust fault reconstruction*.

This chapter begins with preliminary transformations to re-express the descriptor system into a form that facilitates analysis and design. Then, by removing certain states and treating them as unknown inputs, an infinitely observable reduced-order system is formulated. In §4.3, the SMO by Yeu et al. [157] is then applied onto the reduced-order system. The necessary and sufficient conditions for the existence of the observer are also investigated and presented in terms of the original system matrices. The design parameters are then designed using the Bounded Real Lemma [128] such that the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction is minimised. A set of design procedures for the scheme is then outlined. Finally, a numerical example is presented in §4.4 to demonstrate the effectiveness of the proposed scheme.

The work in this chapter has been published; its reference is J. C. L. Chan, C. P. Tan, and H. Trinh, Robust fault reconstruction for a class of infinitely unobservable descriptor systems, *International Journal of Systems Science*, 48(8):1646–1655, 2017.

4.2 Preliminary transformations

Consider the following non-infinitely observable descriptor system (NIODS):

$$E\dot{x} = Ax + Bu + Mf + Q\xi, \tag{4.1}$$

$$y = Cx, \tag{4.2}$$

where $E, A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m},\ M \in \mathbb{R}^{n\times q},\ Q \in \mathbb{R}^{n\times h},\ C \in \mathbb{R}^{p\times n}$. The vectors $x,\ u,\ y,\ f,$ and ξ represent the states, inputs, outputs, faults, and disturbance signals, respectively. Only u and y are measurable. The fault signal f represents an abnormal condition acting upon the system; it is desired to reconstruct f so that information on its shape and magnitude could be obtained such that timely and accurate corrective action can be taken. The unknown disturbance signal ξ (which is not a fault, and could arise from mismatches in modelling or parasitic dynamics [116]) however may cause the fault reconstruction to become erroneous, and raise false positives or even mask the effect of a fault. Thus, it is of interest to reconstruct f while minimising the effect of ξ on its reconstruction. It is assumed generally that E is rank deficient, i.e. $\operatorname{rank}(E) = r < n$, and that M and C are full-column rank and full-row rank, respectively, i.e. $\operatorname{rank}(M) = q,\ \operatorname{rank}(C) = p$. Similar to Lemma 3.1 in §3.2.1, system (4.1)–(4.2) is first re-expressed such that it has a form facilitating further analysis using Lemma 4.1 and Proposition 4.1 as follows:

Lemma 4.1 *Let state equation* (4.1) *be pre-multiplied by* T_1*, state* x_0 *be transformed such that* $x_0 \mapsto T_2x_0$*, and fault* f_0 *be transformed such that* $f_0 \mapsto T_3f_0$*, where* $T_1,\ T_2,$ *and* T_3 *are non-singular matrices (referred to hereafter as a* state equation transformation*, a* state transformation*, and a* fault transformation*, respectively). System* (4.1)–(4.2) *will be transformed to have the form*

$$(T_1ET_2^{-1})(T_2\dot{x}_0) = (T_1AT_2^{-1})(T_2x_0) + (T_1B)u + (T_1MT_3^{-1})(T_3f_0) + (T_1Q)\xi, \tag{4.3}$$

$$y = (CT_2^{-1})(T_2x_0). \tag{4.4}$$

Note that the matrices (E, A, B, M, Q, C) *have been transformed. These transformations are used to change these matrices into forms that are more convenient for analysis and design, but preserve the dynamic properties of the overall system* (4.1)–(4.2). *It can be seen that the inputs and outputs of the system remain unchanged. The transfer function* $G(s)$ *of a descriptor system is given by* $G(s) = C(sE - A)^{-1}B$. *Therefore, if* T_1, T_2, *and* T_3 *are non-singular, the transfer functions of systems* (4.1)–(4.2) *and* (4.3)–(4.4) *are identical. Thus, the dynamic properties such as the inputs, outputs, finite poles, and dynamical order of both systems are the same, and systems* (4.1)–(4.2) *and* (4.3)–(4.4) *are said to be* equivalent. $\sharp$

Proposition 4.1 *There exists a set of transformations for system* (4.1)–(4.2) *such that the system matrices and state vector* x *would have the following structure:*

$$E = \begin{bmatrix} 0 & I_{\bar{n}-p} & E_{21} \\ 0 & 0 & E_{22} \end{bmatrix}, \; A = \begin{array}{c} \begin{bmatrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; \bar{n}-p \\ \updownarrow \; n-\bar{n}+p \end{array},$$
$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \; M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \; Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} \begin{array}{l} \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; p \end{array}, \tag{4.5}$$

where

$$\bar{n} = \operatorname{rank} \begin{bmatrix} E \\ C \end{bmatrix} < n. \tag{4.6}$$

The partitions of E, A, B, M, Q *have the same row dimensions, whilst the column partitions of* E, A, C *are conformable to the partitions of* x. $\sharp$

Proof Introduce a state transformation $T_z = \begin{bmatrix} N_C^T \\ C \end{bmatrix}$ where $N_C C = 0$, i.e. $x \mapsto T_z x$ so that

$$C \mapsto C_z = CT_z^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}, \; E \mapsto E_z = ET_z^{-1} = \begin{bmatrix} E_1 & E_2 \end{bmatrix}, \tag{4.7}$$

where E_z has no particular structure. It can be seen from (4.7) that

$$\operatorname{rank} \begin{bmatrix} E_z \\ C_z \end{bmatrix} = p + \operatorname{rank}(E_1). \tag{4.8}$$

From (4.6) and (4.8), it follows that $\mathrm{rank}(E_1) = \bar{n} - p$. Using singular-value decomposition (SVD), E_1 can be decomposed as

$$E_1 = T_a^T \begin{bmatrix} 0 & I_{\bar{n}-p} \\ 0 & 0 \end{bmatrix} T_{b1}^T, \tag{4.9}$$

where T_a and T_{b1} are orthogonal. Define a state equation transformation T_a and a state transformation T_b^{-1} where $T_b = \begin{bmatrix} T_{b1}^T & 0 \\ 0 & I_p \end{bmatrix}$, and $x \mapsto T_b^{-1}x = x_b$. Apply these transformations onto system (4.1)–(4.2), i.e.

$$\begin{gathered} E \mapsto T_a E T_z^{-1} T_b,\ A \mapsto T_a A T_z^{-1} T_b,\ B \mapsto T_a B,\ M \mapsto T_a M, \\ Q \mapsto T_a Q, C \mapsto C T_z^{-1} T_b. \end{gathered} \tag{4.10}$$

From (4.7)–(4.10), it can be seen that

$$E \mapsto T_a E T_z^{-1} T_b = E_b = \begin{bmatrix} 0 & I_{\bar{n}-p} & E_{21} \\ 0 & 0 & E_{22} \end{bmatrix},\ C \mapsto C T_z^{-1} T_b^{-1} = C_b = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix}, \tag{4.11}$$

It can then be seen that the structures of (E_b, C_b) in (4.11) are identical to the structures of (E, C) in (4.5), thus completing the proof. ■

Define an invertible matrix

$$T = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_p \\ 0 & \bar{T}_p \end{bmatrix}, \tag{4.12}$$

where $T_p \in \mathbb{R}^{p\times(n-\bar{n}+p)}$, $\bar{T}_p \in \mathbb{R}^{(n-\bar{n})\times(n-\bar{n}+p)}$. The matrix T_p represents design freedom to be exploited, while $\bar{T}_p$ is solely to make T full rank. One suitable choice of $\bar{T}_p$ would be the transpose of the right null-space of T_p, i.e. $T_p\bar{T}_p^T = 0$. Apply the state equation transformation T onto the system (that is in the coordinates of (4.5)) to obtain the following:

$$\begin{aligned} \begin{bmatrix} 0 & I_{\bar{n}-p} & E_{21} \\ 0 & 0 & T_pE_{22} \\ 0 & 0 & \bar{T}_pE_{22} \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{y} \end{bmatrix} &= \begin{bmatrix} A_1 & A_2 & A_3 \\ T_pA_4 & T_pA_5 & T_pA_6 \\ \bar{T}_pA_4 & \bar{T}_pA_5 & \bar{T}_pA_6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} + \begin{bmatrix} B_1 \\ T_pB_2 \\ \bar{T}_pB_2 \end{bmatrix} u \\ &\quad + \begin{bmatrix} M_1 \\ T_pM_2 \\ \bar{T}_pM_2 \end{bmatrix} f + \begin{bmatrix} Q_1 \\ T_pQ_2 \\ \bar{T}_pQ_2 \end{bmatrix} \xi, \end{aligned} \tag{4.13}$$

$$y = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} x. \tag{4.14}$$

Next, treat x_1 as an unknown input, i.e. define

$$\bar{x} = \begin{bmatrix} x_2 \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} f \\ x_1 \end{bmatrix}, \tag{4.15}$$

and system (4.13)–(4.14) can then be re-expressed as the following reduced-order system:

$$\underbrace{\begin{bmatrix} I_{\bar{n}-p} & E_{21} \\ 0 & T_pE_{22} \end{bmatrix}}_{\bar{E}} \dot{\bar{x}} = \underbrace{\begin{bmatrix} A_2 & A_3 \\ T_pA_5 & T_pA_6 \end{bmatrix}}_{\bar{A}} \bar{x} + \underbrace{\begin{bmatrix} B_1 \\ T_pB_2 \end{bmatrix}}_{\bar{B}} u + \underbrace{\begin{bmatrix} M_1 & A_1 \\ T_pM_2 & T_pA_4 \end{bmatrix}}_{\bar{M}} \bar{f} + \underbrace{\begin{bmatrix} Q_1 \\ T_pQ_2 \end{bmatrix}}_{\bar{Q}} \xi, \tag{4.16}$$

$$y = \underbrace{\begin{bmatrix} 0 & I_p \end{bmatrix}}_{\bar{C}} \bar{x}. \tag{4.17}$$

Notice that $\operatorname{rank}\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$ (which is full-column rank) and the output of the reduced order system (4.16)–(4.17) is a measurable signal y. Hence the observer by Yeu et al. [157] can now be designed based on $(\bar{E}, \bar{A}, \bar{M}, \bar{C})$ and driven by u and y to estimate $\bar{x}$ and $\bar{f}$, thus estimating x and f. In order to be able to fully reconstruct $\bar{f}$, the top $\bar{n} - p$ rows of $\bar{M}$ $\left(\text{i.e. } \begin{bmatrix} M_1 & A_1 \end{bmatrix}\right)$ need to be made zero, i.e. the following equation must be satisfied:

$$\operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ T_pM_2 & T_pA_4 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix}. \tag{4.18}$$

If (4.18) holds (the satisfaction of (4.18) will be addressed later in Proposition 4.3), then define the following state equation transformation:

$$T_l = \begin{bmatrix} I_{\bar{n}-p} & -X \\ 0 & I_p \end{bmatrix}, \tag{4.19}$$

where $X = \begin{bmatrix} M_1 & A_1 \end{bmatrix} \begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix}^{\dagger}$, and $\begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix}^{\dagger}$ represents the left pseudo-inverse (also known as the Moore-Penrose inverse) of $\begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix}$, where

$$\begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix}^{\dagger} \begin{bmatrix} T_pM_2 & T_pA_4 \end{bmatrix} = \begin{bmatrix} I_{q+r_a} & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.20}$$

where $r_a = \operatorname{rank}(A_4)$. Apply the state equation transformation T_l onto system (4.16)–

(4.17) to obtain following transformed matrices with the structures

$$\begin{aligned} \bar{E} &= \begin{bmatrix} I_{\bar{n}-p} & \bar{E}_{21} \\ 0 & T_p E_{22} \end{bmatrix}, \bar{A} = \begin{bmatrix} \bar{A}_2 & \bar{A}_3 \\ T_p A_5 & T_p A_6 \end{bmatrix}, \bar{B} = \begin{bmatrix} \bar{B}_1 \\ T_p B_2 \end{bmatrix}, \\ \bar{M} &= \begin{bmatrix} 0 & 0 \\ T_p M_2 & T_p A_4 \end{bmatrix}, \bar{Q} = \begin{bmatrix} \bar{Q}_1 \\ T_p Q_2 \end{bmatrix}, \bar{C} = \begin{bmatrix} 0 & I_p \end{bmatrix}, \end{aligned} \tag{4.21}$$

where $\bar{E}_{21} = E_{21} - XT_pE_{22}, \bar{A}_2 = A_2 - XT_pA_5, \bar{A}_3 = A_3 - XT_pA_6, \bar{B}_1 = B_1 - XT_pB_2, \bar{Q}_1 = Q_1 - XT_pQ_2$. The observer by Yeu et al. [157] can now be applied to reconstruct $\bar{f}$, thereby reconstructing f, as will be shown in the next subsection.

4.3 The sliding mode observer for fault reconstruction

The observer by Yeu et al. [157] for system (4.21) is given in compact form as follows:

$$\dot{z} = \left(R\bar{A} - G_l\bar{C}\right) z - R\bar{B}u - \left(G_l(I_p - \bar{C}V) + R\bar{A}V\right) y - G_n\nu, \tag{4.22}$$

$$\hat{\bar{x}} = Vy - z, \tag{4.23}$$

$$\nu = -\rho \frac{e_y}{\|e_y\|},\ e_y = \bar{C}\hat{\bar{x}} - y, \tag{4.24}$$

where $R \in \mathbb{R}^{\bar{n}\times\bar{n}}$ is invertible and $V \in \mathbb{R}^{\bar{n}\times p}$. The following theorem forms the main result of this chapter.

Theorem 4.1 *The observer by Yeu et al. [157] can reconstruct f for system* (4.1)–(4.2) *if and only if the following conditions hold:*

E1. $\operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ M_2 & A_4 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix}$,

E2. $\operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix} = q + r_a$,

E3. $p \geq q + r_a$,

where $r_a = \operatorname{rank}(A_4)$. The following condition is necessary, and also sufficient if $p - q - r_a > 0$:

E4. $\operatorname{rank}\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix} = \bar{n} + q + r_a \ \forall s \in \mathbb{C}^+$.

If $p - q - r_a = 0$*, the necessary condition is E4, and the sufficient condition is that*

E5. the eigenvalues of $\bar{A}_2$ *are stable.* ♯

Proof The remainder of this section presents the constructive proof for Theorem 4.1.

Let the gain matrices G_l and G_n have the following structures

$$G_l = \begin{bmatrix} G_{l1} \\ G_{l2} \end{bmatrix}, \; G_n = \begin{bmatrix} 0 \\ G_{n2} \end{bmatrix} \begin{array}{l} \updownarrow \; \bar{n} - p \\ \updownarrow \; p \end{array}, \tag{4.25}$$

where G_{n2} is invertible. Next, pre-multiply the state-equation (4.16) (with the structures in (4.18)) with R and add $V\dot{y}$ to both sides to obtain

$$R\bar{E}\dot{\bar{x}} + V\dot{y} = \left(R\bar{E} + V\bar{C}\right)\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\dot{y}. \tag{4.26}$$

Suppose that $R\bar{E} + V\bar{C} = I_{\bar{n}}$. Hence (4.26) becomes

$$\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\bar{C}\dot{\bar{x}}. \tag{4.27}$$

Corollary 4.1 *The matrices* R, V *from* (4.22)–(4.23) *will have the following structures:*

$$R = \begin{bmatrix} I_{\bar{n}-p} & R_2 \\ 0 & R_4 \end{bmatrix}, \; V = \begin{bmatrix} -(\bar{E}_{21} + R_2 T_p E_{22}) \\ I_p - R_4 T_p E_{22} \end{bmatrix}, \tag{4.28}$$

where $|R_4| \neq 0$. ♯

Proof Since $\text{rank}\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$, then the matrices R and V can be chosen such that

$$\begin{bmatrix} R & V \end{bmatrix} \begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = I_{\bar{n}}, \tag{4.29}$$

i.e. $\begin{bmatrix} R & V \end{bmatrix}$ is chosen to be the Moore-Penrose inverse of $\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix}$. Partition the matrices R and V generally as follows:

$$\begin{bmatrix} R & V \end{bmatrix} = \left[\begin{array}{cc|c} R_1 & R_2 & V_1 \\ R_3 & R_4 & V_2 \end{array}\right], \tag{4.30}$$

where $R_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$ and $R_4 \in \mathbb{R}^{p\times p}$. By substituting $\bar{E}$ and $\bar{C}$ from (4.21) and R and V from (4.30) into (4.29), it is straightforward to see that R and V would take the forms given in (4.28). ■

Next, substitute z from (4.23) into (4.22) to obtain

$$\begin{aligned}\left(V\dot{y}-\dot{\hat{\bar{x}}}\right) &= \left(R\bar{A}-G_l\bar{C}\right)\left(Vy-\hat{\bar{x}}\right)-R\bar{B}u-\left(G_l\left(I_p-\bar{C}V\right)+R\bar{A}V\right)y-G_n\nu \\ \Rightarrow \dot{\hat{\bar{x}}} &= \left(R\bar{A}-G_l\bar{C}\right)\hat{\bar{x}}+R\bar{B}u+G_l\bar{C}\hat{x}+G_n\nu+V\bar{C}\dot{\bar{x}}. \end{aligned} \tag{4.31}$$

Define the state estimation error e as follows:

$$e=\begin{bmatrix} e_2 \\ e_y \end{bmatrix}\begin{array}{l} \updownarrow\ \ \bar{n}-p \\ \updownarrow\ \ p \end{array}. \tag{4.32}$$

Hence by subtracting (4.27) from (4.31), the error equation (which characterises the performance of the observer) becomes

$$\dot{e}=(R\bar{A}-G_l\bar{C})e-R\bar{M}\bar{f}-R\bar{Q}\xi+G_n\nu. \tag{4.33}$$

Therefore, expand (4.33) using (4.21) and (4.32) to obtain

$$\begin{aligned}\begin{bmatrix}\dot{e}_2\\ \dot{e}_y\end{bmatrix} &= \begin{bmatrix} A_2+(R_2-X)T_pA_5 & A_3+(R_2-X)T_pA_6-G_{l1} \\ R_4T_pA_5 & R_4T_pA_6-G_{l2}\end{bmatrix}\begin{bmatrix} e_2 \\ e_y\end{bmatrix} \\ &\quad -\begin{bmatrix} R_2T_pM_2 & R_2T_pA_4 \\ R_4T_pM_2 & R_4T_pA_4\end{bmatrix}\begin{bmatrix} f \\ x_1\end{bmatrix}-\begin{bmatrix} Q_1+(R_2-X)T_pQ_2 \\ R_4T_pQ_2\end{bmatrix}\xi+\begin{bmatrix} 0 \\ G_{n2}\end{bmatrix}\nu. \end{aligned} \tag{4.34}$$

4.3.1 Convergence of the sliding mode observer

Proposition 4.2 *Suppose there exists a positive-definite matrix P that satisfies*

$$P(R\bar{A}-G_l\bar{C})+(R\bar{A}-G_l\bar{C})^TP<0, \tag{4.35}$$

where $P=\begin{bmatrix} P_1 & 0 \\ 0 & G_{n2}^{-1}\end{bmatrix}$, $P_1\in\mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$. *If ρ in* (4.24) *is chosen as follows:*

$$\begin{aligned} \rho &\geq 2\|G_{n2}^{-1}R_4T_pA_5\|\mu_1\beta/\mu_0+\|G_{n2}^{-1}R_4T_p\begin{bmatrix} M_2 & A_4\end{bmatrix}\|\alpha+\|G_{n2}^{-1}R_4T_pQ_2\|\beta+\eta, \\ \mu_0 &= -\lambda_{\max}\left(P(R\bar{A}-G_l\bar{C})+(R\bar{A}-G_l\bar{C})^TP\right), \\ \mu_1 &= \|PR\bar{Q}\|,\ \alpha\geq\|f\|_{\max}+\|x_1\|_{\max},\ \beta\geq\|\xi\|, \end{aligned} \tag{4.36}$$

where η is an arbitrarily small positive scalar, then an ideal sliding motion takes place on the surface $\mathcal{S}=\{e:\bar{C}e=0\}$ in finite time. ♯

Proof The first portion of the proof is to show that e can be made to be ultimately bounded by setting a suitable value for ρ. Define a candidate Lyapunov function $W = e^T Pe > 0$. Differentiating W with respect to time yields

$$\dot{W} = e^T(P(R\bar{A} - G_l\bar{C}) + (R\bar{A} - G_l\bar{C}^T P))e - 2e^T PR\bar{M}\bar{f} - 2e^T PR\bar{Q}\xi + 2e^T PG_n\nu$$
$$\leq -\mu_0\|e\|^2 + 2\|e\|\mu_1\beta - 2\|e_y\|(\rho - \|G_{n2}^{-1}\bar{C}R\bar{M}\|\alpha). \tag{4.37}$$

If ρ is set to satisfy

$$\rho \geq \|G_{n2}^{-1}\bar{C}R\bar{M}\|\alpha + \eta, \tag{4.38}$$

then (4.37) becomes

$$\dot{W} \leq \|e\|\left(-\mu_0\|e\| + 2\mu_1\beta\right). \tag{4.39}$$

When the magnitude of the error e is smaller than or equal than a certain bound, i.e. $\|e\| \leq \frac{2\mu_1\beta}{\mu_0}$, then (4.39) becomes

$$\dot{W} \leq \kappa, \tag{4.40}$$

where $\kappa \geq 0$, and thus the magnitude of e can grow, shrink, or remain constant within this bound. It can however be seen that if the magnitude of the error e is larger than the aforementioned bound, i.e. $\|e\| > \frac{2\mu_1\beta}{\mu_0}$, then (4.39) becomes

$$\dot{W} < 0, \tag{4.41}$$

and $\|e\|$ would shrink. This implies that the magnitude of e will be bounded $\left(\|e\| \leq \frac{2\mu_1\beta}{\mu_0}\right)$ in finite time if (4.38) is satisfied.

The next (and remaining) portion of the proof aims to show how sliding motion $(e_y, \dot{e}_y = 0)$ is induced. Define another candidate Lyapunov function $W_y = e_y^T G_{n2}^{-1} e_y > 0$. Differentiating W_y with respect to time yields

$$\begin{aligned}\dot{W}_y = {} & e_y^T\left(G_{n2}^{-1}(R_4T_pA_6 - G_{l2}) + (R_4T_pA_6 - G_{l2})^T G_{n2}^{-1}\right)e_y \\ & + 2e_y^T\left(G_{n2}^{-1}R_4T_pA_5e_2 - G_{n2}^{-1}R_4T_p\begin{bmatrix} M_2 & A_4\end{bmatrix}\bar{f} - G_{n2}^{-1}R_4T_pQ_2\xi + \nu\right) \\ \leq {} & -2\|e_y\|(\rho - 2\|G_{n2}^{-1}R_4T_pA_5\|\mu_1\beta/\mu_0 + \|G_{n2}^{-1}R_4T_p\begin{bmatrix} M_2 & A_4\end{bmatrix}\alpha \\ & + \|G_{n2}^{-1}R_4T_pQ_2\|\beta).\end{aligned} \tag{4.42}$$

Notice that

$$\|e_y\|^2 = \left(\sqrt{G_{n2}^{-1}}e_y\right)^T G_{n2}\left(\sqrt{G_{n2}^{-1}}e_y\right) \geq \lambda_{\min}\left(G_{n2}\right)\|\sqrt{G_{n2}^{-1}}e_y\|^2 = \lambda_{\min}\left(G_{n2}\right)W_y. \quad (4.43)$$

Then, by setting ρ such that (4.36) is satisfied (which also satisfies (4.38)), inequality (4.42) becomes

$$\dot{W}_y \leq -2\eta\|e_y\| \leq -2\eta\sqrt{\lambda_{\min}\left(G_{n2}\right)}\sqrt{W_y}, \quad (4.44)$$

which is the *reachability condition* that will result in $e_y = 0$ in finite time, and a sliding motion takes place on the surface $\mathcal{S} = \{e : \bar{C}e = 0\}$ [128]. ■

Remark 4.1 *Notice from* (4.36) *that x, f, ξ need to be bounded, and the bounds need to be known in order to calculate ρ. In practical situations, the bounds could be obtained by knowing the physical operation of the system, or by simulating the system operating under practical scenarios. It is reasonable to assume that x is bounded, as a controller will be put in place to stabilise the system, but the design of the controller is not considered as it has no effect on the fault reconstruction, since u does not appear in the error equation from* (4.33) *onwards.*

Since the upper bounds of $\|x\|, \|f\|, \|\xi\|$ are only required to calculate ρ in (4.36)*, in situations where it is not easy to obtain these upper bounds, ρ can be set to be adaptive, where the magnitude of ρ can be adjusted to achieve convergence of e_y. This way, the bounds of $\|x\|, \|f\|, \|\xi\|$ do not need to be known a-priori, and ρ will also be smaller (less conservative), which can reduce chattering. One such adaptive scheme to calculate ρ is from [47], given as follows:*

$$\rho(t) = \bar{\rho} + \beta\left(e_y(t)\right) + \eta_0, \quad (4.45)$$

where $\beta(\cdot)$ is an exponentially decaying term driven by the output estimation error, η_0 is an arbitrarily small positive scalar, and $\bar{\rho}$ is given by

$$\dot{\bar{\rho}} = \alpha_0\Phi\left(\|e_y(t)\|\right) - \alpha_1\bar{\rho(t)},\ \bar{\rho}(0) = \rho_0, \quad (4.46)$$

where α_0 and α_1 are positive scalars, $\rho_0 \geq 0$, and $\Phi(\cdot)$ is a dead-zone function such that

$$\Phi(x) = \begin{cases} 0 & \textit{if } \|x\| \leq \varepsilon \\ x - \varepsilon\,\mathrm{sign}(x) & \textit{otherwise,} \end{cases}$$

where ε is a positive scalar. If $\bar{f}$ and ξ remain bounded, then e_y would be driven to a boundary layer about the sliding surface $\mathcal{S} = \{e : \bar{C}e = 0\}$. The proof for this design procedure can be found in [47]. ♯

Remark 4.2 *Chattering is an inherent problem arising in SMOs which may corrupt the reconstruction of fault signals. The effects of chattering can be alleviated through proper tuning of ρ, and by introducing a small positive scalar δ such that ν in* (4.24) *becomes*

$$\nu = -\rho \frac{e_y}{\|e_y\| + \delta}. \tag{4.47}$$

Setting δ to be larger and ρ to be smaller (while still satisfying (4.36)*) reduces chattering in the fault reconstruction. The parameter δ, however, causes the sliding motion of e_y to take place within a small boundary layer near 0 instead of perfectly sliding on the surface $\mathcal{S} = \{e : \bar{C}e = 0\}$. A larger δ would result in a larger boundary layer, which in turn diminishes the accuracy of the reconstructed fault [8].* ♯

Remark 4.3 *Note that the requirement for f to be bounded even though $x(0)$ is unknown reduces the scheme's applicability since it does not cater to cases where f is unbounded. This is however not a stringent restriction: practical systems operate around a certain operating point, and by finding reasonable bounds on the faults affecting the system as per Remark 4.1, the scheme can still be applied across a wide range of systems.* ♯

After sliding motion is achieved, $e_y, \dot{e}_y = 0$, and system (4.34) becomes

$$\dot{e}_2 = \left(A_2 + (R_2 - X)\,T_p\bar{A}_5\right)e_2 - R_2T_p\begin{bmatrix} M_2 & A_4 \end{bmatrix}\bar{f} - (Q_1 + (R_2 - X)T_pQ_2)\xi, \tag{4.48}$$

$$0 = R_4T_pA_5e_2 - R_4T_p\begin{bmatrix} M_2 & A_4 \end{bmatrix}\bar{f} - R_4T_pQ_2\xi + G_{n2}\nu. \tag{4.49}$$

4.3.2 Robustly reconstructing the fault

Define a measurable signal $\bar{\nu} = R_4^{-1}G_{n2}\nu$. Equation (4.49) can then be rearranged to become

$$\bar{\nu} = -T_pA_5e_2 + T_pQ_2\xi + T_pA_4x_1 + T_pM_2f. \tag{4.50}$$

For the purpose of fault reconstruction, it is desirable to be able to extract f without influence from x_1, i.e. a linear combination of (4.50) can be formed such that x_1 is eliminated. To achieve this, it is required that

$$\operatorname{rank}(T_p M_2) = q, \tag{4.51}$$

$$\operatorname{rank}\begin{bmatrix} T_p M_2 & T_p A_4 \end{bmatrix} = \operatorname{rank}(T_p M_2) + \operatorname{rank}(T_p A_4). \tag{4.52}$$

Equation (4.51) implies that all components of f can be recovered, while (4.52) implies that the columns of $T_p M_2$ and $T_p A_4$ are independent of each other so that x_1 does not influence f. Combining (4.51)–(4.52) results in

$$\operatorname{rank}\begin{bmatrix} T_p M_2 & T_p A_4 \end{bmatrix} - \operatorname{rank}(T_p A_4) = q. \tag{4.53}$$

Proposition 4.3 *A suitable choice of* T_p *exists to satisfy* (4.18) *and* (4.53) *if and only if E1–E3 hold true.* ♯

Proof To prove the necessity of E1, suppose that it does not hold, i.e.

$$\operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ M_2 & A_4 \end{bmatrix} < \operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix}. \tag{4.54}$$

Notice that (4.54) can never hold, since adding more rows to a matrix does not reduce its rank - it is assumed only to show the necessity of E1. Since $\operatorname{rank}\left(T_p \begin{bmatrix} M_2 & A_4 \end{bmatrix}\right) \leq \operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix}$, if E1 does not hold, then

$$\operatorname{rank}\left(T_p \begin{bmatrix} M_2 & A_4 \end{bmatrix}\right) < \operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ T_p M_2 & T_p A_4 \end{bmatrix}, \tag{4.55}$$

which violates the assumption made in (4.18), thus proving the necessity of E1.

Then to prove the necessity of E2, suppose that it is not met, i.e.

$$\operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix} < q + r_a, \tag{4.56}$$

which implies that M_2 and A_4 would have dependent columns, or that M_2 does not have full-column rank, or both. Hence, the columns of $T_p M_2$ and $T_p A_4$ will be dependent, or even if they did have independent columns, $\operatorname{rank}(M_2) < q$, and (4.53) will never hold, thus proving the necessity of E2.

Since $\operatorname{rank}(A_4) = r_a$, use SVD to decompose A_4 into

$$A_4 = T_{c1}^T \begin{bmatrix} I_{r_a} & 0 \\ 0 & 0 \end{bmatrix} T_{d1}^T, \tag{4.57}$$

where T_{c1} and T_{d1} are orthogonal. Since $A_4 \in \mathbb{R}^{(n-\bar{n}+p)\times(n-\bar{n})}$, then $r_a \le n - \bar{n}$. Introduce a state equation transformation T_c and a state transformation T_d such that $x \mapsto T_d x$, where

$$T_c = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{c1} \end{bmatrix}, \; T_d = \begin{bmatrix} T_{d1} & 0 \\ 0 & I_{\bar{n}} \end{bmatrix}. \tag{4.58}$$

Apply these transformations onto M and A in the coordinates of (4.5) to obtain

$$M \mapsto T_c M = \left[\begin{array}{c} M_1 \\ \hline M_{21} \\ M_{22} \end{array}\right], \; A \mapsto T_c A T_d^{-1} = \left[\begin{array}{cc|c|c} A_{11} & A_{12} & A_2 & A_3 \\ \hline I_{r_a} & 0 & A_{51} & A_{61} \\ 0 & 0 & A_{52} & A_{62} \end{array}\right]. \tag{4.59}$$

Substitute the structures of M and A from (4.59) into E2 to obtain

$$\operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix} - \operatorname{rank}(A_4) = \operatorname{rank}\left[\begin{array}{c|cc} M_{21} & I_{r_a} & 0 \\ M_{22} & 0 & 0 \end{array}\right] - \operatorname{rank}\begin{bmatrix} I_{r_a} & 0 \\ 0 & 0 \end{bmatrix} = q. \tag{4.60}$$

Equation (4.60) implies

$$\operatorname{rank}(M_{22}) = q. \tag{4.61}$$

Thus, perform QR decomposition on M_{22} to obtain

$$T_{e1} M_{22} = \begin{bmatrix} I_q \\ 0 \end{bmatrix}, \tag{4.62}$$

where T_{e1} is invertible. Hence define the state equation transformation T_e such that

$$T_e = \left[\begin{array}{cc|cc} I_{\bar{n}-p} & 0 & 0 & 0 \\ 0 & 0 & I_q & 0 \\ \hline 0 & I_{r_a} & -M_{21} & 0 \\ 0 & 0 & 0 & I_{n-\bar{n}+p-q-r_a} \end{array}\right] \begin{bmatrix} I_{\bar{n}-p+r_a} & 0 \\ 0 & T_{e1} \end{bmatrix}. \tag{4.63}$$

Apply this transformation onto the system in the coordinates of (4.59) to obtain

$$M \mapsto T_e M = \left[\begin{array}{c} M_1 \\ \hline I_q \\ 0 \\ 0 \end{array}\right], \; A \mapsto T_e A = \left[\begin{array}{cc|c|c} A_{11} & A_{12} & A_2 & A_3 \\ \hline 0 & 0 & A_{51} & A_{61} \\ I_{r_a} & 0 & A_{52} & A_{62} \\ 0 & 0 & A_{53} & A_{63} \end{array}\right]. \tag{4.64}$$

The structures of M and A from (4.64) imply the following

$$\begin{aligned}
\operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ M_2 & A_4 \end{bmatrix} &= \operatorname{rank}\left[\begin{array}{c|cccc} M_1 & A_{11} & A_{12} & A_2 & A_3 \\ \hline I_q & 0 & 0 & A_{51} & A_{61} \\ 0 & I_{r_a} & 0 & A_{52} & A_{62} \\ 0 & 0 & 0 & A_{53} & A_{63} \end{array}\right], \\
\operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix} &= \operatorname{rank}\left[\begin{array}{c|cccc} I_q & 0 & 0 & A_{51} & A_{61} \\ 0 & I_{r_a} & 0 & A_{52} & A_{62} \\ 0 & 0 & 0 & A_{53} & A_{63} \end{array}\right].
\end{aligned} \tag{4.65}$$

By substituting for the structures of M and A from (4.65), condition E1 (which states that the LHS terms in (4.65) are equal) implies $A_{12} = 0$. Therefore, set

$$T_p = \begin{bmatrix} I_{q+r_a} & 0 \\ 0 & T_{33} \end{bmatrix}, \tag{4.66}$$

where $T_{33} \in \mathbb{R}^{(p-q-r_a)\times(n-\bar{n}+p-q-r_a)}$ is chosen such that T_p is full-row rank. To show the necessity of E3, suppose it is not satisfied, i.e. $p < q + \operatorname{rank}(A_4)$. Then it is straightforward to see in this case that T_{33} (and therefore T_p) does not exist, violating the proposition that a suitable choice of T_p exists to satisfy (4.18) and (4.53). On the other hand, the sufficiency of E3 is obvious: if E3 is satisfied, i.e. $p \geq q + r_a$, then T_{33} exists and can be freely chosen such that T_p is full-row rank. Thus, Proposition 4.3 is proven. ■

Let R_2 from (4.28) and Q_2 in the coordinates of (4.64) take the forms

$$R_2 = \begin{bmatrix} 0 & L \end{bmatrix}, \; Q_2 = \begin{array}{c} \begin{bmatrix} Q_{21} \\ Q_{22} \\ Q_{23} \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; q \\ \updownarrow \; r_a \\ \updownarrow \; n-\bar{n}+p-q-r_a \end{array}, \tag{4.67}$$

where $L \in \mathbb{R}^{(\bar{n}-p)\times(p-q-r_a)}$ represents design freedom. By substituting for the structures of M and A in (4.64), then $\bar{M}$ from (4.16), X from (4.19), and $\bar{A}_2$ from (4.21) would have the following forms:

$$\bar{M} = \left[\begin{array}{c|cc} M_1 & A_{11} & 0 \\ \hline I_q & 0 & 0 \\ 0 & I_{r_a} & 0 \\ 0 & 0 & 0 \end{array}\right], \; X = \begin{bmatrix} M_1 & A_{11} & 0 \end{bmatrix}, \; \bar{A}_2 = A_2 - M_1 A_{51} - A_{11} A_{52}. \tag{4.68}$$

Therefore, using (4.50) and the structure of $\bar{M}$ from (4.68), define the fault reconstruction signal to be

$$\hat{f} = Y\bar{\nu}, \tag{4.69}$$

where $Y = \begin{bmatrix} I_q & 0 & W \end{bmatrix}$, where $W \in \mathbb{R}^{q\times(p-q-r_a)}$ represents design freedom. Pre-multiply (4.50) with Y, and substitute for A_5 and M_2 from (4.64), T_p from (4.66), Q_2 from (4.67), and A_4 from (4.68) to obtain

$$\hat{f} = \left(-A_{51} - WT_{33}A_{53}\right)e_2 + \left(Q_{21} + WT_{33}Q_{23}\right)\xi + f, \tag{4.70}$$

which estimates f completely and x_1 does not appear. However, e_2 and ξ affect $\hat{f}$; this issue will be addressed later in §4.3.3 when the observer is designed to minimise their effects on $\hat{f}$.

Define the fault reconstruction error

$$e_f = \hat{f} - f. \tag{4.71}$$

Rearranging (4.48)–(4.49) using (4.66)–(4.71) yields

$$\dot{e}_2 = \underbrace{\left(\bar{A}_2 + LT_{33}A_{53}\right)}_{\breve{A}} e_2 + \underbrace{\left(-Q_1 + M_1Q_{21} + A_{11}Q_{22} - LT_{33}Q_{23}\right)}_{\breve{B}}\xi, \tag{4.72}$$

$$e_f = \underbrace{\left(-A_{51} - WT_{33}Q_{53}\right)}_{\breve{C}} e_2 + \underbrace{\left(Q_2 + WT_{33}Q_{23}\right)}_{\breve{D}}\xi, \tag{4.73}$$

which is a state-space system that shows how the disturbance ξ affects the fault reconstruction error e_f. It can be seen that for e_f to be stable, the error system (4.72)–(4.73) needs to be stable.

Proposition 4.4 *For a suitable T_{33} (and hence T_p) to exist such that the error system* (4.72)–(4.73) *is stable, E4 is necessary. If $p - q - r_a > 0$, then E4 is also sufficient; otherwise (for $p - q - r_a = 0$), E5 is sufficient.* ♯

Proof Expand the left-hand side of E4 using (4.10):

$$\operatorname{rank}\underbrace{\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix}}_{R(s)} = \operatorname{rank}\underbrace{\begin{bmatrix} -A_1 & sI - A_2 & M_1 \\ -A_4 & -A_5 & M_2 \end{bmatrix}}_{R_2(s)} + p, \tag{4.74}$$

where $R(s)$ is the Rosenbrock matrix of (E, A, M, C), and the values of s that make it lose rank are the zeros of the system (E, A, M, C) [109]. Substitute for A_1, A_4, A_5, and M_2 from (4.64) to obtain

$$\begin{aligned}\operatorname{rank}(R_2(s)) &= \operatorname{rank}\begin{bmatrix} I_{\bar{n}-p} & -M_1 & -A_{11} & 0 \\ 0 & I_q & 0 & 0 \\ 0 & 0 & I_{r_a} & 0 \\ 0 & 0 & 0 & -I_{p-q-r_a} \end{bmatrix}\begin{bmatrix} A_{11} & 0 & sI_{\bar{n}-p} - A_2 & M_1 \\ 0 & 0 & -A_{51} & I_q \\ -I_{r_a} & 0 & -A_{52} & 0 \\ 0 & 0 & -A_{53} & 0 \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} 0 & sI - \bar{A}_2 & 0 \\ 0 & -A_{51} & I_q \\ I_{r_a} & A_{52} & 0 \\ 0 & A_{53} & 0 \end{bmatrix} = \operatorname{rank}\underbrace{\begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ A_{53} \end{bmatrix}}_{D(s)} + q + r_a. \end{aligned} \tag{4.75}$$

From the Popov-Hautus-Rosenbrock (PHR) rank test [60], if the values of s that make $D(s)$ lose rank (i.e. the unobservable modes of $\bar{A}_2$) are stable, the pair $\left(\bar{A}_2, A_{53}\right)$ is said to be *detectable* - hence E4 can be recast as: $\left(\bar{A}_2, A_{53}\right)$ is *detectable*.

<u>Proof of Necessity:</u>

Recall that for error state equation (4.72) to be stable, it is required that $\breve{A}$ is stable, which implies $\lambda\left(\bar{A}_2 + LT_{33}A_{53}\right) < 0$, i.e. $\left(\bar{A}_2, T_{33}A_{53}\right)$ is *detectable*. Notice that the detectability of $\left(\bar{A}_2, T_{33}A_{53}\right)$ depends also on T_{33}, which constitutes design freedom. Hence, the requirement is recast as: matrix T_{33} exists such that $\left(\bar{A}_2, T_{33}A_{53}\right)$ is detectable. From the PHR rank test, if the values of s that make the following matrix $D_2(s)$ lose rank are stable, then $\left(\bar{A}_2, T_{33}A_{53}\right)$ is said to be *detectable*, whereby

$$D_2(s) = \begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ T_{33}A_{53} \end{bmatrix} = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{33} \end{bmatrix}\underbrace{\begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ A_{53} \end{bmatrix}}_{D(s)}. \tag{4.76}$$

From (4.76), it follows that $\operatorname{rank}(D_2(s)) \leq \operatorname{rank}(D(s))$. Therefore if a value of s makes $D(s)$ lose rank, it will also make $D_2(s)$ lose rank, and hence the zeros of (E, A, M, C) are also the unobservable modes of $\left(\bar{A}_2, T_{33}A_{53}\right)$. This therefore shows that E4 is necessary for $\left(\bar{A}_2, T_{33}A_{53}\right)$ to be detectable.

Proof of Sufficiency:

For the case where $p - q - r_a > 0$, let Z be a matrix containing the generalised right-eigenvectors of $\bar{A}_2$. Therefore $Z^{-1}\bar{A}_2 Z$ is a matrix in the *Jordan canonical form*, where the diagonal elements are the real parts of the eigenvalues of $\bar{A}_2$ [98]. Pre-multiply $D_2(s)$ from (4.76) with $\begin{bmatrix} Z^{-1} & 0 \\ 0 & I_{p-q-r_a} \end{bmatrix}$ and post-multiply with Z, i.e.

$$\begin{bmatrix} Z^{-1} & 0 \\ 0 & I_{p-q-r_a} \end{bmatrix} \begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ T_{33}A_{53} \end{bmatrix} Z = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{33} \end{bmatrix} \begin{bmatrix} Z^{-1}(sI_{\bar{n}-p} - \bar{A}_2)Z \\ A_{53}Z \end{bmatrix}. \tag{4.77}$$

A zero of (E, A, M, C), which is an unobservable mode of $\left(Z^{-1}\bar{A}_2 Z, T_{33}A_{53}Z\right)$, will therefore appear as an element of $Z^{-1}\bar{A}_2 Z$ where its corresponding column in $T_{33}A_{53}Z$ is zero. If E4 is satisfied, however, the columns within $A_{53}Z$ corresponding to positive diagonal elements (which indicate unstable eigenvalues) within $Z^{-1}\bar{A}_2 Z$ will be non-zero. Recall that T_{33} is design freedom. Thus, a single row within T_{33} can be chosen such that $T_{33}A_{53}Z$ has non-zero elements at the columns corresponding to the diagonals of $Z^{-1}\bar{A}_2 Z$ that are unstable in order to guarantee that the unstable modes are observable (and therefore are not a zero of (E, A, M, C)). Thus, the sufficiency of E4 when $p - q - r_a > 0$ is shown.

When $p - q - r_a = 0$, T_p in (4.66) becomes

$$T_p = \begin{bmatrix} I_{q+r_a} & 0 \end{bmatrix}. \tag{4.78}$$

Since T_{33} does not exist, $\breve{A}$ in (4.72) becomes $\bar{A}_2$, so it is sufficient that the eigenvalues of $\bar{A}_2$ are stable (i.e. E5 is satisfied) for the error system (4.72)–(4.73) to be stable. ■

Therefore, Propositions 4.3 and 4.4 have been proven: if E1–E4 (and E5 if necessary) are satisfied, the observer by Yeu et al. [157] can reconstruct f for system (4.1)–(4.2), thus proving Theorem 4.1. □

Remark 4.4 *If* $\xi = 0$ *(i.e. system* (4.1)–(4.2) *is not affected by external disturbances),* *error system* (4.72)–(4.73) *becomes*

$$\dot{e}_2 = (A_2 + LT_{33}A_{53})\, e_2, \tag{4.79}$$

$$e_f = -(A_{51} + WT_{33}Q_{53})\, e_2. \tag{4.80}$$

It can be seen that $e_f \to 0$, which implies $\hat{f} \to f$. Therefore, in the absence of external disturbances, the proposed scheme is able to asymptotically reconstruct the faults. ♯

4.3.3 Observer design for robust fault reconstruction

Recall from (4.72)–(4.73) that disturbance ξ affects the fault reconstruction error e_f and may therefore corrupt the fault reconstruction f. The objective now is to minimise the $\mathcal{L}_2$ gain from ξ to e_f, by choice of L and W. This would be achieved using linear matrix inequalities (LMIs). The following portion introduces LMIs, and motivates it with a simple example.

An introduction to linear matrix inequalities [124]

Many problems in system and control theory (e.g. LQR and $\mathcal{H}_\infty$) can be reduced into a few standard problems involving LMIs. An LMI has the form

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0, \tag{4.81}$$

where $x \in \mathbb{R}^m$ is a vector whose scalar variable(s) and the *symmetric matrices* F_i (where $i = 0, 1, \ldots, m$) are given quantities. The inequality sign in (4.81) implies $F(x)$ is positive definite. Generally, LMI problems can then be split into:

1. feasibility problems, where it is of interest to find a value for x that satisfies the LMI system $\mathcal{A}(x) < \mathcal{B}(x)$, where $\mathcal{A}(x)$ and $\mathcal{B}(x)$ are two affine functions in x (such as in (4.81)).

2. minimisation of a linear objective problems, where the goal is to minimise $f(x)$, where $f(x)$ is an affine function in x that satisfies $\mathcal{A}(x) < \mathcal{B}(x)$.

3. generalised eigenvalue problems, where the scalar λ is to be minimised, subject to $\mathcal{A}(x) < \mathcal{B}(x),\ \mathcal{C}(x) < \lambda\mathcal{D}(x),\ 0 < \mathcal{D}(x)$, where $\mathcal{C}(x)$ and $\mathcal{D}(x)$ are affine functions in x.

Each of these types of problems have their corresponding solvers in the MATLAB LMI Toolbox [52]. To illustrate the concept of LMIs, consider the following example from [124]: it is of interest to find the Lyapunov matrix $P = P^T > 0$ and matrices $G_l,\ A$, and C with the following structures

$$P = \begin{bmatrix} p_1 & 0 \\ 0 & p_o \end{bmatrix},\ G_l = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix},\ A = \begin{bmatrix} -2 & -3 \\ 1 & 3 \end{bmatrix},\ C = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{4.82}$$

where $p_1, p_o, g_1, g_2 \in \mathbb{R}$, such that P and G_l satisfy the inequality

$$P\left(A - G_l C\right) + \left(A - G_l C\right)^T P < 0. \tag{4.83}$$

Substituting (4.82) into (4.83) yields

$$\begin{aligned} &\begin{bmatrix} p_1 & 0 \\ 0 & p_o \end{bmatrix}\begin{bmatrix} -2 & -3-g_1 \\ 1 & 3-g_2 \end{bmatrix} + \begin{bmatrix} -2 & 1 \\ -3-g_1 & -3-g_2 \end{bmatrix}\begin{bmatrix} p_1 & 0 \\ 0 & p_o \end{bmatrix} < 0 \\ &\Rightarrow \begin{bmatrix} -4p_1 & -3p_1 + p_o - p_1 g_1 \\ -3p_1 + p_o - p_1 g_1 & 6p_o - 2p_o g_2 \end{bmatrix} < 0 \\ &\Rightarrow p_1 \begin{bmatrix} -4 & -3 \\ -3 & 0 \end{bmatrix} + p_o \begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix} + p_1 g_1 \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + p_o g_2 \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} < 0. \end{aligned} \tag{4.84}$$

Inequality (4.84) can then be re-written as an LMI of the form (4.81), where the constant matrices are

$$F_0 = 0_{2\times 2},\ F_1 = \begin{bmatrix} 4 & 3 \\ 3 & 0 \end{bmatrix},\ F_2 = \begin{bmatrix} 0 & -1 \\ -1 & -6 \end{bmatrix},\ F_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ F_4 = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \tag{4.85}$$

and the LMI variables are

$$x_1 = p_1,\ x_2 = p_o,\ x_3 = p_1 g_1,\ x_4 = p_o g_2. \tag{4.86}$$

The LMI Toolbox can then be used to solve (4.84), and given solutions for x_1, x_2, x_3, x_4, variables p_1, p_o, g_1, g_2 can be determined. Thus by using the LMI toolbox routine *feasp* (which calculates a feasible solution to (4.84)), the following values are obtained:

$$x_1 = p_1 = 2.130,\ x_2 = p_o = 1.633,\ x_3 = -5.126,\ x_4 = 8.158. \tag{4.87}$$

From the definitions of G_l and P in (4.82), and x_3, x_4 in (4.86), matrices G_l and P are found to be

$$G_l = \begin{bmatrix} -2.406 \\ 4.995 \end{bmatrix}, \quad P = \begin{bmatrix} 2.130 & 0 \\ 0 & 1.633 \end{bmatrix} \tag{4.88}$$

Thus a solution to inequality (4.83) has been found using LMIs. This example therefore serves as motivation to the next portion, where the observer design problem will be framed as an LMI system.

Using LMIs to minimise the effect of disturbances on the fault reconstruction error

To minimise the $\mathcal{L}_2$ gain from ξ to e_f, through the choice of L and W, the following lemma is used to frame the problem as an LMI:

Lemma 4.2 (The Bounded Real Lemma [26, 52, 128]) *Consider the following state-space system:*

$$\dot{x}_r = A_r x_r + B_r u_r, \tag{4.89}$$

$$y_r = C_r x_r + D_r u_r, \tag{4.90}$$

where $u_r \in \mathbb{R}^{m_r}, y_r \in \mathbb{R}^{p_r}$. The $\mathcal{L}_2$ norm from u_r to y_r does not exceed the positive scalar γ_r if the following inequality is satisfied:

$$\begin{bmatrix} P_r A_r + A_r^T P_r & P_r B_r & C_r^T \\ B_r^T P_r & -\gamma_r I_{m_r} & D_r^T \\ C_r & D_r & -\gamma_r I_{p_r} \end{bmatrix} < 0. \tag{4.91}$$

where $P_r = P_r^T > 0$. ♯

By applying the Bounded Real Lemma onto error system (4.72)–(4.73), the $\mathcal{L}_2$ gain from ξ to e_f will not exceed the positive scalar γ if there exists a matrix $P_1 = P_1^T > 0$ such that

$$\begin{bmatrix} P_1 \breve{A} + \breve{A}^T P_1 & P_1 \breve{B} & \breve{C}^T \\ \breve{B}^T P_1 & -\gamma I_h & \breve{D}^T \\ \breve{C} & \breve{D} & -\gamma I_q \end{bmatrix} < 0. \tag{4.92}$$

The objective is therefore to find the solution for γ, P_1, L, and W that minimises γ subject to inequality (4.92), while also satisfying

$$P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0, \tag{4.93}$$

which (as in Proposition 4.2) guarantees that sliding motion will occur. The choice of G_l is not unique; any value that satisfies (4.93) is suitable. In this chapter, G_l will be designed using the method by [128], which is to satisfy the following inequality:

$$\begin{bmatrix} P(R\bar{A}) + (R\bar{A})^T P - \gamma_0 \bar{C}^T(\bar{D}\bar{D}^T)^{-1}\bar{C} & -P(R\bar{Q}) & \bar{F}^T \\ -(R\bar{Q})^T P & -\gamma_0 I_h & \breve{D}^T \\ \bar{F} & \breve{D} & -\gamma_0 I_q \end{bmatrix} < 0, \tag{4.94}$$

where $\bar{D} \in \mathbb{R}^{p\times p}$, and

$$\bar{F} = \begin{bmatrix} \breve{C} & F_1 \end{bmatrix}, \; P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}. \tag{4.95}$$

Finally, G_l is calculated using $G_l = \gamma_0 P^{-1}\bar{C}^T(\bar{D}\bar{D}^T)^{-1}$. Note that some terms in the inequalities are not *affine*, whereby there are instances of non-linearity in the variables, i.e. $P_1 L$. This can be solved by introducing a *change in variables*. Define a new variable

$$J = P_1 L. \tag{4.96}$$

After the software solver returns values for P_1 and J, then L can be calculated as

$$L = P_1^{-1} J. \tag{4.97}$$

By applying this technique onto (4.93), the Bounded Real Lemma for system (4.72)–(4.73) has the form

$$\begin{bmatrix} P_1\bar{A}_2 + JT_{33}A_{53} + (P_1\bar{A}_2 + JT_{33}A_{53})^T & * & * \\ -\breve{Q}^T P_1 - (JT_{33}Q_{23})^T & -\gamma I_h & * \\ -(A_{51} - T_{33}A_{53}) & Q_{21} + WT_{33}Q_{23} & -\gamma I_q \end{bmatrix} < 0, \tag{4.98}$$

where $\breve{Q} = Q_1 - M_1 Q_{21} - A_{11} Q_{22}$, and $*$ are terms that make (4.98) symmetrical. Hence if a set of values for P_1, J, W, γ satisfying (4.98) can be found, the $\mathcal{L}_2$ gain from ξ to e_f will not exceed the positive scalar γ. Therefore, the observer design problem can be recast as:

- Minimise γ with respect to the variables $P_1, P_2, F_1, J, W, \gamma$ subject to (4.94) and (4.98), and $P_1, P_2 > 0$.

- The gain W is extracted as is from the solution to LMI (4.98), L is calculated from (4.97), and the observer gain matrices G_l and G_n are determined using:

$$G_l = \gamma_0 P^{-1}\bar{C}^T(\bar{D}\bar{D}^T)^{-1}, \; G_n = \begin{bmatrix} 0 \\ P_2^{-1} \end{bmatrix}. \tag{4.99}$$

Remark 4.5 *Let* $\breve{X} = P_1\bar{A}_2 + JT_{33}A_{53} + (P_1\bar{A}_2 + JT_{33}A_{53})^T$*. For the LMIs in* (4.94) *and* (4.98) *to have a feasible solution, it is necessary that* $\breve{X} < 0$ *as it is on the main diagonal of the LMI in* (4.98)*. For* $\breve{X} < 0$*, it is required that* $\left(\bar{A}_2, T_{33}A_{53}\right)$ *is detectable (so that* $\breve{A}$ *can be made stable), which has been shown in Proposition 4.4 to be equivalent to E4 (i.e.* (E, A, M, C) *is minimum phase) when* $p - q - r_a > 0$*. Hence E4 is a necessary condition for the LMIs to be feasible when* $p - q - r_a > 0$*.*

However, there is no analytical expression (nor condition) that guarantee the solvability of the LMIs (4.94) *and* (4.98)*, though it is possible to know the effect of the design parameters* $\bar{D}$ *and* γ_0 *on solvability as follows:*

- *Setting a smaller* γ_0 *would cause the diagonals of LMI* (4.94) *to become less negative; since LMI* (4.94) *is negative, this would restrict the feasibility of the LMI.*
- *Setting a smaller value of* $\bar{D}$ *causes the top-left element on the main diagonal of* (4.94) *to become more negative, increasing the feasibility of the LMI.*

Knowing these effects would allow the designer to choose appropriate values for $\bar{D}$ *and* γ_0*.* ♯

The solution of γ to the LMIs in (4.94) and (4.98) would be the upper bound of the $\mathcal{L}_2$ norm from ξ to e_f. As long as there exists an upper bound on $\|\xi\|$, then there would also exist an upper bound on e_f, given by $\|e_f\| \leq \gamma\|\xi\|$.

Remark 4.6 *For the specific case of regular state-space systems (i.e. when* $E = I_n$*), the approach proposed in this chapter would be similar to the method presented in [128]. This can be shown by letting* $E = I_n$ *and proceeding with the method outlined in this chapter. In fact, the proposed scheme in this chapter is also applicable with infinitely observable descriptor systems* $\left(\text{i.e. systems where } \operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n\right)$*. The proposed method is therefore more general than those presented in [128] and [160].* ♯

The observer scheme proposed in this chapter has been shown to be able to reconstruct f robustly against ξ, thus concluding its analysis. The strengths of the proposed scheme over existing works are summarised as follows:

- The proposed scheme is able to asymptotically reconstruct the faults (in the absence of disturbances) for a class of non-infinitely observable descriptor systems. This is not possible in the scheme by Yeu et al. [156, 157, 160], which requires the system to be infinitely observable.

- The existence conditions of the scheme are presented in terms of the original system matrices, allowing the designer to determine from the outset whether the scheme is applicable to the system. This improves on the work by Nikhoukhah et al. [97], where the feasibility of the scheme is analysed in terms of the augmented system matrices, but no analysis was conducted on the original system matrices.

- The proposed scheme is able to reconstruct the faults robustly against disturbances. This is in contrast to the findings in [97] and [100], which did not consider disturbances or their effects on the fault reconstruction.

4.3.4 Design procedure

The design procedure for the proposed observer scheme can be summarised as follows:

1. Use the transformations introduced in Proposition 4.1 to obtain the system with (E, C) taking the forms given in (4.5).

2. Check that the existence conditions E1–E4 (E5 if necessary) given in Theorem 4.1 hold for the system. If they are not satisfied, do not continue as the observer scheme is inapplicable for the system.

3. Apply the state equation transformation $T_e T_c$ and the state transformation $x \mapsto T_d x$, where T_c, T_d, T_e are given in Proposition 4.3.

4. Determine T_{33} such that the pair $\left(\bar{A}_2, T_{33} A_{53}\right)$ is detectable. Apply the state equation transformation T given in (4.12), where T_p is given in (4.66).

5. Derive the reduced-order system using (4.16)–(4.17).

6. Determine X from (4.68), and apply the state equation transformation T_l given in (4.19).

7. Choose an invertible matrix R_4 such that $R_4 T_p E_{22} + V_2 = I_p$.

8. Choose values for LMI design parameters $\bar{D}, \gamma_0$.

9. Use an LMI solver to determine P, J, W, γ from (4.94) and (4.98). Then calculate L from (4.97), and G_l, G_n from (4.99). If necessary, repeat from *step 8* until suitable values of G_l, G_n are obtained. A large γ_0 or small $\bar{D}$ increases the magnitude of G_l, amplifying disturbances, while a small γ_0 or large $\bar{D}$ restricts the feasibilty of the LMIs from (4.94) and (4.98).

10. If the upper bounds of $\|x\|$, $\|f\|$, and $\|\xi\|$ can be determined, calculate ρ from (4.36). Otherwise, define a suitable function $\beta(\cdot)$ in (4.45), and set suitable values for $\alpha_0, \alpha_1, \rho_0, \varepsilon$, and δ from (4.45)–(4.46).

11. If necessary, set a suitable value for δ in (4.47).

4.4 Simulation example

To demonstrate the effectiveness of the scheme in this chapter, consider a modified version of the chemical mixing tank given in [157] described by the following dynamical model:

$$\dot{c}_3 = -0.375c_3 - 0.0667q_3 + 0.1q_1, \tag{4.100}$$

$$0 = -q_3 + q_1, \tag{4.101}$$

$$\dot{c}_5 = 0.3c_3 + 0.0533q_3 - 0.5c_5 - 0.04q_5 + 0.02q_4, \tag{4.102}$$

$$0 = q_3 - q_5 + q_4, \tag{4.103}$$

where q_1 is the flow rate of the influent into the first tank, c_3 and q_3 are the concentration and flow rate of the influent from the first tank into the second tank, respectively, q_4 is the flow rate of influent from another pipe into the second tank, and c_5 and q_5 are the concentration and flow rate of the effluent from the second tank, respectively. The actuator signal for q_1 is generated by a faulty first-order device

$$\dot{q}_1 = -2q_1 + 2q_{1,e} + f_1, \tag{4.104}$$

where $q_{1,e}$ is the reference signal, and f_1 is the fault. Additionally, the pipe connecting the two tanks is known to be leaky, and thus (4.101) becomes

$$0 = -q_3 + q_1 - f_2, \tag{4.105}$$

where f_2 is the fault representing the leak.

4.4.1 System formulation

The system matrices (E, A, B, M) in the framework of (4.1)–(4.2) are

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; A = \begin{bmatrix} -0.375 & -0.0667 & 0 & 0 & 0.1 \\ 0 & -1 & 0 & 0 & 1 \\ 0.3 & 0.0533 & -0.5 & -0.04 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}, \tag{4.106}$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0.02 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}, \; M = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \tag{4.107}$$

for the system variables

$$x = \begin{bmatrix} \text{concentration, } c_3 \text{ (mol/l)} \\ \text{flow rate, } q_3 \text{ (l/s)} \\ \text{concentration, } c_5 \text{ (mol/l)} \\ \text{flow rate, } q_5 \text{ (l/s)} \\ \text{flow rate, } q_1 \text{ (l/s)} \end{bmatrix}, \; f = \begin{bmatrix} f_1 \text{ (l/s)} \\ f_2 \text{ (l/s)} \end{bmatrix}, \tag{4.108}$$

$$u = \begin{bmatrix} \text{flow rate reference, } q_{1,e} \text{ (l/s)} \\ \text{flow rate, } q_4 \text{ (l/s)} \end{bmatrix}.$$

There exists a third inlet pipe into the tank system ideally does not output any fluid, i.e. $q_2 = 0$. To demonstrate the robustness of the observation scheme, assume that the outlet leaks fluid, i.e. q_2 forms a disturbance signal ξ_1. At the system's operating point at equilibrium the leak results in disturbances in the concentration states. Additionally, the chemical concentration entering the first tank is uncertain - this uncertainty is modelled as another disturbance signal ξ_2. Hence define $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$, and assume that measurements are only available for $q_3,\ c_5$, and q_1, so Q and C have the form

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},\ C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.109}$$

Notice that $\operatorname{rank} \begin{bmatrix} E \\ C \end{bmatrix} = 4 < 5$, and therefore A1 is not satisfied, which means that the observer by Yeu et al. [157] is not applicable. We now design the observer scheme proposed in this chapter.

4.4.2 Observer design

To ease readability, the steps in §4.3.4 will be referred to in the following design of the observer scheme.

Step 1: To obtain the structures of the system in (4.5), apply the state equation transformation T_a and state transformation $x \mapsto (T_z T_b)^{-1} x$, where

$$T_a = I_5,\ (T_z T_b)^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.110}$$

Steps 2–5: The following partitions are obtained:

$$M_1 = \begin{bmatrix} 0 & 0 \end{bmatrix},\ M_2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix},\ A_1 = 0,\ A_4 = \begin{bmatrix} 0 \\ 0.04 \\ 1 \\ 0 \end{bmatrix}. \tag{4.111}$$

From (4.107), (4.109), and (4.111), the following can be verified:

$$\begin{aligned} &\operatorname{rank}\begin{bmatrix} M_1 & A_1 \\ M_2 & A_4 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} M_2 & A_4 \end{bmatrix} = 3 \ (= q + r_a),\ p - q - r_a = 0, \\ &\operatorname{rank}\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix} = 7 \ \forall \ s \in \mathbb{C}^+, \end{aligned} \tag{4.112}$$

that is, E1–E4 hold for the system. Note that since $p - q - r_a = 0$, E5 needs to be verified. The transformations $T_e T_c$ and T_d are chosen to be

$$T_e T_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0.04 & -0.0016 & 0 \\ 0 & 0 & 0.0399 & 0.9984 & 0 \\ 0 & 0 & -0.9992 & 0.04 & 0 \end{bmatrix}, \ T_d = I_5. \tag{4.113}$$

Since $p - q - r_a = 0,\ T_{33}$ does not exist, and hence $T = I_5$. Then $\bar{A}_2$ is found to be

$$\bar{A}_2 = -0.375 < 0, \tag{4.114}$$

that is, E5 is also satisfied, and the existence of the observer scheme proposed in this chapter is guaranteed. The proposed observer scheme is now designed.

Steps 5–6: The reduced-order system (4.16)–(4.17) is determined to be

$$\begin{aligned} &\bar{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.04 & 0 \\ 0 & 0 & 0.0399 & 0 \end{bmatrix}, \ \bar{A} = \begin{bmatrix} -0.375 & -0.0667 & 0 & 0.1 \\ 0 & 0 & 0 & -2 \\ 0.012 & 1.0005 & -0.02 & -1 \\ 0.012 & 1.0005 & -0.02 & 0 \end{bmatrix}, \\ &\bar{B} = \begin{bmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 0.0008 \\ 0 & 0.9992 \end{bmatrix}, \ \bar{M} = \begin{bmatrix} 0 \\ I_3 \end{bmatrix}, \bar{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.04 & 0 \\ 0.0399 & 0 \end{bmatrix}, \ \bar{C} = \begin{bmatrix} 0 & I_3 \end{bmatrix}. \end{aligned} \tag{4.115}$$

Since $X = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, T_l from (4.19) is determined to be $T_l = I_4$, and the reduced-order system in (4.21) is identical to (4.115). Notice that $\operatorname{rank}\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = 4$, and A1–A2 are satisfied for (4.115), and thus the observer by Yeu et al. [157] can be applied onto the reduced-order system (4.115). The reduced-order system (4.115) is found to have no invariant zeros.

Steps 7–8: The matrix R_4 and LMI design parameters $\bar{D}$ and γ_0 were set as

$$R_4 = \begin{bmatrix} 0 & 1 & -0.9992 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \ \bar{D} = 0.1 I_3, \ \gamma_0 = 3. \tag{4.116}$$

Step 9: Using the Robust Control Toolbox within MATLAB, the following values for the LMI variables were obtained:

$$P = \begin{bmatrix} 0.5444 & 0 & 0 & 0 \\ 0 & 485.4 & -174.3 & 8.633 \\ 0 & -174.3 & 138.1 & -0.6325 \\ 0 & 8.633 & -0.6325 & 11.35 \end{bmatrix}, \ \gamma = 1.0007. \tag{4.117}$$

Note that the value of γ being close to unity is most likely due to there not being sufficient design freedom to attenuate the effect of ξ on the fault reconstruction. The parameters $G_l, \ G_n, \ R$, and V were calculated to be

$$\begin{gathered} G_l = \begin{bmatrix} 0 & 0 & 0 \\ 1.154 & 1.452 & -0.7964 \\ 1.452 & 4.001 & -0.8818 \\ -0.7964 & -0.8818 & 26.97 \end{bmatrix}, G_n = \begin{bmatrix} 0 & 0 & 0 \\ 0.0038 & 0.0048 & -0.0027 \\ 0.0048 & 0.0133 & -0.0029 \\ -0.0027 & -0.0029 & 0.0899 \end{bmatrix}, \\ R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -0.9992 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \end{bmatrix}, V = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0.9201 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{gathered} \tag{4.118}$$

The poles of the observer were found to be $\{-0.3744, -1.4681, -3.7361, -30.96\}$, while the pole of the sliding motion is $\{-0.375\}$.

Steps 10–11: To design ρ in this example, we used the adaptive method by [47] presented in Remark 4.1. For simplicity, let $\beta(\cdot) = 0$. The remaining parameters were set as $\alpha_0 = 10000, \alpha_1 = 0.003, \rho_0 = 500, \varepsilon = 5 \times 10^{-5}, \eta = 1$, while δ was set to be $\delta = 5 \times 10^{-4}$.

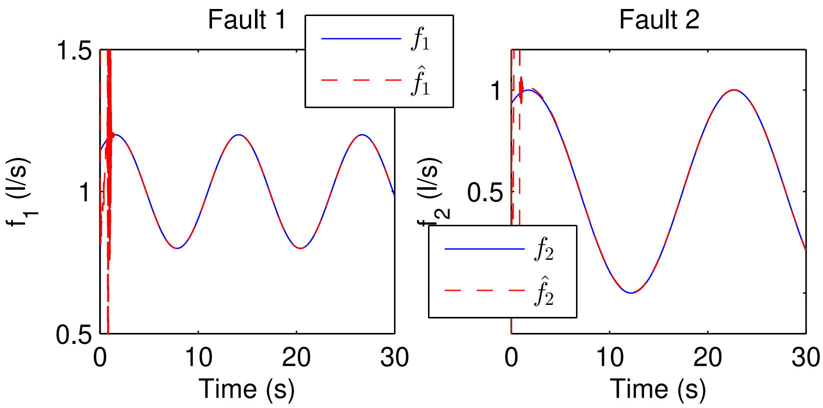

Figure 4.1: Faults (solid) and their reconstructions (dashed) in the disturbance-free scenario.

4.4.3 Simulation results

Two scenarios were simulated to demonstrate the effectiveness of the proposed scheme. The first scenario is disturbance-free ($\xi = 0$) to show the efficacy of the fault reconstruction. In the second scenario, the disturbance signals are non-zero to show the effectiveness of the approach at bounding the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction as follows:

$$\xi_1 = 0.1 \sin(0.1t + \pi) + 0.3, \ \xi_2 = 0.1 \sin\left(0.2t + \frac{\pi}{6}\right) + 0.2, \tag{4.119}$$

In both scenarios for the simulation, the initial condition of the system was set as $x(0) = \{2, 0.0670, 3, 6.067, 1\}$, while the initial condition of the observer was set at zero. The inputs $q_{1,e}$ and q_4 were set as step inputs with magnitudes of 5 and 6, respectively. The fault signals were simulated as

$$f_1 = 0.2 \sin\left(0.5t + \frac{\pi}{4}\right) + 1, \ f_2 = 0.5 \sin\left(0.3t + \frac{\pi}{3}\right) + 0.5. \tag{4.120}$$

Figure 4.1 shows the reconstruction of the faults in the disturbance-free scenario. It can be seen that asymptotic fault reconstruction is achieved, confirming the efficacy of the proposed method of fault reconstruction. Figure 4.2 shows the reconstruction of the faults in the scenario with non-zero disturbances using the same method. It can be seen that the reconstruction of both faults are affected by the disturbance signals. Figure 4.3 however shows that the magnitudes of the fault reconstruction errors for both faults are bounded

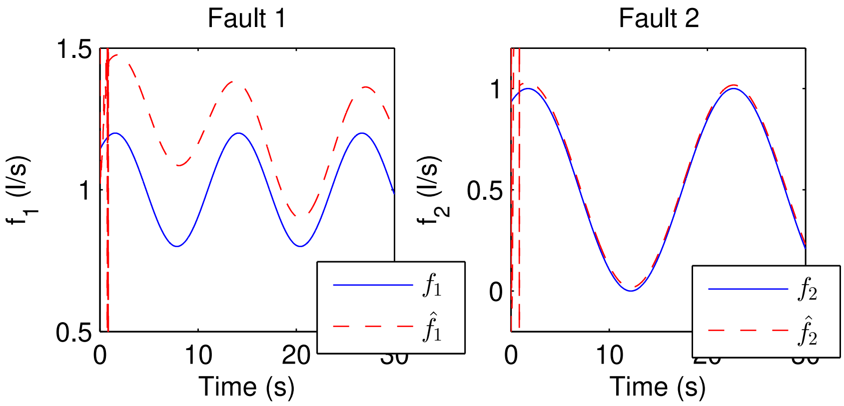

Figure 4.2: Faults (solid) and their reconstructions (dashed) in the scenario with non-zero disturbances.

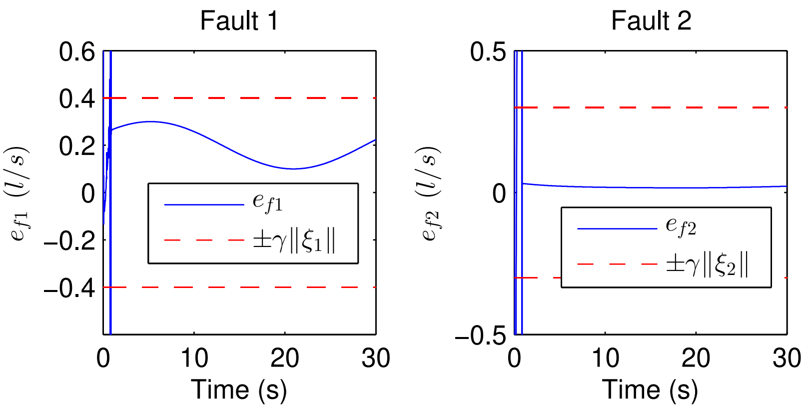

Figure 4.3: Fault reconstruction errors (solid) in the scenario with non-zero disturbances. The dashed lines represent the upper bounds of e_f derived from LMIs (4.94) and (4.98).

within $\pm\gamma\|\xi\|$ (which is indicated by the red dashed lines), verifying that the method is effective at bounding the effect of the disturbances on the fault reconstructions.

4.5 Conclusion

This chapter has presented a robust fault reconstruction scheme for a class of NIODS using SMOs, which improves on previous work that is only applicable to infinitely observable systems, or did not consider robustness in fault reconstruction. This was done by removing certain states and treating them as unknown inputs, thus creating a reduced-order system that is infinitely observable. A standard SMO scheme was then applied onto the reduced-order system to estimate the states and reconstruct the fault signal. The necessary and sufficient conditions for the existence of the scheme were presented in terms of the original system matrices. LMI techniques were used to minimise the effect of disturbances on the fault reconstruction, and the design procedure for the observer scheme was shown. Finally, a simulation was carried out, and the results verify the efficacy of the scheme.

Chapter 5

Estimation for NIODS using two sliding mode observers in cascade

5.1 Introduction

In the previous chapter, the sliding mode observer (SMO) by Yeu et al. [157] was used to reconstruct faults affecting a non-infinitely observable descriptor system (NIODS) such that the reconstructions are robust against disturbances. The scheme presented however requires the fault distribution matrix to satisfy restrictive rank constraints (in particular, E1–E4 in §4.3), limiting its applicability. This chapter thus aims to build on that work by developing a two-SMO scheme for a class of NIODS to relax the necessary and sufficient conditions required for the scheme, and thereby increasing its applicability.

The chapter first introduces a set of preliminary transformations that facilitates further analysis and observer design in §5.2. Similar to the previous chapter, certain states are then removed and treated as unknown inputs, forming an infinitely observable reduced-order system. In §5.3, the observer by Yeu et al. [157] is used to estimate the states of the system and some fault components. The switching term of the observer is then processed, and is found to be the output of an analytical regular (non-descriptor) state-space system that treats the remaining fault components as unknown inputs. The Edwards-Spurgeon observer [8] is then designed based on this analytical system to estimate the remaining faults. The necessary and sufficient existence conditions for the scheme are then presented. A design algorithm is then outlined. Finally, a simulated example is shown in §5.4 to demonstrate the efficacy of the observer scheme.

The work in this chapter has been published; its reference is J. C. L. Chan, C. P. Tan,

H. Trinh, and M. A. S. Kamal, State and fault estimation for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade, *Journal of the Franklin Institute*, 356(5):3010–3029, 2019.

5.2 Preliminary transformations

Consider the following non-infinitely observable descriptor system (NIODS):

$$E\dot{x}_0 = Ax_0 + Bu + Mf_0, \tag{5.1}$$

$$y = Cx_0, \tag{5.2}$$

where $E, A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m},\ M \in \mathbb{R}^{n\times q}$, and $C \in \mathbb{R}^{p\times n}$. The vectors $x_0,\ u,\ y$, and f_0 represent the states, inputs, outputs, and faults, respectively. Only u and y are measurable. The fault signal f_0 represents an abnormal condition acting upon the system that needs to be reconstructed so that information on its shape and magnitude could be obtained such that timely maintenance can be taken. The system is assumed to be regular (that is, there exists a scalar $\gamma \in \mathbb{C}$ such that $|sE - A| \neq 0$ [41]); this is not restrictive as most practical systems are regular [117]. Assume without loss of generality that E is rank deficient, i.e. $\mathrm{rank}(E) = r < n$, and that M and C are full-column rank and full-row rank, respectively, i.e. $\mathrm{rank}(M) = q,\ \mathrm{rank}(C) = p$. The objective of this chapter is to estimate x_0 and f_0 in finite time.

Similar to the previous chapters, system (5.1)–(5.2) is first re-expressed in a form facilitating further analysis using Lemma 4.1 and Proposition 5.1. Introduce a non-singular matrix $T_a = \begin{bmatrix} N_C^T \\ C \end{bmatrix}$, where N_C spans the null-space of C, i.e. $CN_C = 0$. Hence

$$E \mapsto ET_a^{-1} = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix},\ C \mapsto CT_a^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}. \tag{5.3}$$

Since $\mathrm{rank}\begin{bmatrix} E \\ C \end{bmatrix} = \bar{n}$, it follows from (5.3) that $\mathrm{rank}\,(E_{a1}) = \bar{n} - p$. Using singular-value decomposition (SVD), E_{a1} can be decomposed as

$$T_b E_{a1} T_{c1} = \begin{bmatrix} 0 & I_{\bar{n}-p} \\ 0 & 0 \end{bmatrix}, \tag{5.4}$$

where T_b and T_{c1} are non-singular matrices. Define another non-singular matrix $T_c = \begin{bmatrix} T_{c1}^{-1} & 0 \\ 0 & I_p \end{bmatrix}$, and hence matrix E is transformed to become

$$T_b E \left(T_c T_a\right)^{-1} = \begin{bmatrix} 0 & I_{\bar{n}-p} & E_1 \\ 0 & 0 & E_2 \end{bmatrix}. \tag{5.5}$$

In the coordinates of (5.5), matrices M and A are transformed accordingly and partitioned as

$$T_b M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \; T_b A \left(T_c T_a\right)^{-1} = \begin{bmatrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \end{bmatrix}, \tag{5.6}$$

where $M_2 \in \mathbb{R}^{(n-\bar{n}+p)\times q}$, $A_1 \in \mathbb{R}^{(\bar{n}-p)\times(n-\bar{n})}$ and $A_5 \in \mathbb{R}^{(n-\bar{n}+p)\times(\bar{n}-p)}$. Define $\mathrm{rank}(M_2) = q_2 < q$ and perform SVD on M_2 to obtain

$$T_{d1} M_2 T_e^{-1} = \begin{bmatrix} 0 & I_{q_2} \\ 0 & 0 \end{bmatrix}, \; T_{d1} \begin{bmatrix} A_4 & A_5 \end{bmatrix} = \begin{bmatrix} A_{41} & A_{51} \\ A_{42} & A_{52} \end{bmatrix}, \tag{5.7}$$

where $A_{52} \in \mathbb{R}^{(n-\bar{n}+p-q_2)\times(\bar{n}-p)}$, and T_{d1} and T_e are non-singular matrices.

Proposition 5.1 *Assume that*

$$\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p, \tag{5.8}$$

that is, $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ *is full-column rank (the satisfaction of* (5.8) *will be addressed later in Proposition 5.4). Then there exists a set of transformations introduced in Lemma 4.1 for system* (5.1)–(5.2) *such that system* (5.1)–(5.2) *would have the following structure:*

$$\begin{bmatrix} 0 & I_{\bar{n}-p} & E_1 \\ 0 & 0 & \bar{E}_2 \end{bmatrix} \dot{x} = \begin{bmatrix} A_1 & A_2 & A_3 \\ \bar{A}_4 & \bar{A}_5 & \bar{A}_6 \end{bmatrix} x + \begin{bmatrix} B_1 \\ \bar{B}_2 \end{bmatrix} u + \begin{bmatrix} M_{11} & M_{12} \\ 0 & \bar{M}_2 \end{bmatrix} f, \tag{5.9}$$

$$y = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} x, \tag{5.10}$$

$$\bar{E}_2 = \begin{bmatrix} E_{21} \\ E_{221} \\ E_{222} \\ E_{223} \end{bmatrix}, \; \bar{A}_4 = \begin{bmatrix} A_{41} \\ I_{n-\bar{n}} \\ 0 \\ 0 \end{bmatrix}, \; \bar{A}_5 = \begin{bmatrix} A_{51} \\ 0 \\ I_{\bar{n}-p} \\ 0 \end{bmatrix}, \; \bar{M}_2 = \begin{bmatrix} I_{q2} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{array}{l} \updownarrow \; q_2 \\ \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; 2p-\bar{n}-q_2 \end{array}, \tag{5.11}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} \begin{array}{l} \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; p \end{array}, \; f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{array}{l} \updownarrow \; q-q_2 \\ \updownarrow \; q_2 \end{array}, \tag{5.12}$$

where the partitions of $\bar{E}_2, \bar{A}_4, \bar{A}_5, \bar{M}_2$ have the same row dimensions, the dimensions of (E, A) are conformable to the partitions of x, whilst the column partitions of M are conformable to the partitions of f. ♯

Proof Since $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ is full column rank, apply QR decomposition on $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ such that

$$T_{f1} \begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}. \tag{5.13}$$

Define two non-singular matrices $T_d = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{d1} \end{bmatrix}$ and $T_f = \begin{bmatrix} I_{\bar{n}-p+q_2} & 0 \\ 0 & T_{f1} \end{bmatrix}$. Hence apply the state equation transformation $T_f T_d T_b$, the state transformation $T_c T_a$, and the fault transformation T_e such that

$$\begin{gathered} E \mapsto (T_f T_d T_b) E (T_c T_a)^{-1},\ A \mapsto (T_f T_d T_b) A (T_c T_a)^{-1},\ B \mapsto (T_f T_d T_b) B, \\ M \mapsto (T_f T_d T_b) M (T_e)^{-1},\ C \mapsto C (T_c T_a)^{-1},\ x_0 \mapsto (T_c T_a) x_0 = x,\ f_0 \mapsto T_e f_0 = f. \end{gathered} \tag{5.14}$$

It can be seen from the structures of M_2 from (5.7) and $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ from (5.13) that the system in the coordinates of (5.14) is identical to the structures given in (5.9)–(5.12), thus completing the proof. ■

Note that (5.8) is the necessary condition for the structure in (5.13) to be obtained. The system in the form of (5.9)–(5.12) allow the following comparison to be made between the proposed scheme and existing SMO schemes in the literature.

Remark 5.1 *Existing schemes that use a single SMO (such as [157] from Chapter 3, and [17, 100, 101]) are able to estimate all faults if and only if* $\operatorname{rank}(M_2) = \operatorname{rank}(M)$*, which is clearly not satisfied in this case, as shown in* (5.7)*. On the other hand, Ooi et al. [99] showed that a scheme consisting of a linear observer and a SMO in cascade can estimate* x_0 *and* f_0 *asymptotically, if and only if the following conditions are satisfied:*

F1. $$\operatorname{rank} \begin{bmatrix} M & A & 0 & E & 0 \\ 0 & E & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 \\ 0 & 0 & M & A & E \\ 0 & 0 & 0 & C & 0 \end{bmatrix} - \operatorname{rank} \begin{bmatrix} M & A & E \\ 0 & E & 0 \\ 0 & C & 0 \end{bmatrix} = \operatorname{rank} \begin{bmatrix} M & A \\ 0 & E \\ 0 & C \end{bmatrix},$$

F2. $\operatorname{rank}\begin{bmatrix} sE-A & M \\ C & 0 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} M & A \\ 0 & E \\ 0 & C \end{bmatrix}$,

F3. $\operatorname{rank}\begin{bmatrix} A & E \\ E & 0 \\ C & 0 \end{bmatrix} = n + r.$

By substituting (5.9)–(5.12) *into F1–F3, the following conditions are found to be required for F1 and F3 to be satisfied:*

G1. $\operatorname{rank}\begin{bmatrix} M_{11} & E_{222} \\ 0 & E_{223} \end{bmatrix} - \operatorname{rank}\begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix} = q - q_2,$

G2. $\operatorname{rank}\begin{bmatrix} E_{21} \\ E_{221} \\ E_{222} \\ E_{223} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} E_{21} - A_{41}E_{221} \\ E_{222} \\ E_{223} \end{bmatrix}.$

Additionally, Yu et al. [161] showed that by treating some components of the control input u as measurable outputs and then feeding these input and output signals into two cascaded SMOs, x_0 and f_0 can be estimated in finite time, if and only if the following conditions are satisfied:

H1. $\operatorname{rank}\begin{bmatrix} E & A & M \\ 0 & C & 0 \end{bmatrix} = n + p,$

H2. $\operatorname{rank}\begin{bmatrix} E & A & M \\ 0 & E & 0 \\ 0 & C & 0 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} E & M \end{bmatrix} + n,$

H3. $\operatorname{rank}\begin{bmatrix} E & A & 0 & M & 0 & 0 \\ 0 & E & A & 0 & M & 0 \\ 0 & 0 & E & 0 & 0 & M \\ 0 & C & 0 & 0 & 0 & 0 \\ 0 & 0 & C & 0 & 0 & 0 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} E & A & M & 0 \\ 0 & E & 0 & M \\ 0 & C & 0 & 0 \end{bmatrix} + n + q,$

H4. $\operatorname{rank}\begin{bmatrix} sE - A & M \\ C & 0 \end{bmatrix} = n + q \ \forall s \in \mathbb{C}^{+}.$

By substituting (5.9)–(5.12) *into H1–H4, the following conditions are found to be required for H1 and H2 to be satisfied:*

J1. $\operatorname{rank}(E_{223}) = 2p - \bar{n} - q_2,$

J2. $\operatorname{rank}\begin{bmatrix} E_{221} \\ E_{222} \\ E_{223} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix}.$ ♯

Conditions G1–G2 and J1–J2 present additional constraints for the estimation of x_0 and f_0. Thus, the contribution of this chapter is to improve on the works by [17, 99, 100, 101, 161] by estimating x_0 and f_0 in finite time and requiring less stringent conditions, as will be shown in Theorem 5.1.

Define an invertible matrix

$$T = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_p \\ 0 & \bar{T}_p \end{bmatrix}, \tag{5.15}$$

where $T_p \in \mathbb{R}^{p \times (n - \bar{n} + p)}$ is design freedom to be exploited. Apply the state equation transformation T onto system (5.9)–(5.12) to obtain:

$$\begin{bmatrix} 0 & I_{\bar{n}-p} & E_1 \\ 0 & 0 & T_p\bar{E}_2 \\ 0 & 0 & \bar{T}_p\bar{E}_2 \end{bmatrix}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 \\ T_p\bar{A}_4 & T_p\bar{A}_5 & T_p\bar{A}_6 \\ \bar{T}_p\bar{A}_4 & \bar{T}_p\bar{A}_5 & \bar{T}_p\bar{A}_6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} + \begin{bmatrix} B_1 \\ T_p\bar{B}_2 \\ \bar{T}_p\bar{B}_2 \end{bmatrix} u + \begin{bmatrix} M_{11} & M_{12} \\ 0 & T_p\bar{M}_2 \\ 0 & \bar{T}_p\bar{M}_2 \end{bmatrix}\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \tag{5.16}$$

$$y = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} x. \tag{5.17}$$

Then treat x_1 as an unknown input, and define

$$\bar{x} = \begin{bmatrix} x_2 \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} f \\ x_1 \end{bmatrix}. \tag{5.18}$$

System (5.16)–(5.17) can then be re-expressed as the following reduced-order system:

$$\underbrace{\begin{bmatrix} I_{\bar{n}-p} & E_1 \\ 0 & T_p\bar{E}_2 \end{bmatrix}}_{\bar{E}} \begin{bmatrix} \dot{x}_2 \\ \dot{y} \end{bmatrix} = \underbrace{\begin{bmatrix} A_2 & A_3 \\ T_p\bar{A}_5 & T_p\bar{A}_6 \end{bmatrix}}_{\bar{A}} \begin{bmatrix} x_2 \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} B_1 \\ T_p\bar{B}_2 \end{bmatrix}}_{\bar{B}} u$$

$$+ \underbrace{\left[\begin{array}{cc|c} M_{11} & M_{12} & A_1 \\ 0 & T_p\bar{M}_2 & T_p\bar{A}_4 \end{array}\right]}_{\bar{M}} \left[\begin{array}{c} f_1 \\ f_2 \\ \hline x_1 \end{array}\right], \tag{5.19}$$

$$y = \underbrace{\begin{bmatrix} 0 & I_p \end{bmatrix}}_{\bar{C}} \bar{x}. \tag{5.20}$$

Notice that $\operatorname{rank}\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$ (which is full-column rank) and the output of the reduced order system (5.19)–(5.20) is a measurable signal y. The observer by Yeu et al. [157] can therefore be designed based on $(\bar{E}, \bar{A}, \bar{M}, \bar{C})$ and driven by u and y, and applied onto system (5.19)–(5.20) to estimate $\bar{x}$ and $\bar{f}$, thereby estimating x_0 and f_0.

5.3 The two-observer scheme

Figure 5.1 shows a schematic diagram of the cascaded observer scheme proposed in this chapter.

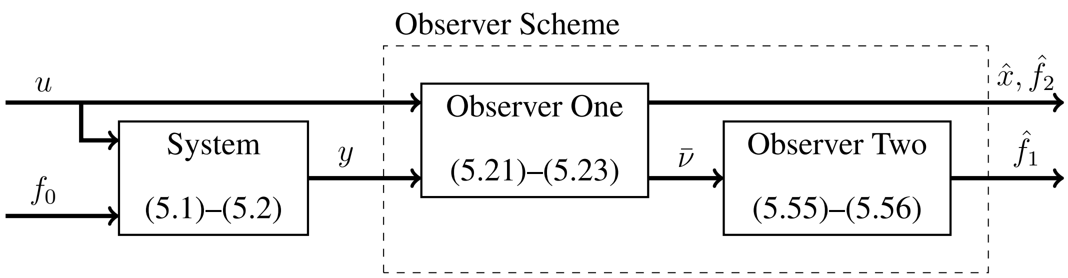

Figure 5.1: Schematic diagram of the scheme proposed in the chapter.

5.3.1 Observer One

Observer One for system (5.19)–(5.20) has the following structure [157]:

$$\dot{z} = \left(R\bar{A} - G_l\bar{C}\right)z - R\bar{B}u - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu, \tag{5.21}$$

$$\hat{\bar{x}} = Vy - z, \tag{5.22}$$

$$\nu = -\rho\frac{e_y}{\|e_y\|},\ e_y = \bar{C}\hat{\bar{x}} - y,\ \hat{\bar{x}} = \begin{bmatrix}\hat{x}_{2,0}\\ \hat{y}\end{bmatrix} \tag{5.23}$$

where $R \in \mathbb{R}^{\bar{n}\times\bar{n}}$ is invertible, and $V \in \mathbb{R}^{\bar{n}\times p}$. The gain matrices G_l and G_n are chosen such that

$$G_l = \begin{bmatrix}A_3 + R_2T_p\bar{A}_6\\ J_1 + R_4T_p\bar{A}_6\end{bmatrix},\ G_n = \begin{bmatrix}0\\ G_{n2}\end{bmatrix}, \tag{5.24}$$

where $J_1 \in \mathbb{R}^{p\times p} > 0$ and $G_{n2} = G_{n2}^T > 0$. Pre-multiply (5.19) with R and add $V\dot{y}$ to both sides to obtain

$$R\bar{E}\dot{\bar{x}} + V\dot{y} = \left(R\bar{E} + V\bar{C}\right)\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + V\dot{y}. \tag{5.25}$$

Next, suppose $R\bar{E} + V\bar{C} = I_{\bar{n}}$. Equation (5.25) becomes

$$\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + V\bar{C}\dot{\bar{x}}. \tag{5.26}$$

Corollary 5.1 *The matrices R, V from* (5.21)–(5.22) *will have the following structures:*

$$R = \begin{bmatrix}I_{\bar{n}-p} & R_2\\ 0 & R_4\end{bmatrix},\ V = \begin{bmatrix}-(\bar{E}_1 + R_2T_p\bar{E}_2)\\ I_p - R_4T_p\bar{E}_2\end{bmatrix}, \tag{5.27}$$

where $|R_4| \neq 0$. ♯

Proof Since $\operatorname{rank}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = \bar{n}$, then the matrices R and V can be chosen such that

$$\begin{bmatrix}R & V\end{bmatrix}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = I_{\bar{n}}, \tag{5.28}$$

that is, $\begin{bmatrix}R & V\end{bmatrix}$ is chosen to be the Moore-Penrose inverse of $\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix}$. Partition the matrices R and V generally as follows:

$$\begin{bmatrix}R & V\end{bmatrix} = \left[\begin{array}{cc|c}R_1 & R_2 & V_1\\ R_3 & R_4 & V_2\end{array}\right], \tag{5.29}$$

where $R_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$ and $R_4 \in \mathbb{R}^{p\times p}$. By substituting $\bar{E}$ and $\bar{C}$ from (5.19)–(5.20) and R and V from (5.29) into (5.28), it can be seen that R and V would take the forms given in (5.27). ■

Then substitute z from (5.22) into (5.21) to obtain

$$\begin{aligned}\left(V\dot{y} - \dot{\hat{\bar{x}}}\right) &= \left(R\bar{A} - G_l\bar{C}\right)\left(Vy - \hat{\bar{x}}\right) - R\bar{B}u - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu \\ \dot{\hat{\bar{x}}} &= \left(R\bar{A} - G_l\bar{C}\right)\hat{\bar{x}} + R\bar{B}u + G_l\bar{C}\bar{x} + G_n\nu + V\bar{C}\dot{\bar{x}}.\end{aligned} \tag{5.30}$$

Define the state estimation error for Observer One as follows:

$$e = \hat{\bar{x}} - \bar{x} = \begin{bmatrix} e_2 \\ e_y \end{bmatrix} \begin{matrix} \updownarrow & \bar{n}-p \\ \updownarrow & p \end{matrix}. \tag{5.31}$$

Therefore, by subtracting (5.26) from (5.30), the error equation for Observer One (which characterises its performance) is given by:

$$\dot{e} = \left(R\bar{A} - G_l\bar{C}\right)e - R\bar{M}\bar{f} + G_n\nu. \tag{5.32}$$

Substitute G_l and G_n from (5.24) into (5.32), and partition according to (5.19)–(5.20) to obtain

$$\begin{aligned}\begin{bmatrix} \dot{e}_2 \\ \dot{e}_y \end{bmatrix} &= \begin{bmatrix} A_2 + R_2T_p\bar{A}_5 & 0 \\ R_4T_p\bar{A}_5 & -J_1 \end{bmatrix}\begin{bmatrix} e_2 \\ e_y \end{bmatrix} + \begin{bmatrix} 0 \\ G_{n2} \end{bmatrix}\nu \\ &\quad - \begin{bmatrix} M_{11} & M_{12} + R_2T_p\bar{M}_2 & A_1 + R_2T_p\bar{A}_4 \\ 0 & R_4T_p\bar{M}_2 & R_4T_p\bar{A}_4 \end{bmatrix}\bar{f}.\end{aligned} \tag{5.33}$$

Proposition 5.2 *Suppose there exists a matrix* $P = P^T > 0$ *that satisfies*

$$P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0, \tag{5.34}$$

where $P = \begin{bmatrix} P_1 & 0 \\ 0 & G_{n2}^{-1} \end{bmatrix}$, $P_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$. *If* ρ *in* (5.23) *is chosen as follows:*

$$\begin{aligned}\rho &\geq \left(\tfrac{2\|G_{n2}^{-1}R_4T_p\bar{A}_5\|\|PR\bar{M}\|}{\mu} + \|G_{n2}^{-1}R_4T_p\begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\|\right)\alpha_1 + \eta, \\ \mu &= -\lambda_{\max}\left(P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P\right) > 0, \\ \alpha_1 &> \|f\|_{\max} + \|x_1\|_{\max},\ \eta > 0,\end{aligned} \tag{5.35}$$

then an ideal sliding motion $(e_y, \dot{e}_y = 0)$ *takes place in finite time.* ♯

Proof The proof of convergence will be an adaptation of the one in [128], and will consist of two parts: the first part would show that by setting ρ appropriately, e can be made to be ultimately bounded. Define a candidate Lyapunov function $W = e^T Pe > 0$. Differentiating W with respect to time yields

$$\begin{aligned}\dot{W} &= e^T \left(P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P\right)e - 2e^T PR\bar{M}\bar{f} + 2e^T PG_n\nu \\ &\leq -\mu\|e\|^2 + 2\|e\|\|PR\bar{M}\|\alpha_1 - 2\rho\|e_y\|. \end{aligned} \tag{5.36}$$

By setting $\rho \geq \eta$, (5.36) becomes

$$\dot{W} \leq \|e\|\left(-\mu\|e\| + 2\|PR\bar{M}\|\alpha_1\right). \tag{5.37}$$

When the magnitude of the error e is smaller than or equal to a certain bound, i.e. $\|e\| \leq \frac{2\|PR\bar{M}\|\alpha_1}{\mu}$, then (5.37) becomes

$$\dot{W} \leq \kappa, \tag{5.38}$$

where $\kappa \geq 0$, and thus the magnitude of e can increase, decrease, or remain unchanged within this bound. If, however, the magnitude of the error e is larger than that bound, i.e. $\|e\| \geq \frac{2\|PR\bar{M}\|\alpha_1}{\mu}$, then (5.37) becomes

$$\dot{W} < 0, \tag{5.39}$$

and $\|e\|$ will shrink. This implies that the magnitude of e would be bounded $\left(\|e\| \leq \frac{2\|PR\bar{M}\|\alpha_1}{\mu}\right)$ in finite time if ρ is set such that $\rho \geq \eta$.

The next (and remaining) part of the proof would show how sliding motion $(e_y, \dot{e}_y = 0)$ is induced. Define another candidate Lyapunov function $W_y = e_y^T G_{n2}^{-1} e_y > 0$, where

$$G_{n2}^{-1} J_1 + J_1^T G_{n2}^{-1} > 0. \tag{5.40}$$

Differentiating W_y with respect to time yields

$$\begin{aligned}\dot{W}_y &= -e_y^T\left(G_{n2}^{-1}J_1 + J_1^T G_{n2}^{-1}\right)e_y + 2e_y^T G_{n2}^{-1} R_4 T_p\left(\bar{A}_5 e_2 - \begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\bar{f}\right) + 2e_y^T\nu \\ &\leq -2\|e_y\|\left(\rho - \left(\frac{2\|G_{n2}^{-1}R_4T_p\bar{A}_5\|\|PR\bar{M}\|}{\mu} + \|G_{n2}^{-1}R_4T_p\begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\|\right)\alpha_1\right). \end{aligned} \tag{5.41}$$

Notice that

$$\|e_y\|^2 = \left(\sqrt{G_{n2}^{-1}}e_y\right)^T G_{n2}\left(\sqrt{G_{n2}^{-1}}e_y\right) \geq \lambda_{\min}\left(G_{n2}\right)\|\sqrt{G_{n2}^{-1}}e_y\|^2 = \lambda_{\min}\left(G_{n2}\right)W_y. \quad (5.42)$$

Then by setting ρ such that (5.35) is satisfied (which also satisfies $\rho \geq \eta$), (5.41) becomes

$$\dot{W}_y \leq -2\eta\|e_y\| \leq -2\eta\sqrt{\lambda_{\min}\left(G_{n2}\right)}\sqrt{W_y}, \quad (5.43)$$

which is the *reachability condition* [128], resulting in $e_y = 0$ in finite time and a sliding motion taking place on the surface $\mathcal{S}_1 = \{e : \bar{C}e = 0\}$, thus proving the proposition. ■

After sliding motion is achieved, (5.33) becomes

$$\dot{e}_2 = \left(A_2 + R_2T_p\bar{A}_5\right)e_2 - \begin{bmatrix} M_{11} & M_{12} + R_2T_p\bar{M}_2 & A_1 + R_2T_p\bar{A}_4 \end{bmatrix}\bar{f}, \quad (5.44)$$

$$0 = R_4T_p\left(\bar{A}_5e_2 - \begin{bmatrix} 0 & \bar{M}_2 & \bar{A}_4 \end{bmatrix}\bar{f}\right) + G_{n2}\nu. \quad (5.45)$$

5.3.1.1 Estimating x and f_2 in finite time

Let T_p from (5.15) and R_2 take the structure

$$T_p = \begin{bmatrix} I_{q_2} & -A_{41} & -A_{51} & T_{14} \\ 0 & I_{n-\bar{n}} & 0 & T_{24} \\ 0 & 0 & -I_{\bar{n}-p} & T_{34} \\ 0 & 0 & 0 & T_{44} \end{bmatrix}, \; R_2 = \begin{bmatrix} -M_{21} & -A_1 & A_2 + J_2 & R_{24} \end{bmatrix}, \quad (5.46)$$

where $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$ is chosen such that T_p is full-row rank, and $J_2 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)} > 0$. Define a measurable signal

$$\bar{\nu} = R_4^{-1}G_{n2}\nu. \quad (5.47)$$

Substitute for T_p and R_2 from (5.46) into (5.44)–(5.45), and rearrange the terms to obtain

$$\dot{e}_2 = -J_2e_2 - \begin{bmatrix} M_{11} & 0 & 0 \end{bmatrix}\bar{f}, \quad (5.48)$$

$$\bar{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I_{\bar{n}-p} \\ 0 \end{bmatrix} e_2 + \begin{bmatrix} 0 & I_{q_2} & 0 \\ 0 & 0 & I_{n-\bar{n}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\bar{f} \; \begin{array}{l} \updownarrow \; q_2 \\ \updownarrow \; n-\bar{n} \\ \updownarrow \; \bar{n}-p \\ \updownarrow \; 2p-n-q_2 \end{array}. \quad (5.49)$$

Define the following measurable signals:

$$\hat{x}_1 = \nu_2, \ \hat{x}_2 = \hat{x}_{2,0} - \nu_3, \ \hat{f}_2 = \nu_1. \tag{5.50}$$

From (5.49) and the structure of $\bar{f}$ in (5.18), it is obvious that

$$\nu_1 = f_2, \ \nu_2 = x_1, \ \nu_3 = e_2. \tag{5.51}$$

Hence the estimates for $x_1, \ x_2$ and f_2 can be obtained from the signals in (5.50) as

$$\hat{x}_1 = x_1, \ \hat{x}_2 = \hat{x}_{2,0} - e_2 = x_2, \ \hat{f}_2 = f_2. \tag{5.52}$$

Furthermore, recall from Proposition 5.2 that (5.44)–(5.45) (and hence (5.48)–(5.49) as well) and $\hat{y} = y$ occur in finite time; therefore x and f_2 are estimated in finite time.

Remark 5.2 *Notice that only f_2 can be estimated (in finite time) from Observer One, while f_1 cannot be obtained from the measurable signals in* (5.49)*. This limitation is also found in other schemes utilising a single SMO [17, 100, 101]. Hence a second observer is required to estimate the remaining fault component f_1.* ♯

5.3.2 Observer Two

To reconstruct f_1, a system based on measurable signals from (5.48)–(5.49) first needs to be formulated. By rearranging (5.48)–(5.49), the following system with measurable output ν_3 is formed:

$$\dot{e}_2 = -J_2 e_2 - M_{11} f_1, \tag{5.53}$$

$$\nu_3 = e_2. \tag{5.54}$$

Notice that (5.53)–(5.54) is a state-space system, and therefore the Edwards-Spurgeon SMO can be designed based on the triple $(-J_2, -M_{11}, I_{\bar{n}-p})$ and driven by ν_3 to estimate f_1. Observer Two for system (5.53)–(5.54) has the following structure [8]:

$$\dot{\hat{e}}_2 = -J_2 \hat{e}_2 - \tilde{G}_l \tilde{e} + \tilde{G}_n \tilde{\nu}, \tag{5.55}$$

$$\tilde{\nu} = -\tilde{\rho} \frac{\tilde{e}}{\|\tilde{e}\|}, \tag{5.56}$$

where $\tilde{e} = \hat{e}_2 - \nu_3$, $\tilde{\rho}$ will be defined later in (5.58), and the gain matrices $\tilde{G}_l, \tilde{G}_n$ are chosen such that

$$\lambda\left(J_2 + \tilde{G}_l\right) > 0,\ \tilde{G}_n = \tilde{G}_n^T > 0. \tag{5.57}$$

Notice that all states of system (5.53)–(5.54) are measurable, and hence observer (5.55)–(5.56) is a specific case of the Edwards-Spurgeon SMO whereby the number of states is the same as the number of outputs.

Remark 5.3 *Note that the purpose of Observer Two is not to estimate e_2 (which is already estimated by Observer One). Rather, it is designed and implemented based on system* (5.53)–(5.54) *to estimate f_1 which is the unknown input.* ♯

Proposition 5.3 *By setting $\tilde{\rho}$ such that*

$$\tilde{\rho} \geq \|\tilde{G}_n^{-1} M_{11}\|\alpha_2 + \zeta,\ \alpha_2 \geq \|f_1\|_{\max},\ \zeta > 0, \tag{5.58}$$

then sliding motion $(\tilde{e}, \dot{\tilde{e}} = 0)$ occurs in finite time. ♯

Proof By subtracting (5.53) from (5.55), the error equation for Observer Two is

$$\dot{\tilde{e}} = -\left(J_2 + \tilde{G}_l\right)\tilde{e} + M_{11} f_1 + \tilde{G}_n \tilde{\nu}. \tag{5.59}$$

Define a candidate Lyapunov function $\tilde{W} = \tilde{e}^T \tilde{G}_n^{-1} \tilde{e} > 0$, where

$$\tilde{G}_n^{-1}\left(J_2 + \tilde{G}_l\right) + \left(J_2 + \tilde{G}_l\right)^T \tilde{G}_n^{-1} > 0. \tag{5.60}$$

Differentiating $\tilde{W}$ with respect to time gives

$$\begin{aligned}\dot{\tilde{W}} &= -\tilde{e}^T\left(\tilde{G}_n^{-1}\tilde{A} + \tilde{A}^T\tilde{G}_n^{-1}\right)\tilde{e} + 2\tilde{e}^T\tilde{G}_n^{-1}M_{11}f_1 + +2\tilde{e}^T\tilde{\nu} \\ &\leq -2\|\tilde{e}\|\left(\tilde{\rho} - \|\tilde{G}_n^{-1}M_{11}\|\alpha_2\right).\end{aligned} \tag{5.61}$$

Note that by using similar arguments as in [128], the following is obtained

$$\|\tilde{e}\|^2 = \left(\sqrt{\tilde{G}_n^{-1}}\tilde{e}\right)^T \tilde{G}_n \left(\sqrt{\tilde{G}_n^{-1}}\tilde{e}\right) \geq \lambda_{\min}\left(\tilde{G}_n\right)\|\sqrt{\tilde{G}_n^{-1}}\tilde{e}\|^2 = \lambda_{\min}\left(\tilde{G}_n\right)\tilde{W}. \tag{5.62}$$

By setting $\tilde{\rho}$ such that (5.58) is satisfied, (5.61) becomes

$$\dot{\tilde{W}} \leq -2\zeta\|\tilde{e}\| \leq -2\zeta\sqrt{\lambda_{\min}\left(\tilde{G}_n\right)}\sqrt{\tilde{W}}, \tag{5.63}$$

which is the reachability condition. This condition will then result in $\tilde{e} = 0$ and a sliding motion on the surface $\mathcal{S}_2 = \{\tilde{e} : \tilde{e} = 0\}$ occurring in finite time, thus proving the proposition. ∎

After sliding motion is achieved, (5.59) becomes

$$M_{11} f_1 + \tilde{G}_n \tilde{\nu} = 0. \tag{5.64}$$

Recall the structure of M in (5.9); since M is full-column rank, M_{11} is also full-column rank. Hence $M_{11}^{\dagger}$ (which is the Moore-Penrose inverse of M_{11}) exists, where $M_{11}^{\dagger} M_{11} = I_{q-q_2}$. Therefore, define the following signal as the estimate of f_1:

$$\hat{f}_1 = -M_{11}^{\dagger} \tilde{G}_n \tilde{\nu}. \tag{5.65}$$

Recall from Proposition 5.3 that (5.65) occurs in finite time, and hence f_1 is estimated in finite time.

5.3.3 Existence conditions

Theorem 5.1 *The proposed cascaded SMO scheme can estimate x_0 and f_0 in finite time for system* (5.1)–(5.2) *if and only if the following conditions hold:*

K1. $\mathrm{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \mathrm{rank}(M_2) = n - p$,

K2. $p + q - \bar{n} \leq \mathrm{rank}(M_2) \leq 2p - n$. ♯

Proof The remainder of this subsection forms the constructive proof for Theorem 5.1.

Remark 5.4 *Note that the scheme does not directly require minimum phase restrictions. This is because K1–K2 already imply (as will be shown in Corollary 5.2) that system* (5.53)–(5.54) *is minimum phase, and that error system* (5.44)–(5.45) *can always be made stable by choice of R_2 in* (5.46). ♯

In the formulation of the cascaded SMO scheme (in the preceding subsections), the following assumptions were made:

L1. $\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p$ in (5.8), so T_{f1} in (5.13) exists such that $T_{f1}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}$,

L2. Matrix T_{44} in (5.46) exists such that it has the dimensions $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$,

L3. $\mathrm{rank}(M_{11}) = q - q_2$, so that $M_{11}^{\dagger}$ in (5.65) exists such that $M_{11}^{\dagger}M_{11} = I_{q-q_2}$,

L4. The error systems (5.44)–(5.45) (and therefore (5.48)–(5.49)) and (5.59) can be made stable.

It is thus of interest to re-express these conditions in terms of the original system matrices so that it is easier for the designer to check whether the proposed scheme is applicable from the outset.

Proposition 5.4 *Condition K1 is necessary and sufficient for L1 to be satisfied.* ♯

Proof Apply the transformations T_a, T_b, T_c, T_d, T_e in (5.3)–(5.7) to obtain

$$M_2 = \begin{bmatrix} 0 & I_{q_2} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} A_4 & A_5 \end{bmatrix} = \begin{bmatrix} A_{41} & A_{51} \\ A_{42} & A_{52} \end{bmatrix}. \tag{5.66}$$

Suppose K1 is not satisfied, i.e.

$$\mathrm{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \mathrm{rank}(M_2) < n - p. \tag{5.67}$$

Equation (5.67) implies

$$\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} < n - p, \tag{5.68}$$

that is, L1 is not satisfied, which shows the necessity of K1. On the other hand, if K1 is satisfied, then from the structures of M_2 and $\begin{bmatrix} A_4 & A_5 \end{bmatrix}$ from (5.66), it can easily be deduced that

$$\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p, \tag{5.69}$$

that is, L1 is satisfied, thus proving the sufficiency of K1. ■

Proposition 5.5 *Conditions L2 and L3 are satisfied if and only if K2 is satisfied.* ♯

Proof The necessity of K2 is shown through the two cases where K2 is not satisfied, i.e.

$$\mathrm{rank}(M_2) > 2p - n, \tag{5.70}$$

$$\mathrm{rank}(M_2) < p + q - \bar{n} \tag{5.71}$$

The matrix T_{44} is assigned the dimensions $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$ in (5.46), where $q_2 = \mathrm{rank}(M_2)$. Equation (5.70) would imply T_{44} does not exist, and thus L2 is not satisfied. Next, (5.71) implies

$$\bar{n} - p < q - q_2. \tag{5.72}$$

Since $M_{11} \in \mathbb{R}^{(\bar{n}-p)\times(q-q_2)}$,

$$\mathrm{rank}(M_{11}) = \min\{\bar{n} - p, q - q_2\} \leq \bar{n} - p < q - q_2, \tag{5.73}$$

that is, M_{11} is not full-column rank, and L3 can never be satisfied, thus showing the necessity of K2. On the other hand, if K2 is satisfied, since $n - \bar{n} > 0$ it is straightforward to see that T_{44} can be assigned the dimensions in (5.46) and M_{11} would be full-column rank, satisfying L2–L3 and showing the sufficiency of K2. ■

Corollary 5.2 *Condition L4 is always satisfied.* ♯

Proof For error system (5.44)–(5.45) (and therefore (5.48)–(5.49)) to be stable, it is required that there exists a matrix $J_2 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)} > 0$ such that

$$\lambda\left(A_2 + R_2 T_p \bar{A}_5\right) = \lambda\left(-J_2\right) < 0. \tag{5.74}$$

If $\bar{n} = p$ (all states of system (5.19)–(5.20) are measurable outputs), then e_2 (and therefore J_2) does not exist. If $\bar{n} > p$, then J_2 exists and can be freely chosen such that $\lambda\left(-J_2\right) < 0$, i.e. error (5.44) (and (5.48)) is stable.

To stabilise error (5.59), the zeros of $(-J_2, -M_{11}, I_{\bar{n}-p})$ need to be stable. Define the following matrix $D(s)$:

$$D(s) = \begin{bmatrix} sI + J_2 & -M_{11} \\ I_{\bar{n}-p} & 0 \end{bmatrix}, \tag{5.75}$$

where $D(s)$ is the Rosenbrock matrix for $(-J_2, -M_{11}, I_{\bar{n}-p})$ [109], and the values of s that make $D(s)$ lose rank are zeros of $(-J_2, -M_{11}, I_{\bar{n}-p})$. Recall however, from the satisfaction of K2,

$$\mathrm{rank}(M_{11}) = q - q_2 \leq \bar{n} - p, \tag{5.76}$$

that is, M_{11} is always full-column rank. It is thus straightforward from (5.75) that $D(s)$ will always be full-column rank, and therefore $(-J_2, -M_{11}, I_{\bar{n}-p})$ has no zeros, and error (5.59) can always be stabilised. Thus L4 is shown to always be satisfied. ■

Remark 5.5 *Propositions 5.4 and 5.5 re-express the assumptions L1–L3 as conditions K1–K2 (which are necessary and sufficient for the feasibility of the scheme proposed in this chapter), that are expressed in terms of the original system matrices (which greatly helps the designer know at the outset whether the scheme is applicable to a particular system or not). Then Corollary 5.2 shows that the observers can always be designed to be stable (i.e. L4 is always satisfied) if K1–K2 are satisfied.*

Remark 5.6 *The existence conditions in this chapter can be summarised as follows:*

- *Conditions F1–F3 are the existence conditions for the scheme by Ooi et al. [99]. Likewise, G1–G4 are required for the scheme by Yu et al. [161] to be applicable.*

- *Conditions L1–L4 represent the assumptions made in designing the scheme proposed in this chapter. These assumptions are required for the transformations in §5.2 to exist, and for the scheme to be applicable.*

- *After applying the transformations in §5.2 (i.e. assuming L1–L4 are valid), the schemes by Ooi et al. [99] and Yu et al. [161] still require the satisfaction of G1–G2 and J1–J2 to be applicable, respectively. Thus the scheme proposed in this chapter is more relaxed than those by Ooi et al. and Yu et al.*

- *Finally, K1–K2 are conditions in terms of the original system matrices that are shown to be equivalent to L1–L4. Conditions K1–K2 allow system designers to quickly verify if the scheme proposed in this chapter is applicable to their system.*

Thus, Theorem 5.1 is proven. □

The following summarises the strengths of the proposed scheme over previous works:

- The proposed scheme is able to estimate all states and faults for a class of non-infinitely observable descriptor systems where $\mathrm{rank}(M_2) \neq \mathrm{rank}(M)$. This is not possible in schemes utilising a single SMO [17, 100, 101], as shown in Remark 5.2.
- The convergence of the estimates occur in finite time, unlike in [161] where the estimates converge only asymptotically.
- The existence conditions of the scheme are found to be more relaxed than those in existing cascaded SMO schemes [99, 161], as shown in Remark 5.1. Condition K2 only limits the number of faults that can be estimated by the proposed scheme. This is in contrast to G1–G2 and J1–J2 required by existing schemes, which form stringent rank conditions on matrix E.

5.3.4 Design procedure

The design procedure for the proposed cascaded observer scheme can be summarised as follows:

1. Determine T_a, T_b, T_c in (5.3)–(5.5), and apply the state equation transformation T_b and state transformation T_cT_a.
2. Check that K1–K2 hold for the system. If they do are not satisfied, do not continue as the observer scheme is not applicable.
3. Determine T_d, T_e, T_f from (5.7)–(5.13), and apply the state equation transformation T_fT_d and the fault transformation T_e.
4. Set values for $T_{14}, T_{24}, T_{34}, T_{44}$ in (5.46). Apply the state equation transformation T given in (5.15).
5. Form the reduced-order system in (5.19)–(5.20).
6. Set R_2 from (5.46), choose R_4 to be an invertible symmetric positive-definite matrix. If $\bar{n} > p$, choose a value for J_2 such that $J_2 > 0$.
7. Calculate G_l and G_n from (5.24), and V from (5.27).

8. Form the reduced-order system in (5.53)–(5.54).

9. Set $\tilde{G}_n$ as an invertible symmetric positive-definite matrix, and choose $\tilde{G}_l$ such that $\lambda\left(J_2 + \tilde{G}_l\right) > 0$.

10. Choose ρ and $\tilde{\rho}$ to satisfy (5.35) and (5.58), respectively. If the bounds α_1 and α_2 are difficult to determine, ρ and $\tilde{\rho}$ can be designed to be adaptive according to the method outlined in [47].

5.4 Simulation example

To demonstrate the effectiveness of the scheme in this chapter, consider a modified version of the chemical mixing tank given in [157]. Note that the example is deliberately set to be similar to that in Chapter 4 to ease the comparison between the schemes presented in Chapter 4 and this chapter. The dynamics of the mixing tank are described by the following dynamical model:

$$\dot{c}_3 = -0.375c_3 - 0.0667q_3 + 0.1q_1, \tag{5.77}$$

$$0 = -q_3 + q_1, \tag{5.78}$$

$$\dot{c}_5 = 0.3c_3 + 0.0533q_3 - 0.5c_5 - 0.04q_5 + 0.02q_4, \tag{5.79}$$

$$0 = q_3 - q_5 + q_4, \tag{5.80}$$

where q_1 is the flow rate of the influent into the first tank, c_3 and q_3 are the concentration and flow rate of the influent from the first tank into the second tank, respectively, q_4 is the flow rate of influent from another pipe into the second tank, and c_5 and q_5 are the concentration and flow rate of the effluent from the second tank, respectively. The actuator signal for q_1 is generated by a faulty first-order device

$$\dot{q}_1 = -2q_1 + 2q_{1,e} + f_a, \tag{5.81}$$

where $q_{1,e}$ is the reference signal, and f_a is the fault. Furthermore, the input chemical concentration c_3 is known to fluctuate about its equilibrium value, and hence its dynamical equation becomes:

$$\dot{c}_3 = -0.375c_3 - 0.0667q_3 + 0.1q_1 - f_b, \tag{5.82}$$

where f_b is the fault representing the fluctuations.

5.4.1 System formulation

The system matrices (E, A, B, M) in the framework of (5.1)–(5.2) are

$$
\begin{gathered}
E = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; A = \begin{bmatrix} -0.375 & -0.0667 & 0 & 0 & 0.1 \\ 0 & -1 & 0 & 0 & 1 \\ 0.3 & 0.0533 & -0.5 & -0.04 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}, \\
B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0.02 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}, \; M = \begin{bmatrix} 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix},
\end{gathered} \tag{5.83}
$$

for the system variables

$$
\begin{gathered}
x = \begin{bmatrix} \text{concentration, } c_3 \text{ (mol/l)} \\ \text{flow rate, } q_3 \text{ (l/s)} \\ \text{concentration, } c_5 \text{ (mol/l)} \\ \text{flow rate, } q_5 \text{ (l/s)} \\ \text{flow rate, } q_1 \text{ (l/s)} \end{bmatrix}, \; f = \begin{bmatrix} f_a \text{ (l/s)} \\ f_b \text{ (mol/l.s)} \end{bmatrix}, \\
u = \begin{bmatrix} \text{flow rate reference, } q_{1,e} \text{ (l/s)} \\ \text{flow rate, } q_4 \text{ (l/s)} \end{bmatrix}.
\end{gathered} \tag{5.84}
$$

It can be verified that the system is regular. Assume that measurements are only available for $q_3,\ c_5$, and q_1, so C has the form

$$
C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.85}
$$

5.4.2 Design of observers

To ease readability, the steps in §5.3.4 will be referred to in the following design of the cascaded observer scheme.

Step 1: To obtain the system in the form of (5.9)–(5.12), apply the state equation transformation T_b and the state transformation T_cT_a, where

$$T_b = I_5,\ T_cT_a = \begin{bmatrix} 0 & -1 & 0 \\ I_3 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.86}$$

Step 2: The following partitions are then obtained:

$$M_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix},\ A_4 = \begin{bmatrix} 0 \\ 0.04 \\ 1 \\ 0 \end{bmatrix},\ A_5 = \begin{bmatrix} 0 \\ 0.3 \\ 0 \\ 0 \end{bmatrix}. \tag{5.87}$$

Equation (5.87) allows the following to be verified:

$$\operatorname{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \operatorname{rank}(M_2) = 2,\ p + q - \bar{n} \le \operatorname{rank}(M_2) \le 2p - n, \tag{5.88}$$

that is, K1–K2 hold for the system, thereby guaranteeing the existence of the proposed observer scheme. The scheme proposed in the chapter is now designed.

Step 3: The state equation transformation T_fT_d and the fault transformation T_e were calculated to be

$$T_fT_d = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3.333 & -0.1333 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{bmatrix},\ T_e = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{5.89}$$

Step 4: The parameters for T_p in (5.46) were set as $T_{14} = T_{24} = T_{34} = 1$.

Step 5: The reduced-order system (5.19)–(5.20) is found to be

$$\bar{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3.333 & 0 \end{bmatrix},\ \bar{A} = \begin{bmatrix} -0.375 & -0.0667 & 0 & 0.1 \\ 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & -1 \\ -1 & 0.9557 & 1.667 & -1 \end{bmatrix},$$
$$\bar{B} = \begin{bmatrix} 0 & 0 \\ -2 & 0 \\ 0 & 1 \\ 0 & 0.0667 \end{bmatrix},\ \bar{M} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ \bar{C} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.90}$$

Steps 6–7: The poles of Observer One were chosen to be $\{-10, -10, -10, -10\}$, while the pole of the sliding motion was chosen as $\{-10\}$. The parameters for Observer One are then calculated to be

$$R = \begin{bmatrix} 1 & 0 & 0 & 9.625 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ V = \begin{bmatrix} 0 & 32.08 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 3.333 & 1 \end{bmatrix},$$
$$G_l = \begin{bmatrix} 9.131 & 16.04 & -9.525 \\ 11 & 0 & 1 \\ 2 & 10 & -1 \\ 0.9557 & 1.667 & 9 \end{bmatrix}, \ G_n = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.91}$$

Step 8: The reduced-order system (5.53)–(5.54) is found to be

$$\dot{e}_2 = -10e_2 + f_1, \ \nu_3 = e_2. \tag{5.92}$$

Step 9: The pole for Observer Two was chosen as $\{-20\}$, and hence the parameters for Observer Two are

$$\tilde{G}_l = 10, \ \tilde{G}_n = 1. \tag{5.93}$$

Step 10: The bounds $\alpha_1, \ \alpha_2$ are assumed to be unknown, and hence $\rho, \ \tilde{\rho}$ are designed to be adaptive according to the method given in [47] as follows. Define the two time-varying scalars ρ_1 and ρ_2, where

$$\dot{\rho}_1(t) = \alpha_{0,1}\Phi\left(\|e_y(t)\|\right) - \alpha_{1,1}\rho_1(t), \ \dot{\rho}_2(t) = \alpha_{0,2}\Phi\left(\|\tilde{e}(t)\|\right) - \alpha_{1,2}\rho_2(t),$$
$$\alpha_{0,1}, \alpha_{0,2}, \alpha_{1,1}, \alpha_{1,2} > 0, \ \rho_1(0) = \rho_0, \ \rho_2(0) = \tilde{\rho}_0, \tag{5.94}$$

and $\Phi(\cdot)$ represents a dead-zone function such that

$$\Phi(x) = \begin{cases} 0 & \text{if } |x| \leq \varepsilon \\ x - \varepsilon\,\mathrm{sign}(x) & \text{otherwise,} \end{cases}$$

where $\varepsilon > 0$. Thus define $\rho, \tilde{\rho}$ as

$$\rho(t) = \rho_1(t) + \delta_1, \ \tilde{\rho}(t) = \rho_2(t) + \delta_2, \delta_1, \delta_2 > 0. \tag{5.95}$$

The parameters for ρ are set as $\rho_0 = 10,\ \alpha_{0,1} = 100,\ \alpha_{1,1} = 0.1,\ \varepsilon = 0.001$, and $\delta_1 = 0.001$, whereas the parameters for $\tilde{\rho}$ are set to be$\tilde{\rho}_0 = 50,\ \alpha_{0,2} - 10,\ \alpha_{1,2} = 0.1,\ \varepsilon = 0.0001$, and $\delta_2 = 0.01$.

Remark 5.7 *The design of the observers (i.e. the scheme in this chapter) is complete. It will now be shown that previous methods for state/fault estimation in descriptor systems cannot be applied to the example in* (5.83) *and* (5.85)*, as follows:*

- *The system is not infinitely observable, i.e.* $\operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = 4 < 5$*, and therefore the schemes in [54, 55, 77, 78, 153, 157, 160] are not applicable.*

- *From* (5.87)*,* $\operatorname{rank}(M_2) = 1 < 2$*. Hence the schemes utilising a single SMO in [17, 100, 101] cannot estimate all components of* f_0*, and are thus inapplicable.*

- *In the coordinates of* (5.89)*, the following partitions are obtained:*

$$M_{11} = -1,\quad \begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix} = \left[\begin{array}{ccc} 0 & 3.333 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]. \tag{5.96}$$

  Therefore it can be shown that

$$\operatorname{rank}\begin{bmatrix} M_{11} & E_{222} \\ 0 & E_{223} \end{bmatrix} - \operatorname{rank}\begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix} = 0 < 1,\ \operatorname{rank}(E_{223}) = 0 < 1, \tag{5.97}$$

  that is, G1 and J1 are not satisfied, and so the schemes [99, 161] are also not applicable. $\sharp$

5.4.3 Simulation results

The initial conditions of the system was set at $x(0) = \{2, 1, 3, 6, 1\}$, while the initial condition of the observers was set at zero. The inputs $q_{1,e}$ and q_4 were set as step inputs with magnitudes of 10 and 5, respectively. The fault signals were simulated as

$$f_a = 0.1\sin(1.2t) + 0.6,\ f_b = 0.1\sin\left(t + \frac{2\pi}{3}\right) + 0.15. \tag{5.98}$$

Figures 5.2–5.6 show the states and their estimates, while Figures 5.7–5.8 show the fault signals and their estimates. It can be seen that the estimates converge to the real signals in finite time, proving the efficacy of the scheme.

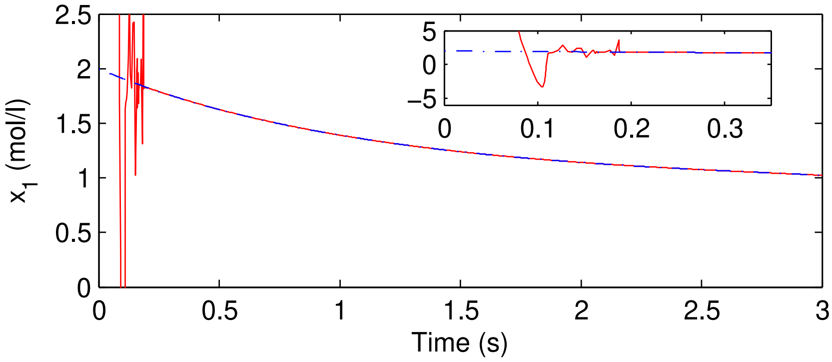

Figure 5.2: The first state (dash-dotted) and its estimate (solid).

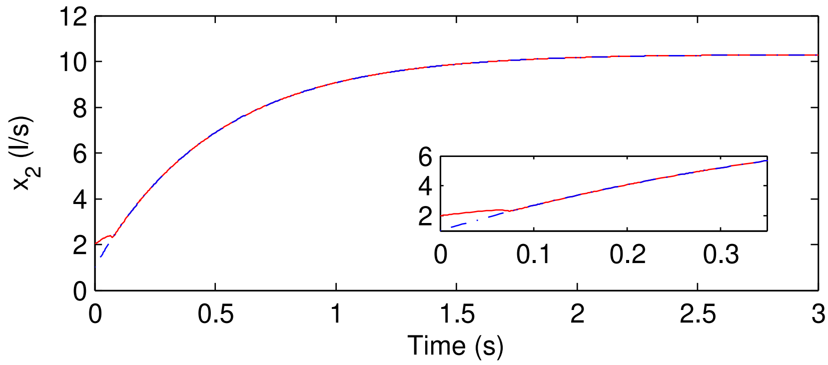

Figure 5.3: The second state (dash-dotted) and its estimate (solid).

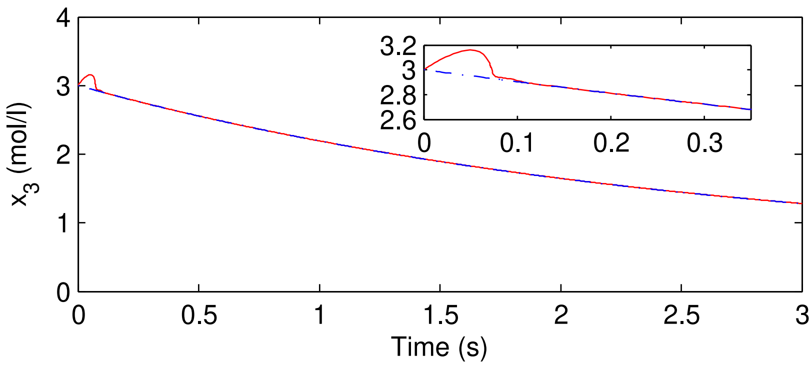

Figure 5.4: The third state (dash-dotted) and its estimate (solid).

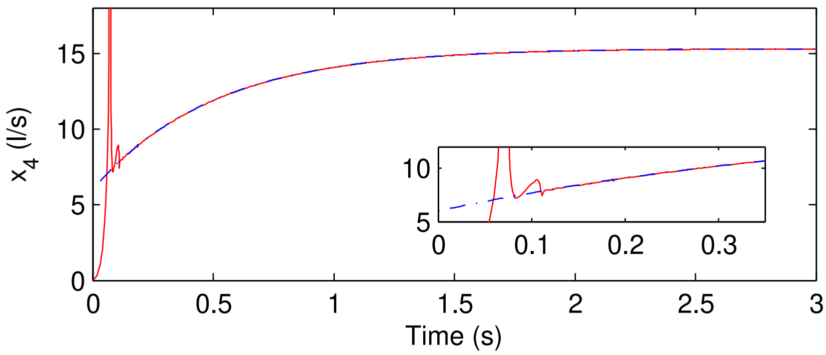

Figure 5.5: The fourth state (dash-dotted) and its estimate (solid).

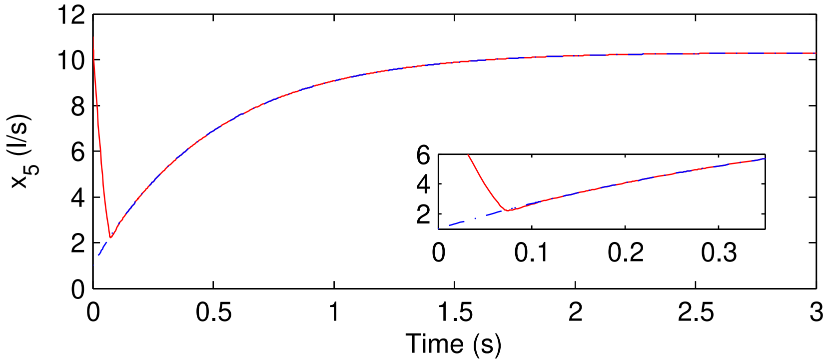

Figure 5.6: The fifth state (dash-dotted) and its estimate (solid).

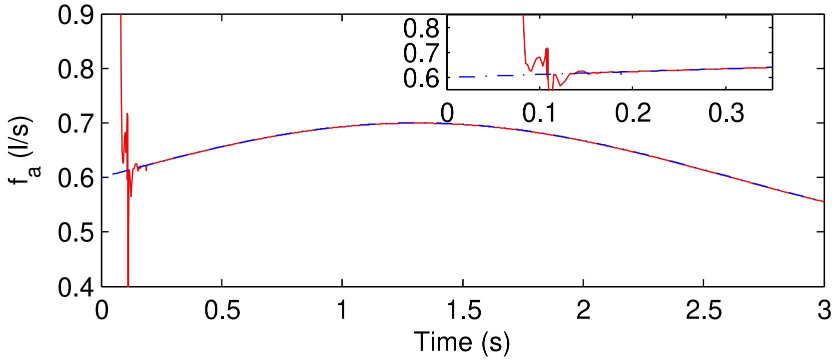

Figure 5.7: The first fault signal (dash-dotted) and its estimate (solid).

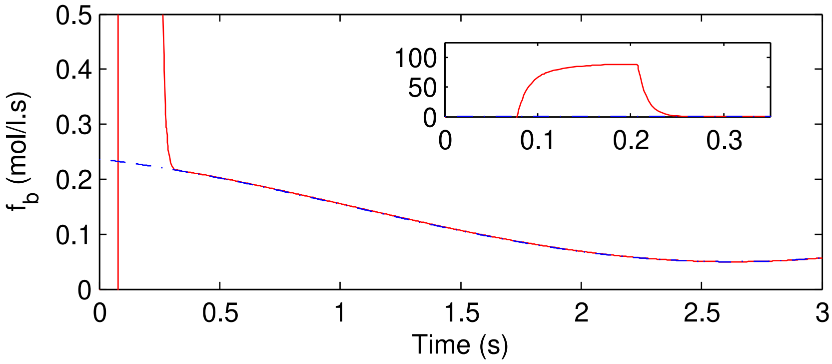

Figure 5.8: The second fault signal (dash-dotted) and its estimate (solid).

5.5 Conclusion

This chapter has presented an observer scheme consisting of two cascaded SMOs for a class of NIODS. Certain states were treated as unknown inputs to form an infinitely observable reduced-order system. The observer by Yeu et al. [157] was then applied onto this reduced-order system to estimate the states and some components of the fault in finite time. The switching term from the first SMO is found to be the output of an analytical standard (non-descriptor) state-space system which treats the remaining faults as unknown inputs. Thus the switching term is fed into a standard Edwards-Spurgeon SMO to estimate the remaining faults in finite time. The necessary and sufficient conditions for the existence of the scheme were presented in terms of the original system matrices, and are found to be more relaxed than those for existing state and fault estimation schemes for NIODS. Finally, a simulation was carried out, and the results verify the efficacy of the scheme.

Chapter 6

Robust fault reconstruction for NIODS using two SMOs in cascade

6.1 Introduction

The previous chapter introduced a scheme to estimate states and faults in a class of non-infinitely observable descriptor systems (NIODS) using two sliding mode observers (SMOs) [45, 157] in cascade. The work in that scheme however does not fully exploit the design freedom inherent in the system. Furthermore, in practical systems, there may be unmodelled dynamics or external disturbances acting upon the system, which may corrupt the fault reconstruction and raise false positives, or even mask the effects of the fault. Hence this chapter aims to build on the work in the previous chapter by developing a fault reconstruction scheme for a class of NIODS utilising two SMOs in cascade such that the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction is minimised, therefore achieving *robust fault reconstruction*.

This chapter begins by introducing the preliminary transformations to re-express the descriptor system in a form facilitating analysis and gain design in §6.2. The dynamics of the disturbances are augmented into the system. Certain states are then removed and treated as unknown inputs, formulating an infinitely observable reduced-order system. Next, in §6.3, the observer by Yeu et al. is applied onto the system to process the signals from the system and reconstruct a component of the fault. The switching signal from the observer is found to be the output of an analytical system that treats the remaining fault signals as an unknown input. A second SMO [45] is then implemented to reconstruct the remaining faults. Then in §6.3.3, linear matrix inequality (LMI) techniques and the

Bounded Real Lemma [128] are used to design the gains of the observers such that the $\mathcal{L}_2$ gain from the external disturbances onto the fault reconstruction is minimised. The existence conditions of the scheme are investigated and presented in terms of the original system matrices. Interestingly, the gains for both cascaded observers can be designed in a single LMI pair. Previous works on LMI-designed SMOs [17, 91, 128, 129] designed each SMO using a separate LMI pair, which would increase conservativeness in the solution. The LMI solution in this paper would therefore be less conservative. A set of design procedures is outlined. Finally, a numerical example is presented in §6.4 to demonstrate the effectiveness of the observer scheme.

The work in this chapter has been published; its reference is J. C. L. Chan, C. P. Tan, H. Trinh, M. A. S. Kamal, and Y. S. Chiew, Robust fault reconstruction for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade,' *Applied Mathematics and Computation*, 350:78–92, 2019.

6.2 Preliminary transformations

Consider the following non-infinitely observable descriptor system (NIODS):

$$E_0\dot{x}_0 = A_0x_0 + B_0u + M_0f_0 + Q_0\xi_0, \tag{6.1}$$

$$y = C_0x_0, \tag{6.2}$$

where $E_0, A_0 \in \mathbb{R}^{n\times n}$, $B_0 \in \mathbb{R}^{n\times m}$, $M_0 \in \mathbb{R}^{n\times q}$, $Q_0 \in \mathbb{R}^{n\times h}$, $C_0 \in \mathbb{R}^{p\times n}$, whereas x_0, u, y, f_0, and ξ_0 represent the states, inputs, outputs, faults, and disturbance signals, respectively. Only u and y are measurable. The fault signal f_0 represents anomalous behaviour in the system which needs to be reconstructed to obtain information regarding its shape and magnitude so that timely and accurate maintenance can be performed. The unknown disturbance signal ξ_0 (which is not a fault, and could arise from unmodelled dynamics or deviations from the assumptions made during the modelling process [83]) may however corrupt the fault reconstruction. The erroneous reconstruction could raise false positives, or even worse, mask the effects of a fault. Therefore it is of interest to reconstruct f_0 while minimising the effect of ξ_0 on its reconstruction. It is assumed that

E_0 is rank deficient, i.e. $\text{rank}(E_0) = r < n$, and that M_0 and C_0 are full-column rank and full-row rank, respectively, i.e. $\text{rank}(M_0) = q$, $\text{rank}(C_0) = p$. Finally, ξ_0 is assumed to be piecewise continuous, and the upper bound of its bandwidth (hereafter labelled ω_c) is assumed to be known. Hence the dynamics of ξ_0 can be modelled as a first-order low-pass filter with cut-off frequency $\omega_f \gg \omega_c$ [112] as follows:

$$\dot{\xi}_0 = A_\Omega \xi_0 + B_\Omega \xi, \tag{6.3}$$

where $A_\Omega, B_\Omega \in \mathbb{R}^{h \times h}$ are known, ξ is a bounded unknown signal generating the disturbance signal ξ_0, and $\lambda(A_\Omega) < 0$. The objective is to reconstruct f_0 while rejecting the influence of ξ_0.

Remark 6.1 *Note that no 'physical' filtration of the disturbance actually occurs; filter* (6.3) *only implies that ξ_0 can be considered to be the output of a low-pass filter $G_\Omega(s) = (sI_h - A_\Omega)^{-1} B_\Omega$ driven by ξ. The choice of (A_Ω, B_Ω) is also not unique; the first-order linear filter realisation is chosen in this chapter, but other higher order filters could equally have been selected. The more important consideration is the bandwidth of the filter: if the upper bound of the bandwidth of ξ_0 is ω_c, then choose*

$$A_\Omega = -\omega_f I_h, B_\Omega = \omega_f I_h, \tag{6.4}$$

where $\omega_f \gg \omega_c$. A choice of $\frac{\omega_f}{\omega_c} \geq 10$ would be sufficient to ensure $\xi \approx \xi_0$ [129]. Hence by making the reconstruction of f robust against ξ, the reconstruction is also made to be robust against ξ_0. ♯

Similar to previous chapters, system (6.1)–(6.2) is first re-expressed such that it takes on a form that eases further analysis and gain design using Lemma 4.1 and Proposition 6.1. First, define a non-singular matrix $T_a = \begin{bmatrix} N_C^T \\ C_0 \end{bmatrix}$, where $C_0 N_C = 0$ and hence

$$E_0 \mapsto E_0 T_a^{-1} = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix}, \; C_0 \mapsto C_0 T_a^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}. \tag{6.5}$$

Since $\text{rank} \begin{bmatrix} E_0 \\ C_0 \end{bmatrix} = \bar{n}$, it follows from (6.5) that $\text{rank}(E_{a1}) = \bar{n} - p$. Using singular-value decomposition (SVD), E_{a1} can be decomposed as

$$T_b E_{a1} T_{c1} = \begin{bmatrix} 0 & I_{\bar{n}-p} \\ 0 & 0 \end{bmatrix}, \tag{6.6}$$

where T_b and T_{c1} are non-singular matrices. Define another non-singular matrix $T_c = \begin{bmatrix} T_{c1}^{-1} & 0 \\ 0 & I_p \end{bmatrix}$, and let (E_0, M_0, A_0) in the coordinates of (6.1) be transformed to attain the following structures:

$$\begin{aligned} T_b E_0 (T_c T_a)^{-1} &= \begin{bmatrix} 0 & I_{\bar{n}-p} & E_1 \\ 0 & 0 & E_2 \end{bmatrix}, \; T_b M_0 = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \\ T_b A_0 (T_c T_a)^{-1} &= \begin{bmatrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \end{bmatrix}, \end{aligned} \tag{6.7}$$

where $E_2 \in \mathbb{R}^{(n-\bar{n}+p)\times p}$, $A_1 \in \mathbb{R}^{(\bar{n}-p)\times(n-\bar{n})}$, and $A_5 \in \mathbb{R}^{(n-\bar{n}+p)\times(\bar{n}-p)}$. Define $\operatorname{rank}(M_2) = q_2 < q$, and perform SVD on M_2 such that

$$\begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{d1} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} T_e^{-1} = \begin{bmatrix} M_{11} & M_{12} \\ 0 & I_{q2} \\ 0 & 0 \end{bmatrix}, \; T_{d1} \begin{bmatrix} A_4 & A_5 \end{bmatrix} = \begin{bmatrix} A_{41} & A_{51} \\ A_{42} & A_{52} \end{bmatrix}, \tag{6.8}$$

where $A_{52} \in \mathbb{R}^{(n-\bar{n}+p-q_2)\times(\bar{n}-p)}$, and T_{d1} and T_e are non-singular matrices.

Proposition 6.1 *Assume that*

$$\operatorname{rank} \begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p, \tag{6.9}$$

that is, $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ *is full-column rank (the satisfaction of* (6.9) *will be addressed later in Proposition 6.5). Then there exists a set of transformations introduced in Lemma 4.1 for* x_0 *and* f_0 *such that system* (6.1)–(6.2) *can be rewritten as follows:*

$$\underbrace{\begin{bmatrix} 0 & I_h & 0 & 0 \\ 0 & -Q_{222} & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & 0 & \bar{E}_2 \end{bmatrix}}_{E} \dot{x} = \underbrace{\begin{bmatrix} 0 & A_\Omega & 0 & 0 \\ 0 & \bar{Q}_1 & 0 & \bar{A}_3 \\ \bar{A}_4 & \bar{Q}_2 & \bar{A}_5 & \bar{A}_6 \end{bmatrix}}_{A} x + \underbrace{\begin{bmatrix} 0 \\ \bar{B}_1 \\ \bar{B}_2 \end{bmatrix}}_{B} u + \underbrace{\begin{bmatrix} 0 & 0 \\ M_{11} & 0 \\ 0 & \bar{M}_2 \end{bmatrix}}_{M} f, \tag{6.10}$$

$$+ \underbrace{\begin{bmatrix} B_\Omega \\ 0 \\ 0 \end{bmatrix}}_{Q} \xi, \; x = \begin{bmatrix} x_1 \\ \xi_0 \\ x_2 \\ y \end{bmatrix} \begin{matrix} \updownarrow & n - \bar{n} \\ \updownarrow & h \\ \updownarrow & \bar{n} - p \\ \updownarrow & p \end{matrix}, \; f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{matrix} \updownarrow & q - q_2 \\ \updownarrow & q_2 \end{matrix},$$

$$y = \underbrace{\begin{bmatrix} 0 & I_p \end{bmatrix}}_{C} x, \tag{6.11}$$

$$\bar{E}_2 = \begin{bmatrix} \bar{E}_{21} \\ E_{221} \\ E_{222} \\ E_{223} \end{bmatrix}, \bar{A}_4 = \begin{bmatrix} 0 \\ I_{n-\bar{n}} \\ 0 \\ 0 \end{bmatrix}, \bar{Q}_2 = \begin{bmatrix} \bar{Q}_2 \\ 0 \\ 0 \\ Q_{223} \end{bmatrix}, \bar{A}_5 = \begin{bmatrix} 0 \\ 0 \\ I_{\bar{n}-p} \\ 0 \end{bmatrix}, \bar{M}_2 = \begin{array}{c} \begin{bmatrix} I_{q2} \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \;\; q_2 \\ \updownarrow \;\; n-\bar{n} \\ \updownarrow \;\; \bar{n}-p \\ \updownarrow \;\; 2p-\bar{n}-q_2 \end{array}, \tag{6.12}$$

where the partitions of $\bar{E}_2, \bar{A}_4, \bar{A}_5, \bar{M}_2, \bar{Q}_2$ have the same row dimensions, the dimensions of the partitions of E, A are conformable to the partitions of x, while the column partitions of M are conformable to the partitions of f. ♯

Proof The first portion of the proof transforms system (6.1)–(6.2) into the form in (6.16)–(6.17) which has special structures facilitating further analysis. Since $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ is full-column rank, apply QR decomposition on $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ such that

$$T_{f1} \begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}. \tag{6.13}$$

Define three non-singular matrices T_d, T_f, and T_g, where

$$T_d = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{d1} \end{bmatrix}, \; T_f = \begin{bmatrix} I_{\bar{n}-p+q_2} & 0 \\ 0 & T_{f1} \end{bmatrix},$$
$$T_g = \begin{bmatrix} I_{\bar{n}-p} & -M_{12} & -\bar{A}_1 & -\bar{A}_2 & 0 \\ 0 & I_{q2} & -A_{41} & -A_{51} & 0 \\ 0 & 0 & I_{n-\bar{n}} & 0 & 0 \\ 0 & 0 & 0 & I_{\bar{n}-p} & 0 \\ 0 & 0 & 0 & 0 & I_{2p-\bar{n}-q2} \end{bmatrix}, \tag{6.14}$$

where $\bar{A}_1 = A_1 - M_{12}A_{41}$ and $\bar{A}_2 = A_2 - M_{12}A_{51}$. Therefore, apply the state equation transformation $T_{pre} = T_g T_f T_d T_b$, the state transformation $T_{aft}^{-1} = T_c T_a$, and the fault transformation T_e such that

$$\begin{gathered} E_0 \mapsto T_{pre} E_0 T_{aft}, \; A_0 \mapsto T_{pre} A_0 T_{aft}, \; B_0 \mapsto T_{pre} B_0, \; M_0 \mapsto T_{pre} M_0 T_e^{-1}, \\ Q_0 \mapsto T_{pre} Q_0, \; C_0 \mapsto C_0 T_{aft}, \; x_0 \mapsto x_f = T_{aft}^{-1} x_0, \; f_0 \mapsto f = T_e f_0. \end{gathered} \tag{6.15}$$

From the structures of M_2 from (6.8) and $\begin{bmatrix} A_{42} & A_{52} \end{bmatrix}$ from (6.13), the system in the coordinates of (6.15) can be expressed as follows:

$$\underbrace{\begin{bmatrix} 0 & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & \bar{E}_{21} \\ 0 & 0 & E_{221} \\ 0 & 0 & E_{222} \\ 0 & 0 & E_{223} \end{bmatrix}}_{\check{E}} \dot{x}_f = \underbrace{\begin{bmatrix} 0 & 0 & \bar{A}_3 \\ 0 & 0 & \bar{A}_{61} \\ I_{n-\bar{n}} & 0 & A_{621} \\ 0 & I_{\bar{n}-p} & A_{622} \\ 0 & 0 & A_{623} \end{bmatrix}}_{\check{A}} x_f + \underbrace{\begin{bmatrix} \bar{B}_1 \\ \bar{B}_{21} \\ B_{221} \\ B_{222} \\ B_{223} \end{bmatrix}}_{\check{B}} u + \underbrace{\begin{bmatrix} M_{11} & 0 \\ 0 & I_{q2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\check{M}} f + \underbrace{\begin{bmatrix} \bar{Q}_1 \\ \bar{Q}_{21} \\ Q_{221} \\ Q_{222} \\ Q_{223} \end{bmatrix}}_{\check{Q}} \xi_0, \tag{6.16}$$

$$y = \underbrace{\begin{bmatrix} 0 & 0 & I_p \end{bmatrix}}_{\check{C}} x_f. \tag{6.17}$$

Note that system (6.16)–(6.17) is unrelated to filter (6.4), and thus the disturbances affecting it are still denoted as ξ_0. In the next (and remaining) portion of the proof, the dynamics of the disturbance signals ξ_0 are augmented into the system to ease observer design in the later sections of the chapter. Define an augmented state x_g as follows:

$$x_g = \begin{bmatrix} \xi_0 \\ x_f \end{bmatrix}. \tag{6.18}$$

Using the dynamics of ξ_0 from (6.3)-(6.4) and (6.16)–(6.17), the following augmented system is formulated:

$$\begin{bmatrix} I_h & 0 \\ 0 & \check{E} \end{bmatrix} \dot{x}_g = \begin{bmatrix} A_\Omega & 0 \\ \check{Q} & \check{A} \end{bmatrix} x_g + \begin{bmatrix} 0 \\ \check{B} \end{bmatrix} u + \begin{bmatrix} 0 \\ \check{M} \end{bmatrix} f + \begin{bmatrix} B_\Omega \\ 0 \end{bmatrix} \xi, \tag{6.19}$$

$$y = \begin{bmatrix} 0 & \check{C} \end{bmatrix} x_g. \tag{6.20}$$

By substituting from the structures in (6.16)–(6.17), system (6.19)–(6.20) becomes

$$\underbrace{\begin{bmatrix} I_h & 0 & 0 & 0 \\ 0 & 0 & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & 0 & \bar{E}_{21} \\ 0 & 0 & 0 & E_{221} \\ 0 & 0 & 0 & E_{222} \\ 0 & 0 & 0 & E_{223} \end{bmatrix}}_{E} \dot{x}_g = \underbrace{\begin{bmatrix} A_\Omega & 0 & 0 & 0 \\ \bar{Q}_1 & 0 & 0 & \bar{A}_3 \\ \bar{Q}_{21} & 0 & 0 & \bar{A}_{61} \\ Q_{221} & I_{n-\bar{n}} & 0 & A_{621} \\ Q_{222} & 0 & I_{\bar{n}-p} & A_{622} \\ Q_{223} & 0 & 0 & A_{623} \end{bmatrix}}_{A} x_g + \underbrace{\begin{bmatrix} 0 \\ \bar{B}_1 \\ \bar{B}_{21} \\ B_{221} \\ B_{222} \\ B_{223} \end{bmatrix}}_{B} u + \underbrace{\begin{bmatrix} B_\Omega \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{Q} \xi$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 \\ M_{11} & 0 \\ 0 & I_{q2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{M} f, \tag{6.21}$$

$$y = \underbrace{\begin{bmatrix} 0 & 0 & 0 & I_p \end{bmatrix}}_{C} x_g. \tag{6.22}$$

Finally, define an invertible matrix T_h such that

$$T_h = \begin{bmatrix} Q_{221} & I_{n-\bar{n}} & 0 & 0 \\ I_h & 0 & 0 & 0 \\ Q_{222} & 0 & I_{\bar{n}-p} & 0 \\ 0 & 0 & 0 & I_p \end{bmatrix}. \tag{6.23}$$

Apply the state transformation T_h onto system (6.21)–(6.22) as follows:

$$E \mapsto ET_h^{-1},\ A \mapsto AT_h^{-1},\ C \mapsto CT_h^{-1},\ x_g \mapsto x = T_{aft}^{-1} x_f = \begin{bmatrix} x_1 \\ \xi_0 \\ x_2 \\ y \end{bmatrix}. \tag{6.24}$$

It can then be seen by comparison that system (6.10)–(6.12) and system (6.21)–(6.22) in the coordinates of (6.24) are identical, thus completing the proof. ∎

Note that (6.9) is the necessary condition for the structure in (6.13) to be obtained. The system in the form of (6.10)–(6.12) allows the following comparison to be made between the proposed scheme and existing SMO schemes in the literature.

Remark 6.2 *In earlier work on fault reconstruction for descriptor systems, the schemes utilising a single SMO (such as [157] from Chapter 3, and [17, 100, 101]) require* $\mathrm{rank}(M_2) = \mathrm{rank}(M)$. *This condition sets constraints on the faults and satisfying it may require more measurable outputs. This condition, however, is not required by the scheme proposed in this chapter. Furthermore, the cascaded observer schemes by [99, 161] also require the following additional conditions to be satisfied:*

M1. $\mathrm{rank}\begin{bmatrix} \bar{E}_{21} \\ E_{222} \\ E_{223} \end{bmatrix} = r - \bar{n} + p$ *(for [99]),*

M2. $\mathrm{rank}\begin{bmatrix} E_{221} \\ E_{222} \\ E_{223} \end{bmatrix} = \mathrm{rank}\begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix}$ *(for [161]).* ♯

Conditions M1–M2 limit the applicability of the schemes in [99, 161]. Moreover, the schemes in [99, 100, 101, 161] do not consider the effect of disturbances on the fault reconstruction. The contribution of this chapter is therefore to improve on the schemes presented in [17, 99, 100, 101, 161] by relaxing the conditions required to reconstruct f, and also to make the reconstruction robust against ξ_0. Define an invertible matrix

$$T = \begin{bmatrix} I_{h+\bar{n}-p} & 0 \\ 0 & T_p \\ 0 & \bar{T}_p \end{bmatrix}, \tag{6.25}$$

where $T_p \in \mathbb{R}^{p\times(n-\bar{n}+p)}$. Then apply the state equation transformation T onto system

(6.10)–(6.12) to obtain

$$\begin{bmatrix} 0 & I_h & 0 & 0 \\ 0 & -Q_{222} & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & 0 & T_p\bar{E}_2 \\ 0 & 0 & 0 & \bar{T}_p\bar{E}_2 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{\xi}_0 \\ \dot{x}_2 \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & A_\Omega & 0 & 0 \\ 0 & \bar{Q}_1 & 0 & \bar{A}_3 \\ T_p\bar{A}_4 & T_p\bar{Q}_2 & T_p\bar{A}_5 & T_p\bar{A}_6 \\ \bar{T}_p\bar{A}_4 & \bar{T}_p\bar{Q}_2 & \bar{T}_p\bar{A}_5 & \bar{T}_p\bar{A}_6 \end{bmatrix} \begin{bmatrix} x_1 \\ \xi_0 \\ x_2 \\ y \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ \bar{B}_1 \\ T_p\bar{B}_2 \\ \bar{T}_p\bar{B}_2 \end{bmatrix} u + \begin{bmatrix} 0 & 0 \\ M_{11} & 0 \\ 0 & T_p\bar{M}_2 \\ 0 & \bar{T}_p\bar{M}_2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} B_\Omega \\ 0 \\ 0 \\ 0 \end{bmatrix} \xi, \quad (6.26)$$

$$y = \begin{bmatrix} 0 & 0 & 0 & I_p \end{bmatrix} \begin{bmatrix} x_1 \\ \xi_0 \\ x_2 \\ y \end{bmatrix}. \quad (6.27)$$

Then treat x_1 as an unknown input, and define

$$\bar{x} = \begin{bmatrix} \xi_0 \\ x_2 \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} f \\ x_1 \end{bmatrix}, \quad (6.28)$$

which respectively are the state and fault of the following reduced-order system (that has been re-expressed from (6.26)–(6.27)):

$$\underbrace{\begin{bmatrix} I_h & 0 & 0 \\ -Q_{222} & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & T_p\bar{E}_2 \end{bmatrix}}_{\bar{E}} \dot{\bar{x}} = \underbrace{\begin{bmatrix} A_\Omega & 0 & 0 \\ \bar{Q}_1 & 0 & \bar{A}_3 \\ T_p\bar{Q}_2 & T_p\bar{A}_5 & T_p\bar{A}_6 \end{bmatrix}}_{\bar{A}} \bar{x} + \underbrace{\begin{bmatrix} 0 \\ \bar{B}_1 \\ T_p\bar{B}_2 \end{bmatrix}}_{\bar{B}} u$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ M_{11} & 0 & 0 \\ 0 & T_p\bar{M}_2 & T_p\bar{A}_4 \end{bmatrix}}_{\bar{M}} \bar{f} + \underbrace{\begin{bmatrix} B_\Omega \\ 0 \\ 0 \end{bmatrix}}_{\bar{Q}} \xi, \quad (6.29)$$

$$y = \underbrace{\begin{bmatrix} 0 & 0 & I_p \end{bmatrix}}_{\bar{C}} \bar{x}. \quad (6.30)$$

Note that the rows pre-multiplied with $\bar{T}_p$ are linearly dependent on those pre-multiplied with T_p (from its definition in (4.12)), and are hence omitted from further analysis. Furthermore, notice that $\operatorname{rank}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = h + \bar{n}$ (which is full-column rank) and the output of the reduced-order system (6.29)–(6.30) is a measurable signal y. The observer by Yeu et al. [157] can therefore be designed based on $\left(\bar{E}, \bar{A}, \bar{B}, \bar{M}, \bar{Q}, \bar{C}\right)$ and driven by u and y, and applied onto system (6.29)–(6.30) to reconstruct $\bar{f}$, and therefore reconstruct f_0.

6.3 Observer formulation

Figure 6.1 shows a schematic diagram of the cascaded observer scheme proposed in this chapter.

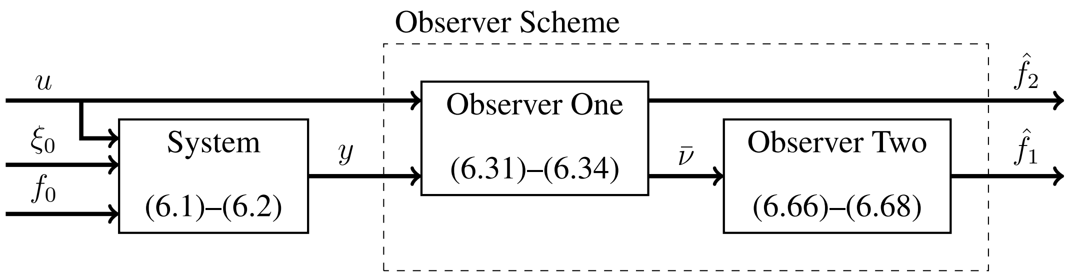

Figure 6.1: Schematic diagram of the scheme proposed in the chapter.

6.3.1 Observer One

Observer One for system (6.29)–(6.30) has the following structure [157]:

$$\dot{z} = \left(R\bar{A} - G_l\bar{C}\right)z - R\bar{B}u - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu, \tag{6.31}$$

$$\hat{\bar{x}} = Vy - z, \tag{6.32}$$

$$\nu = -\rho\frac{e_y}{\|e_y\|}, \tag{6.33}$$

$$e_y = \bar{C}\hat{\bar{x}} - y,\ G_l = \begin{bmatrix}G_{l1}\\ G_{l2}\\ G_{l3}\end{bmatrix},\ G_n = \begin{bmatrix}0\\ G_{n3}\end{bmatrix}, \tag{6.34}$$

where $G_{l1} \in \mathbb{R}^{h\times p},\ G_{l2} \in \mathbb{R}^{(\bar{n}-p)\times p},\ |G_{n3}| \neq 0,\ R \in \mathbb{R}^{(h+\bar{n})\times(h+\bar{n})}$ is invertible, and $V \in \mathbb{R}^{(h+\bar{n})\times p}$. Pre-multiply (6.29) with R and add $V\dot{y}$ to both sides to obtain

$$R\bar{E}\dot{\bar{x}} + V\dot{y} = \left(R\bar{E} + V\bar{C}\right)\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\dot{y}. \tag{6.35}$$

Suppose $R\bar{E} + V\bar{C} = I_{\bar{n}}$. Then (6.35) becomes

$$\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\bar{C}\dot{\bar{x}}. \tag{6.36}$$

Corollary 6.1 *The matrices* R, V *from* (6.31)–(6.32) *will have the following structures:*

$$R = \begin{bmatrix} I_h & 0 & R_3 \\ Q_{222} & I_{\bar{n}-p} & R_6 \\ 0 & 0 & R_9 \end{bmatrix}, \; V = \begin{bmatrix} -R_3 T_p \bar{E}_2 \\ -\left(\bar{E}_1 + R_6 T_p \bar{E}_2\right) \\ I_p - R_9 T_p \bar{E}_2 \end{bmatrix} \tag{6.37}$$

where $R_3 \in \mathbb{R}^{h \times p}$ *and* $R_6 \in \mathbb{R}^{(\bar{n}-p) \times p}$ *have specific structures that will be defined later in* (6.55)*, and* $R_9 \in \mathbb{R}^{p \times p}$, $|R_9| \neq 0$. ♯

Proof Since $\operatorname{rank} \begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = h + \bar{n}$, then the matrices R and V can be chosen such that

$$\begin{bmatrix} R & V \end{bmatrix} \begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = I_{h+\bar{n}}, \tag{6.38}$$

that is, $\begin{bmatrix} R & V \end{bmatrix}$ is chosen to be the Moore-Penrose inverse of $\begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix}$. Partition the matrices R and V generally as follows:

$$\begin{bmatrix} R & V \end{bmatrix} = \left[\begin{array}{ccc|c} R_1 & R_2 & R_3 & V_1 \\ R_4 & R_5 & R_6 & V_2 \\ R_7 & R_8 & R_9 & V_3 \end{array}\right], \tag{6.39}$$

where $R_1 \in \mathbb{R}^{h \times h}$, $R_5 \in \mathbb{R}^{(\bar{n}-p) \times (\bar{n}-p)}$, and $R_9 \in \mathbb{R}^{p \times p}$. By substituting $\bar{E}$ and $\bar{C}$ from (6.29)–(6.30) and R and V from (6.39) into (6.38), it is straightforward to see that R and V would have the forms given in (6.37). ■

Substitute z from (6.32) into (6.31) to obtain the following analytical structure for the observer:

$$\begin{aligned} \left(V\dot{y} - \dot{\hat{\bar{x}}}\right) &= \left(R\bar{A} - G_l\bar{C}\right)\left(Vy - \hat{\bar{x}}\right) - R\bar{B}u - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu \\ \dot{\hat{\bar{x}}} &= \left(R\bar{A} - G_l\bar{C}\right)\hat{\bar{x}} + R\bar{B}u + G_l\bar{C}\bar{x} + G_n\nu + V\bar{C}\dot{\bar{x}}. \end{aligned} \tag{6.40}$$

Define the state estimation error for Observer One as

$$e = \hat{\bar{x}} - \bar{x} = \left[\begin{array}{c} e_\xi \\ e_2 \\ \hline e_y \end{array}\right] \begin{array}{l} \updownarrow \ h \\ \updownarrow \ \bar{n}-p \\ \updownarrow \ p \end{array} = \left[\begin{array}{c} \bar{e} \\ \hline e_y \end{array}\right]. \tag{6.41}$$

Hence by subtracting (6.36) from (6.40), the error equation for Observer One (which characterises its performance) is given by

$$\dot{e} = \left(R\bar{A} - G_l\bar{C}\right)e - R\bar{M}\bar{f} - R\bar{Q}\xi + G_n\nu, \tag{6.42}$$

Define $\tilde{Q}_1 = \bar{Q}_1 + Q_{222}A_\Omega$, and substitute the structures of $\bar{A}, \bar{M}, \bar{Q}, \bar{C}$ from (6.29), G_l, G_n from (6.34), and R from (6.37) into (6.42) to obtain

$$\dot{e} = \begin{bmatrix} A_\Omega + R_3T_p\bar{Q}_2 & R_3T_p\bar{A}_5 & R_3T_p\bar{A}_6 - G_{l1} \\ \tilde{Q}_1 + R_6T_p\bar{Q}_2 & R_6T_p\bar{A}_5 & \bar{A}_3 + R_6T_p\bar{A}_6 - G_{l2} \\ R_9T_p\bar{Q}_2 & R_9T_p\bar{A}_5 & R_9T_p\bar{A}_6 - G_{l3} \end{bmatrix} e - \begin{bmatrix} B_\Omega \\ Q_{222}B_\Omega \\ 0 \end{bmatrix}\xi + \begin{bmatrix} 0 \\ 0 \\ G_{n3} \end{bmatrix}\nu$$
$$- \begin{bmatrix} 0 & R_3T_p\bar{M}_2 & R_3T_p\bar{A}_4 \\ M_{11} & R_6T_p\bar{M}_2 & R_6T_p\bar{A}_4 \\ 0 & R_9T_p\bar{M}_2 & R_9T_p\bar{A}_4 \end{bmatrix}\bar{f}. \tag{6.43}$$

Proposition 6.2 *Suppose there exists a matrix $P = P^T > 0$ that satisfies*

$$P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0, \tag{6.44}$$

where $P = \begin{bmatrix} P_{11} & 0 & 0 \\ 0 & P_{12} & 0 \\ 0 & 0 & G_{n3}^{-1} \end{bmatrix}$, $P_{11} \in \mathbb{R}^{h\times h}$, $P_{12} \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$. *If ρ in* (6.33) *is chosen as follows:*

$$\rho > \frac{2}{\mu_1}\|Y\begin{bmatrix}\bar{Q}_2 & \bar{A}_5\end{bmatrix}\|\delta + \|Y\begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\|\alpha_1 + \eta,$$
$$\mu_1 = -\lambda_{\max}\left(P(R\bar{A} - G_l\bar{C}) + (R\bar{A} - G_l\bar{C})^T P\right),\ \delta = \left(\|PR\bar{M}\|\alpha_1 + \|PR\bar{Q}\|\beta\right), \tag{6.45}$$
$$\alpha_1 > \|f\|_{\max} + \|x_1\|_{\max},\ \beta > \|\xi\|_{\max},\ Y = G_{n3}^{-1}R_9T_p,\ \eta \in \mathbb{R}^+,$$

then sliding motion $(e_y, \dot{e}_y = 0)$ occurs in finite time.

Proof As (6.43) has the same form as the error system for the Edwards-Spurgeon SMO for a system represented by the quadruple $(R\bar{A}, R\bar{M}, R\bar{Q}, \bar{C})$, the proof of convergence is adapted from [128]. There are two portions to the proof: the first portion aims to show that by setting ρ appropriately, e can be made to be ultimately bounded. Define a candidate Lyapunov function $W = e^T Pe > 0$, and differentiate it with respect to time as follows:

$$\begin{aligned}\dot{W} &= e^T \left(P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P\right)e - 2e^T\left(PR\bar{M}\bar{f} + PR\bar{Q}\xi\right) + 2e_y^T\nu, \\ &\leq -\mu_1\|e\|^2 + 2\|e\|\delta - 2\rho\|e_y\|. \end{aligned} \tag{6.46}$$

By setting $\rho > \eta$, (6.46) becomes

$$\dot{W} \leq \|e\|\left(-\mu_1\|e\| + 2\delta\right). \tag{6.47}$$

When the magnitude of the error e is smaller than or equal to a certain bound, i.e. $\|e\| \leq \frac{2\delta}{\mu_1}$, then (6.47) becomes

$$\dot{W} \leq \kappa_1, \tag{6.48}$$

where $\kappa_1 \geq 0$, and thus the magnitude of e can increase, decrease, or remain unchanged within this bound. If the magnitude of the error e is larger than that bound however, i.e. $\|e\| \geq \frac{2\delta}{\mu_1}$, then (6.47) becomes

$$\dot{W} < 0, \tag{6.49}$$

and $\|e\|$ will shrink. This implies that the magnitude of e would be bounded $\left(\|e\| \leq \frac{2\delta}{\mu_1}\right)$ in finite time.

The succeeding (and remaining) portion of the proof aims to show how sliding motion $(e_y, \dot{e}_y = 0)$ is induced. Define another candidate Lyapunov function $W_y = e_y^T G_{n3}^{-1} e_y > 0$, and differentiating it with respect to time to obtain

$$\begin{aligned}\dot{W}_y &= e_y^T\left(Y\bar{A}_6 - G_{n3}^{-1}G_{l3} + \left(Y\bar{A}_6 - G_{n3}^{-1}G_{l3}\right)^T\right)e_y \\ &\quad + 2e_y^T Y\left(\begin{bmatrix}\bar{Q}_2 & \bar{A}_5\end{bmatrix}\bar{e} - \begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\bar{f}\right) + 2e_y^T\nu \\ &\leq -2\|e_y\|\left(\rho - \frac{2}{\mu_1}\|Y\begin{bmatrix}\bar{Q}_2 & \bar{A}_5\end{bmatrix}\|\delta - \|Y\begin{bmatrix}0 & \bar{M}_2 & \bar{A}_4\end{bmatrix}\|\alpha_1\right). \end{aligned} \tag{6.50}$$

Notice that

$$\|e_y\|^2 = \left(\sqrt{G_{n3}^{-1}}e_y\right)^T G_{n3}\left(\sqrt{G_{n3}^{-1}}e_y\right) \geq \lambda_{\min}\left(G_{n3}\right)\|\sqrt{G_{n3}^{-1}}e_y\|^2 = \lambda_{\min}\left(G_{n3}\right)W_y. \tag{6.51}$$

Hence by setting ρ such that (6.45) is satisfied (which also satisfies $\rho > \eta$), (6.50) becomes

$$\dot{W}_y \leq -2\eta\|e_y\| \leq -2\eta\sqrt{\lambda_{\min}\left(G_{n3}\right)}\sqrt{W_y}, \tag{6.52}$$

which is the *reachability condition* [128] resulting in $e_y = 0$ in finite time and a sliding motion taking place on the surface $\mathcal{S}_1 = \{e : \bar{C}e = 0\}$, thus proving the proposition. ∎

Remark 6.3 *Notice that from Proposition 6.2, x, f, and ξ need to be bounded, and these bounds are required to calculate ρ. Knowledge of these bounds can be obtained through several methods, e.g. from simulations of the system operating under practical conditions, by physically operating the system, or by utilising a controller to stabilise the system [17]. On the other hand, in cases where these bounds are not easy to determine, the magnitude of ρ can be set to be adaptive to achieve convergence of e_y without* a-priori *knowledge of the upper bounds of $\|x\|$, $\|f\|$, and $\|\xi\|$ [47].* ♯

After sliding motion $\left(e_y, \dot{e}_y = 0\right)$ occurs, (6.43) becomes

$$\dot{\bar{e}} = \begin{bmatrix} A_\Omega + R_3T_p\bar{Q}_2 & R_3T_p\bar{A}_5 \\ \tilde{Q}_1 + R_6T_p\bar{Q}_2 & R_6T_p\bar{A}_5 \end{bmatrix}\bar{e} - \begin{bmatrix} 0 & R_3T_p\bar{M}_2 & R_3T_p\bar{A}_4 \\ M_{11} & R_6T_p\bar{M}_2 & R_6T_p\bar{A}_4 \end{bmatrix}\bar{f} - \begin{bmatrix} B_\Omega \\ Q_{222}B_\Omega \end{bmatrix}\xi, \tag{6.53}$$

$$0 = R_9T_p\left(\begin{bmatrix} \bar{Q}_2 & \bar{A}_5 \end{bmatrix}\bar{e} - \begin{bmatrix} 0 & \bar{M}_2 & \bar{A}_4 \end{bmatrix}\bar{f}\right) + G_{n3}\nu. \tag{6.54}$$

6.3.1.1 Reconstructing f_2

Let T_p from (6.25) and R_3, R_6 from (6.37) have the following structures:

$$T_p = \begin{bmatrix} I_{q2+n-\bar{n}} & 0 & 0 \\ 0 & -I_{\bar{n}-p} & 0 \\ 0 & 0 & T_{44} \end{bmatrix}, \begin{bmatrix} R_3 \\ R_6 \end{bmatrix} = \left[\begin{array}{ccc} 0 & 0 & R_{34} \\ \hline 0 & J & 0 \end{array}\right], \tag{6.55}$$

where $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$, $R_{34} \in \mathbb{R}^{h\times(2p-n-q_2)}$, and $J \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)} > 0$. Next, define

$$\begin{bmatrix} R_{3t} \\ R_{6t} \end{bmatrix} = \left[\begin{array}{ccc} 0 & 0 & R_{34}T_{44} \\ \hline 0 & -J & 0 \end{array}\right], \ \bar{\nu} = R_9^{-1}G_{n3}\nu, \ \tilde{A}_\Omega = A_\Omega + R_{34}T_{44}Q_{223}, \tag{6.56}$$

where $\bar{\nu}$ is a measurable signal. Substituting (6.55)–(6.56) into (6.53)–(6.54) yields

$$\dot{\bar{e}} = \begin{bmatrix} A_\Omega + R_{3t}\bar{Q}_2 & R_{3t}\bar{A}_5 \\ \tilde{Q}_1 + R_{6t}\bar{Q}_2 & R_{6t}\bar{A}_5 \end{bmatrix} \bar{e} - \begin{bmatrix} 0 & R_{3t}\bar{M}_2 & R_{3t}\bar{A}_4 \\ M_{11} & R_{6t}\bar{M}_2 & R_{6t}\bar{A}_4 \end{bmatrix} \bar{f} - \begin{bmatrix} B_\Omega \\ Q_{222}B_\Omega \end{bmatrix} \xi, \tag{6.57}$$

$$\bar{\nu} = T_p\left(- \begin{bmatrix} \bar{Q}_2 & \bar{A}_5 \end{bmatrix} \bar{e} + \begin{bmatrix} 0 & \bar{M}_2 & \bar{A}_4 \end{bmatrix} \bar{f} \right). \tag{6.58}$$

Finally, substitute the structures of $\bar{A}_4,\ \bar{A}_5,\ \bar{M}_2$, and $\bar{Q}_2$ from (6.12) into (6.57)–(6.58) to obtain

$$\dot{\bar{e}} = \begin{bmatrix} \tilde{A}_\Omega & 0 \\ \tilde{Q}_1 & -J \end{bmatrix} \bar{e} - \begin{bmatrix} 0 & 0 & 0 \\ M_{11} & 0 & 0 \end{bmatrix} \bar{f} - \begin{bmatrix} B_\Omega \\ Q_{222}B_\Omega \end{bmatrix} \xi, \tag{6.59}$$

$$\bar{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix} = \begin{bmatrix} -\bar{Q}_{21} & 0 \\ 0 & 0 \\ 0 & I_{\bar{n}-p} \\ -T_{44}Q_{223} & 0 \end{bmatrix} \bar{e} + \begin{bmatrix} 0 & I_{q2} & 0 \\ 0 & 0 & I_{n-\bar{n}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bar{f}. \tag{6.60}$$

Define the reconstruction for f_2 as

$$\hat{f}_2 = \nu_1 + L\nu_4, \tag{6.61}$$

and define the corresponding reconstruction error as

$$e_{f2} = \hat{f}_2 - f_2. \tag{6.62}$$

From (6.60)–(6.61) it follows that

$$e_{f2} = -\left(\bar{Q}_{21} + LT_{44}Q_{223}\right)\bar{e}. \tag{6.63}$$

Remark 6.4 *Notice that only f_2 can be reconstructed from Observer One, as f_1 cannot be obtained from a linear combination of measurable signals in* (6.60). *This is a limitation common to other schemes that utilise a single SMO [17, 100, 101]. A second observer is therefore required to reconstruct the remaining faults.* ♯

6.3.2 Observer Two

To reconstruct f_1, a system based on the measurable signals from (6.59)–(6.60) is first formulated. Rearrange (6.59)–(6.60) to form the following system with the measurable output ν_3:

$$\dot{\bar{e}} = \underbrace{\begin{bmatrix} \tilde{A}_\Omega & 0 \\ \tilde{Q}_1 & -J \end{bmatrix}}_{\tilde{A}} \bar{e} + \underbrace{\begin{bmatrix} 0 \\ -M_{11} \end{bmatrix}}_{\tilde{M}} f_1 + \underbrace{\begin{bmatrix} -B_\Omega \\ -Q_{222}B_\Omega \end{bmatrix}}_{\tilde{Q}} \xi, \tag{6.64}$$

$$\nu_3 = \underbrace{\begin{bmatrix} 0 & I_{\bar{n}-p} \end{bmatrix}}_{\tilde{C}} \bar{e}. \tag{6.65}$$

It can be seen that (6.64)–(6.65) is a regular state-space system: hence the Edwards-Spurgeon SMO (see p. 134 of [45]) can be designed based on the quadruple $\left(\tilde{A}, \tilde{M}, \tilde{Q}, \tilde{C}\right)$ and driven by ν_3 to reconstruct f_1 robustly against ξ_0. The observer (Observer Two) for (6.64)–(6.65) has the following structure [8]:

$$\dot{\hat{\bar{e}}} = \tilde{A}\hat{\bar{e}} - \tilde{G}_n \tilde{\nu}, \tag{6.66}$$

$$\tilde{\nu} = -\tilde{\rho}\frac{\tilde{e}_2}{\|\tilde{e}_2\|}, \tag{6.67}$$

$$\tilde{e}_2 = \tilde{C}\tilde{e},\ \tilde{e} = \bar{e} - \hat{\bar{e}} = \begin{bmatrix} \tilde{e}_\xi \\ \tilde{e}_2 \end{bmatrix},\ \tilde{G}_n = \begin{bmatrix} 0 \\ \tilde{G}_{n2} \end{bmatrix}. \tag{6.68}$$

where $|\tilde{G}_{n2}| \neq 0$.

Proposition 6.3 *Suppose there exists a matrix $P_1 = P_1^T > 0$ that satisfies*

$$P_1\tilde{A} + \tilde{A}^T P_1 < 0, \tag{6.69}$$

where $P_1 = \begin{bmatrix} P_{11} & 0 \\ 0 & \tilde{G}_{n2}^{-1} \end{bmatrix}$. *If $\tilde{\rho}$ in* (6.67) *is chosen as follows:*

$$\begin{aligned} \tilde{\rho} &\geq \frac{2\beta}{\mu_2}\|\tilde{G}_{n2}^{-1}\tilde{Q}_1\|\|P_1\tilde{Q}\| + \|\tilde{G}_{n2}^{-1}M_{11}\|\alpha_2 + \|\tilde{G}_{n2}^{-1}Q_{222}B_\Omega\|\beta + \zeta, \\ \mu_2 &= -\lambda_{\max}\left(P_1\tilde{A} + \tilde{A}^T P_1\right) > 0,\ \alpha_2 > \|f_1\|_{\max},\ \zeta > 0, \end{aligned} \tag{6.70}$$

then sliding motion $\left(\tilde{e}_2, \dot{\tilde{e}}_2 = 0\right)$ *occurs in finite time.* ♯

Proof From the definition of $\tilde{e}$ in (6.68), subtracting (6.66) from (6.64) yields

$$\dot{\tilde{e}} = \tilde{A}\tilde{e} + \tilde{M}f_1 + \tilde{Q}\xi + \tilde{G}_n\tilde{\nu}. \tag{6.71}$$

Define a candidate Lyapunov function $\tilde{W} = \tilde{e}^T P_1 \tilde{e} > 0$, and differentiating it with respect to time yields

$$\begin{aligned}\dot{\tilde{W}} &= \tilde{e}^T \left(P_1 \tilde{A} + \tilde{A}^T P_1\right) \tilde{e} + 2\tilde{e}^T P_1 \tilde{M} f_1 + 2\tilde{e}^T P_1 \tilde{Q}_1 \xi + 2\tilde{e}_2 \tilde{\nu}, \\ &\le -\mu_2 \|\tilde{e}\|^2 + 2\|\tilde{e}\| \|P_1 \tilde{Q}\| \beta - 2\|\tilde{e}_2\| \left(\tilde{\rho} - \|\tilde{G}_{n2}^{-1} M_{11}\| \alpha_2\right). \end{aligned} \tag{6.72}$$

By setting $\tilde{\rho} > \|\tilde{G}_{n2}^{-1} M_{11}\| \alpha_2$, (6.72) becomes

$$\dot{\tilde{W}} \le \|\tilde{e}\| \left(-\mu_2 \|\tilde{e}\| + 2\|P_1 \tilde{Q}\| \beta\right). \tag{6.73}$$

If the magnitude of the error $\tilde{e}$ is smaller than or equal to a certain bound, i.e. $\|\tilde{e}\| \le \frac{2\|P_1 \tilde{Q}\|\beta}{\mu_2}$, then (6.73) becomes

$$\dot{\tilde{W}} \le \kappa_2, \tag{6.74}$$

where $\kappa_2 \ge 0$, and the magnitude of $\tilde{e}$ can become smaller, larger, or remain constant within this bound. If the magnitude of the error $\tilde{e}$ is however larger than that bound, i.e. $\|\tilde{e}\| > \frac{2\|P_1 \tilde{Q}\|\beta}{\mu_2}$, then (6.73) becomes

$$\dot{\tilde{W}} < 0, \tag{6.75}$$

and $\|\tilde{e}\|$ would shrink. This implies that the magnitude of $\tilde{e}$ would be bounded $\left(\|e\| \le \frac{2\|P_1 \tilde{Q}\|\beta}{\mu_2}\right)$ in finite time. The succeeding portion of the proof aims to show how sliding motion $\left(\tilde{e}_2, \dot{\tilde{e}}_2 = 0\right)$ is induced.

Next, define another candidate Lyapunov function $\tilde{W}_y = \tilde{e}_2^T \tilde{G}_{n2}^{-1} \tilde{e}_2 > 0$. Differentiating $\tilde{W}_y$ with respect to time yields

$$\begin{aligned}\dot{\tilde{W}}_y &= \tilde{e}_2^T \left(\tilde{G}_{n2}^{-1} J + J^T \tilde{G}_{n2}^{-1}\right) \tilde{e}_2 + 2\tilde{e}_2^T \tilde{G}_{n2}^{-1} \left(\tilde{Q}_1 \tilde{e}_\xi - M_{11} f_1 - Q_{222} B_\Omega \xi\right) + 2\tilde{e}_2^T \tilde{\nu}, \\ &\le -2\|\tilde{e}_2\| \left(\tilde{\rho} - \frac{2\beta}{\mu_2} \|\tilde{G}_{n2}^{-1} \tilde{Q}_1\| \|P_1 \tilde{Q}\| - \|\tilde{G}_{n2}^{-1} M_{11}\| \alpha_2 - \|\tilde{G}_{n2}^{-1} Q_{222} B_\Omega\| \beta\right). \end{aligned} \tag{6.76}$$

Notice that

$$\|\tilde{e_2}\|^2 = \left(\sqrt{\tilde{G}_{n2}^{-1}} \tilde{e}_2\right)^T \tilde{G}_{n2} \left(\sqrt{\tilde{G}_{n2}^{-1}} \tilde{e}_2\right) \ge \lambda_{\min}\left(\tilde{G}_{n2}\right) \|\sqrt{\tilde{G}_{n2}^{-1}} \tilde{e}_2\|^2 = \lambda_{\min}\left(\tilde{G}_{n2}\right) \tilde{W}_y. \tag{6.77}$$

Hence by setting $\tilde{\rho}$ such that (6.70) is satisfied, (6.77) becomes

$$\dot{\tilde{W}}_y \le -2\zeta \|\tilde{e}_2\| \le -2\zeta \sqrt{\lambda_{\min}\left(\tilde{G}_{n2}\right)} \sqrt{\tilde{W}_y}, \tag{6.78}$$

which is the *reachability condition* [128] resulting in $\tilde{e}_2 = 0$ (and sliding motion taking place on the surface $\mathcal{S}_2 = \{\tilde{e} : \tilde{C}\tilde{e} = 0\}$) in finite time, thus proving the proposition. ■

After sliding motion occurs, (6.71) becomes

$$\dot{\tilde{e}}_\xi = \tilde{A}_\Omega \tilde{e}_\xi - B_\Omega \xi, \tag{6.79}$$

$$0 = \tilde{Q}_1 \tilde{e}_\xi - M_{11} f_1 - Q_{222} B_\Omega \xi + \tilde{G}_{n2} \tilde{\nu}. \tag{6.80}$$

It can be seen that f_1 can be reconstructed from $\tilde{\nu}$ (which is measurable) from (6.80). Hence define the reconstruction for f_1 as

$$\hat{f}_1 = K M_{11}^{\dagger} \tilde{G}_{n2} \tilde{\nu}, \tag{6.81}$$

where $M_{11}^{\dagger}$ is the Moore-Penrose inverse of M_{11} and is defined (along with K) as

$$M_{11}^{\dagger} M_{11} = \begin{bmatrix} 0 \\ I_{q-q_2} \end{bmatrix}, \; K = \begin{bmatrix} K_1 & I_{q-q_2} \end{bmatrix}, \tag{6.82}$$

where $K_1 \in \mathbb{R}^{(q-q_2)\times(n-\bar{n}-q+q_2)}$. Define also the reconstruction error for f_1 as follows:

$$e_{f1} = \hat{f}_1 - f_1. \tag{6.83}$$

From (6.80)–(6.82), (6.83) becomes

$$e_{f1} = -K\tilde{Q}_1 \tilde{e}_\xi + K Q_{222} B_\Omega \xi. \tag{6.84}$$

Thus the reconstructions for f (i.e. $\hat{f}_1$ and $\hat{f}_2$) have been derived. The signals $\bar{e}$, $\tilde{e}_\xi$, and ξ however affect $\hat{f}_1$ and $\hat{f}_2$. The goal now is to design the observer gains such that the effect of ξ on e_{f1} and e_{f2} will be minimised, which will be addressed in the next subsection.

6.3.3 Observer design

Define the following:

$$e_f = \begin{bmatrix} e_{f1} \\ e_{f2} \end{bmatrix}, \; \tilde{C} = \begin{bmatrix} -K\tilde{Q}_1 \\ -\left(\bar{Q}_{21} + L T_{44} Q_{223}\right) \end{bmatrix}, \; H = \begin{bmatrix} K Q_{222} B_\Omega \\ 0 \end{bmatrix}. \tag{6.85}$$

where e_f is referred to as the *fault reconstruction error*. From (6.59), (6.63), (6.69), and (6.84), the following state-space system can be formed:

$$\dot{\breve{e}} = \tilde{A}_\Omega \breve{e} - B_\Omega \xi, \tag{6.86}$$

$$e_f = \tilde{C}\breve{e} + H\xi, \tag{6.87}$$

where $\breve{e}$ is a state variable representing e_ξ from (6.41) and $\tilde{e}_\xi$ from (6.68).

Remark 6.5 *System* (6.86)–(6.87) *is a state-space system describing how the disturbance ξ affects the fault reconstruction error e_f. If $\xi = 0$ (i.e. system* (6.1)–(6.2) *is not affected by disturbances),* (6.86)–(6.87) *becomes*

$$\dot{\breve{e}} = \tilde{A}_\Omega \breve{e}, \tag{6.88}$$

$$e_f = \tilde{C}\breve{e}, \tag{6.89}$$

which implies $e_f \to 0$, which in turn implies $\hat{f}_1 \to f_1,\ \hat{f}_2 \to f_2$. Therefore, in the absence of disturbances, the proposed scheme is able to asymptotically reconstruct the faults. Thus, the scheme proposed in this chapter are also applicable to systems without considerations for external disturbances, and is therefore more general than the scheme proposed in the previous chapter. ♯

The effect of ξ on e_f is thus minimised using the Bounded Real Lemma, which has been described in Lemma 4.2. Define $P_1 = \begin{bmatrix} P_{11} & 0 \\ 0 & P_{12} \end{bmatrix}$, where $P_1 = P_1^T > 0,\ P_{12} = \tilde{G}_{n2}^{-1}$. By applying the Bounded Real Lemma onto error system (6.86)–(6.87), the $\mathcal{L}_2$ gain from ξ to e_f will not exceed the positive scalar γ if there exists a matrix $P_{11} = P_{11}^T > 0$ such that

$$\begin{bmatrix} P_{11}\tilde{A}_\Omega + \tilde{A}_\Omega P_{11} & * & * \\ -B_\Omega^T P_{11} & -\gamma I_h & * \\ \tilde{C} & H & -\gamma I_q \end{bmatrix} < 0. \tag{6.90}$$

where $*$ are terms that make LMI (6.90) symmetrical. The objective is therefore to find solutions for $\gamma,\ P_{11},\ K$, and L to minimise γ subject to inequality (6.90), while also satisfying

$$P > 0,\ P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0,\ P_1\tilde{A} + \tilde{A}^T P_1 < 0, \tag{6.91}$$

which are required to guarantee sliding motion in both observers (refer to Propositions 6.2 and 6.3). The choice of G_l and J is not unique; the only condition is that (6.91) is satisfied. In this paper, G_l and J are designed using a modified version of the method in [128] that will be outlined in Proposition 6.4 below.

Proposition 6.4 *Define $\bar{D} \in \mathbb{R}^{p\times p}$ and $\gamma_0 \in \mathbb{R}^+$ as user-defined variables. Then define*

$$P_{34} = P_{11}R_{34},\ P_J = P_{12}J,\ P_9 = P_2R_9,\ F = \begin{bmatrix} \breve{C} & F_1 & F_2 \end{bmatrix}, \tag{6.92}$$

where $F_1 \in \mathbb{R}^{q\times(\bar{n}-p)}$ *and* $F_2 \in \mathbb{R}^{q\times p}$*. Suppose there exists matrices* $P,\ K_1,\ L,\ F_1,\ F_2$*, and* γ *that satisfy the following inequalities:*

$$\begin{bmatrix} X & * & * \\ -\left(R\bar{Q}\right)^T P & -\gamma_0 I_h & * \\ F & H & -\gamma_0 I_q \end{bmatrix} < 0, \tag{6.93}$$

$$\begin{bmatrix} \begin{matrix} P_{11}A_\Omega + P_{34}T_{44}Q_{223} \\ +A_\Omega^T P_{11} + (T_{44}Q_{223})^T P_{34} \end{matrix} & * & * \\ -B_\Omega^T P_{11} & -\gamma I_h & * \\ \breve{C} & H & -\gamma I_q \end{bmatrix} < 0, \tag{6.94}$$

where $*$ *are the terms that make* (6.93)–(6.94) *symmetric, and*

$$X = P\left(R\bar{A}\right) + \left(R\bar{A}\right)^T P - \gamma_0 \bar{C}^T \left(\bar{D}\bar{D}^T\right)^{-1} \bar{C}. \tag{6.95}$$

If the observer parameters are chosen as

$$R_{34} = P_{11}^{-1}P_{34},\ J = P_{12}^{-1}P_J,\ R_9 = P_2^{-1}P_9,\ G_l = \frac{\gamma_0}{2}P^{-1}\bar{C}^T\left(\bar{D}\bar{D}^T\right)^{-1}, \tag{6.96}$$

then (6.91) *is satisfied, and* $\|e_f\| \le \gamma\|\xi\|$. ♯

Proof The structure of G_l in (6.96) implies

$$PG_l\bar{C} + (G_l\bar{C})^T P = \gamma_0 \bar{C}^T(\bar{D}\bar{D}^T)^{-1}\bar{C}, \tag{6.97}$$

that is, the top-left element of LMI (6.93) (i.e. X) becomes

$$X = P(R\bar{A} - G_l\bar{C}) + (R\bar{A} - G_l\bar{C})^T P. \tag{6.98}$$

The structure of LMI (6.93) and X in (6.98) imply $P\left(R\bar{A}\right) + \left(R\bar{A}\right)^T P < 0$. Furthermore, from the structures of $\bar{A}$ and $\bar{C}$ from (6.29)–(6.30), and G_l and R from (6.37), the following expression is obtained:

$$R\bar{A} - G_l\bar{C} = \begin{bmatrix} I_h & 0 & R_3 \\ Q_{222} & I_{\bar{n}-p} & R_6 \\ 0 & 0 & R_9 \end{bmatrix} \begin{bmatrix} A_\Omega & 0 & 0 \\ \bar{Q}_1 & 0 & \bar{A}_3 \\ T_p\bar{Q}_2 & T_p\bar{A}_5 & T_p\bar{A}_6 \end{bmatrix} - \begin{bmatrix} G_{l1} \\ G_{l2} \\ G_{l3} \end{bmatrix} \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} = \left[\begin{array}{cc|c} \tilde{A}_\Omega & 0 & * \\ \tilde{Q}_1 & -J & * \\ \hline * & * & * \end{array}\right], \tag{6.99}$$

where $*$ represents terms that do not contribute to succeeding analysis. From (6.64), it can be seen that the top-left partition of $R\bar{A} - G_l\bar{C}$ in (6.99) is equivalent to $\tilde{A}$. The structure of P from Propositions 6.2 and 6.3, and X from (6.98) therefore imply that

$$\begin{aligned} X &= P(R\bar{A} - G_l\bar{C}) + (R\bar{A} - G_l\bar{C})^T P \\ &= \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} \begin{bmatrix} \tilde{A} & * \\ * & * \end{bmatrix} + \begin{bmatrix} \tilde{A}^T & * \\ * & * \end{bmatrix} \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} = \begin{bmatrix} P_1\tilde{A}_1 + \tilde{A}_1^T P_1 & X_1^T \\ X_1 & X_2 \end{bmatrix}. \end{aligned} \tag{6.100}$$

Since $X < 0$, $P_1\tilde{A}_1 + \tilde{A}_1^T P_1 < 0$. Thus both inequalities in (6.91) are satisfied, and sliding motion is guaranteed in both observers. Lastly, by substituting P_{11} and P_{34} from (6.96) into LMI (6.94), it can be seen that (6.94) has the same form as (6.90), which describes the Bounded Real Lemma bounding the $\mathcal{L}_2$ gain from ξ to e_f. Thus, the proof is complete. ∎

Remark 6.6 *Compared to the LMIs presented in [17], the formulation in this chapter has enabled the LMI pair* (6.93)–(6.94) *to simultaneously calculate the gains of both observers in cascade. This is because the formulation in the paper has caused the system from ξ to both e_{f1} and e_{f2} to have the same 'state equation' - this can be verified by comparing* (6.86)–(6.87) *with* (6.59)–(6.63) *and* (6.79)–(6.84). *If the approach in [17] is used, the LMI pair* (6.93)–(6.94) *would calculate the gain of only one observer, and the gain of the second observer will have to be calculated using another LMI pair [91], or through a different method [129]. This may not be optimal, causing γ to be larger than necessary and hence be conservative. Therefore, the LMI design method in this chapter reduces the conservatism by design. The only remaining conservatism is caused by the block-diagonal structure of P, which is required in Propositions 6.2 and 6.3 to reconstruct the faults.* ♯

6.3.4 Existence conditions

Theorem 6.1 *The proposed cascaded SMO scheme can reconstruct f_0 for system* (6.1)–(6.2) *if and only if the following conditions hold:*

N1. $\operatorname{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \operatorname{rank}(M_2) = n - p,$

N2. $p + q - \bar{n} \leq \mathrm{rank}(M_2) \leq 2p - n$. $\sharp$

Proof The remainder of this subsection forms the constructive proof for Theorem 6.1.

The following assumptions were made during the analysis of the cascaded SMO scheme in the preceding subsections:

P1. $\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p$ in (6.9), so that T_{f1} in Proposition 6.1 exists such that $T_{f1}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}$.

P2. Matrix T_{44} in (6.55) exists such that it has the dimensions $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$.

P3. $\mathrm{rank}(M_{11}) = q - q_2$, so that $M_{11}^{\dagger}$ in (6.81) exists such that $M_{11}^{\dagger}M_{11} = \begin{bmatrix} 0 \\ I_{q-q2} \end{bmatrix}$.

P4. The error systems (6.53)–(6.54) (and therefore (6.59)–(6.60)) and (6.79)–(6.80) can be made stable.

Therefore, it is of interest to re-express these conditions in terms of the original system matrices so that it is easier for designers to verify whether the proposed scheme is applicable from the outset.

Proposition 6.5 *Condition P1 is satisfied if and only if N1 is satisfied.* $\sharp$

Proof Pre-multiply (M_0, A_0) with T_dT_b and post-multiply A_0 and M_0 with $(T_cT_a)^{-1}$ and T_e^{-1} in Proposition 6.1 to obtain

$$M_2 = \begin{bmatrix} 0 & I_{q2} \\ 0 & 0 \end{bmatrix}, \; A_4 = \begin{bmatrix} A_{41} \\ A_{42} \end{bmatrix}, \; A_5 = \begin{bmatrix} A_{51} \\ A_{52} \end{bmatrix}. \tag{6.101}$$

Suppose N1 is not satisfied, i.e.

$$\mathrm{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \mathrm{rank}(M_2) < n - p. \tag{6.102}$$

Equation (6.102) implies

$$\mathrm{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} < n - p, \tag{6.103}$$

that is, P1 is also not satisfied, which shows the necessity of N1. If N1 is satisfied, however, it is straightforward to deduce from the structure of $\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix}$ in (6.101) that

$$\operatorname{rank}\begin{bmatrix} A_{42} & A_{52} \end{bmatrix} = n - p, \tag{6.104}$$

which shows that P1 is satisfied, thus proving the sufficiency of N1. ■

Proposition 6.6 *Condition N2 is necessary and sufficient for P2–P3 to be satisfied.* ♯

Proof Condition N2 is shown to be necessary through the two cases where it is not satisfied, i.e.

$$\operatorname{rank}(M_2) > 2p - n, \tag{6.105}$$

$$\operatorname{rank}(M_2) < p + q - \bar{n}. \tag{6.106}$$

Recall from (6.7) that $q_2 = \operatorname{rank}(M_2)$, and from (6.55) that $T_{44} \in \mathbb{R}^{(2p-n-q_2)\times(2p-\bar{n}-q_2)}$, where $q_2 = \operatorname{rank}(M_2)$. Equation (6.105) implies T_{44} does not exist, and P2 is not satisfied. Next, (6.106) implies

$$\bar{n} - p < q - q_2, \tag{6.107}$$

Since $M_{11} \in \mathbb{R}^{(\bar{n}-p)\times(q-q_2)}$,

$$\operatorname{rank}(M_{11}) \le \min\{\bar{n} - p, q - q_2\} \le \bar{n} - p < q - q_2, \tag{6.108}$$

that is, M_{11} is not full-column rank, and P3 can never be satisfied, thus showing the necessity of N2. On the other hand, if N2 is satisfied, since $n - \bar{n} > 0$, it is straightforward to see that T_{44} can be assigned the dimensions in (6.55), and M_{11} would be full-column rank, satisfying P2–P3 and showing the sufficiency of N2. ■

Corollary 6.2 *Condition P4 is always satisfied.* ♯

Proof The eigenvalues of system (6.53) are $\{\lambda(\tilde{A}_\Omega), \lambda(J)\}$. Recall from (6.3) and (6.56) that

$$\lambda\left(\tilde{A}_\Omega\right) = \lambda\left(A_\Omega + R_{34}T_{44}Q_{223}\right), \ \lambda\left(A_\Omega\right) < 0, \tag{6.109}$$

where $R_{34} \in \mathbb{R}^{h\times(2p-n-q_2)}$. Condition N2 implies

$$2p - n - q_2 \ge 0. \tag{6.110}$$

In the case $2p - n - q_2 = 0,\ R_{34}$ does not exist, and hence

$$\lambda\left(\tilde{A}_\Omega\right) = \lambda\left(A_\Omega\right) < 0. \tag{6.111}$$

If $2p - n - q_2 > 0$, then R_{34} can be freely chosen such that $\lambda\left(\tilde{A}_\Omega\right) < 0$. Therefore $\lambda\left(\tilde{A}_\Omega\right) < 0$ is satisfied for all cases. It is also required that $J \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$ exists such that $\lambda\left(J\right) > 0$ for (6.53)–(6.54) (and therefore (6.59)–(6.60)) to be stable. If $\bar{n} = p$ (all states of system (6.29)–(6.30) are measurable outputs), then e_2 (and therefore J) does not exist. If $\bar{n} > p$ then J exists and can be freely chosen such that $\lambda\left(J\right) > 0$. Therefore error system (6.53)–(6.54) (and therefore (6.59)–(6.60)) can always be made stable.

To stabilise error system (6.79)–(6.80), the zeros of $(\tilde{A}, \tilde{M}, \tilde{C})$ need to be stable, i.e. $D(s)$ is full rank for all $s \in \mathbb{C}^+$, where

$$D(s) = \begin{bmatrix} sI - \tilde{A} & \tilde{M} \\ \tilde{C} & 0 \end{bmatrix} = \left[\begin{array}{cc|c} sI_h - \tilde{A}_\Omega & 0 & 0 \\ -\tilde{Q}_1 & sI_{\bar{n}-p} + J & -M_{11} \\ \hline 0 & I_{\bar{n}-p} & 0 \end{array}\right]. \tag{6.112}$$

Since $\operatorname{rank}\left(\tilde{A}_\Omega\right) < 0$ and $\operatorname{rank}(M_{11}) = q - q_2 \leq \bar{n} - p$ (from N2), $D(s)$ will always be full-column rank for $s \in \mathbb{C}^+$ and therefore system (6.79)–(6.80) can always be stabilised. Thus P4 is always satisfied. ■

Thus, N1–N2 are shown to be necessary and sufficient for P1–P4 to be satisfied (and consequently no additional conditions are required). Therefore, Theorem 6.1 has been proven. □

The strengths of the proposed scheme over existing work (and hence the contribution of this chapter) are summarised as follows:

- The scheme considers disturbances (that will corrupt the fault reconstruction) and makes the fault reconstruction robust against them. Disturbances were not considered in the prior schemes in [99, 100, 101, 161].

- The proposed scheme is able to asymptotically reconstruct the faults (in the absence of disturbances) for a larger class of non-infinitely observable descriptor systems compared to the works in [17, 99, 100, 101, 161].

- Only a single LMI pair (i.e. (6.93)–(6.94)) is required to design the gains of both observers simultaneously, which reduces conservatism when compared to previous works [17, 91, 128, 129] where each observer was designed separately.

6.3.5 Design procedure

The design procedure for the proposed cascaded observer scheme is as follows:

1. Determine T_a, T_b, T_c in Proposition 6.1, and apply the state equation transformation T_b and apply the state transformation $T_{aft}^{-1} = T_c T_a$.

2. Check that N1–N2 hold; if they are not satisfied then the proposed scheme is not applicable.

3. Calculate $T_{pre} = T_g T_f T_d T_b$ and T_e in Proposition 6.1, and apply the state equation transformation T_{pre}, the state transformation T_{aft}^{-1}, and fault transformation T_e.

4. Determine suitable values for (A_Ω, B_Ω) in (6.3)–(6.4) based on knowledge of the disturbance ξ_0.

5. Augment the dynamics of ξ_0 into the system using (6.19)–(6.20). Then determine T_h in (6.23), and apply the state transformation T_h.

6. Set a value for T_{44} in (6.55) such that $T_{44}Q_{223}$ has as many non-zero columns as possible. Apply the state equation transformation T given in (6.25).

7. Form the reduced-order system in (6.29)–(6.30).

8. Choose values for LMI parameters $\bar{D}$ and γ_0 in Proposition 6.4, and use a LMI solver to determine P and γ from (6.93)–(6.95). Note that the gains for both observers are simultaneously designed here.

9. Using (6.96), determine R and V from (6.37), G_l and G_n from (6.34), L from (6.61), $\tilde{G}_n$ from (6.68), and K from (6.82).

10. Set ρ and $\tilde{\rho}$ to satisfy (6.45) and (6.69), respectively, and determine $M_{11}^{\dagger}$ from (6.82).

11. Reconstruct f according to (6.61) and (6.82).

6.4 Simulation example

The effectiveness of the proposed scheme is demonstrated through the following example: consider a modified version of the generalised RLC circuit in [135] described by the following dynamical model:

$$L_1\dot{I}_{L1} = -R_1 I_{L1} - V_{C1} + V_1, \tag{6.113}$$

$$C_1\dot{V}_{C1} = I_{L1}, \tag{6.114}$$

$$L_2\dot{I}_{L2} = -R_2 I_{L2} - V_{C2} + V_2, \tag{6.115}$$

$$C_2\dot{V}_{C2} = I_{L2}, \tag{6.116}$$

$$L_{12}\dot{I}_{L12} = -R_{12} I_{L12} - V_{C12} + V_1 - V_2, \tag{6.117}$$

$$C_{12}\dot{V}_{C12} = I_{L12}, \tag{6.118}$$

$$0 = -I_{L1} - I_{L12} + I, \tag{6.119}$$

$$0 = -I_{L2} + I_{L12}, \tag{6.120}$$

where I is the input current, $I_{L1},\ I_{L2},\ I_{L12}$ are the currents flowing across the inductors $L_1,\ L_2$, and L_{12}, respectively, $V_{C1},\ V_{C2}$, and V_{C12} are the voltage drops across the capacitors $C_1,\ C_2$, and C_{12} respectively, and V_1 and V_2 are the voltages at nodes 1 and 2, respectively. The inductor L_1 is faulty, causing fluctuations in the voltage drop and current across it (labelled as f_a and f_b respectively), and hence (6.113)–(6.114) become

$$L_1\dot{I}_{L1} = -R_1 I_{L1} - V_{C1} + V_1 - f_a, \tag{6.121}$$

$$C_1\dot{V}_{C1} = I_{L1} - f_b, \tag{6.122}$$

Let the components have the following values:

$$L_1 = L_2 = L_{12} = 1\text{ H},\ C_1 = C_2 = C_{12} = 1\text{ F},\ R_1 = R_2 = R_{12} = 1\ \Omega. \tag{6.123}$$

6.4.1 System formulation

The system matrices (E_0, A_0, B_0, M_0) in the framework of (6.1)–(6.2) are

$$E_0 = \begin{bmatrix} I_6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A_0 = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},\ B_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},\ M_0 = \begin{bmatrix} -I_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{6.124}$$

for the system variables $u = \begin{bmatrix} \text{input current, } I \text{ (A)} \end{bmatrix}$ and

$$x_0 = \begin{bmatrix} \text{current across } L_1,\ I_{L1} \text{ (A)} \\ \text{voltage across } C_1,\ V_{L1} \text{ (V)} \\ \text{current across } L_2,\ I_{L2} \text{ (A)} \\ \text{voltage across } L_2,\ V_{L2} \text{ (V)} \\ \text{current across } L_{12},\ I_{L12} \text{ (A)} \\ \text{voltage across } C_{12},\ V_{L12} \text{ (V)} \\ \text{voltage at node } 1,\ V_1 \text{ (V)} \\ \text{voltage at node } 2,\ V_2 \text{ (V)} \end{bmatrix},\ f_0 = \begin{bmatrix} f_a \text{ (A/s)} \\ f_b \text{ (V/s)} \end{bmatrix}. \tag{6.125}$$

To demonstrate the robustness of the observer scheme, assume that the currents flowing across capacitors C_1 and C_{12} fluctuate by ξ_1 and ξ_2. Hence define the disturbance signal $\xi_0 = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$, and assume that measurements are only available for $V_{C1},\ V_{C2},\ I_{L12},\ V_{C12},$

and V_1, so Q_0 and C_0 have the form

$$Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \; C_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_4 & 0 \end{bmatrix}. \tag{6.126}$$

6.4.2 Design of observers

To ease readability, the steps to design the cascaded observer scheme in §6.3.5 will be referred to in the following subsection.

Step 1: To obtain the system in the form (6.10)–(6.12), apply the state equation transformation T_b and state equation transformation T_{aft}^{-1}, where

$$T_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_5 \end{bmatrix}, \; T_{aft} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_4 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.127}$$

Step 2: The following partitions are obtained:

$$M_2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \; A_4 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; A_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{6.128}$$

From (6.128), the following can then be verified:

$$\mathrm{rank}\begin{bmatrix} M_2 & A_4 & A_5 \end{bmatrix} - \mathrm{rank}(M_2) = 3, \; p + q - \bar{n} \le \mathrm{rank}(M_2) \le 2p - n. \tag{6.129}$$

Equation (6.129) shows that N1–N2 hold for the system, and therefore guarantees the existence of the proposed observer scheme. The observer scheme proposed in the chapter is now designed.

Step 3: Apply the state equation transformation T_{pre}, the state transformation T_{aft}^{-1}, and fault transformation T_e, where

$$T_{pre} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \; T_e = -I_2. \tag{6.130}$$

Step 4: The disturbance ξ_0 is assumed to have a frequency content of $\omega < 2$ rad/s. Therefore, choose $\omega_f = 20$ rad/s, and hence

$$A_\Omega = -20I_2, \; B_\Omega = 20I_2. \tag{6.131}$$

Step 5: The dynamics of ξ_0 are augmented into the system using (6.18)–(6.20). Finally, the state transformation T_h is applied onto the augmented system, where

$$T_h = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_5 \end{bmatrix}. \tag{6.132}$$

The following augmented system in the coordinates of (6.24) is obtained:

$$E = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; M = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \; Q = \begin{bmatrix} 20I_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$A = \begin{bmatrix} 0 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \; C = \begin{bmatrix} 0 & I_5 \end{bmatrix}. \tag{6.133}$$

It can be seen from (6.133) that $Q_{223} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$; hence for T_p in (6.55), set $T_{44} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ (so $T_{44}Q_{223} = \begin{bmatrix} -1 & -1 \end{bmatrix}$ has non-zero columns). The reduced-order system (6.29)–(6.30) is then found to be

$$\bar{E} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \bar{M} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & I_2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ \bar{Q} = \begin{bmatrix} 20I_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\bar{A} = \begin{bmatrix} -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \end{bmatrix}, \ \bar{C} = \begin{bmatrix} 0 & I_5 \end{bmatrix}. \tag{6.134}$$

The LMI design parameters in Proposition 6.4 were chosen as $\bar{D} = I_2$ and $\gamma_0 = 1$. By using the LMI Control Toolbox within MATLAB on LMIs (6.93)–(6.94), the following values of P and γ were obtained:

$$P = \begin{bmatrix} 0.0053 & -0.0043 & 0 & 0 & 0 \\ -0.0043 & 0.0053 & 0 & 0 & 0 \\ 0 & 0 & 0.4687 & -0.001 & 0 \\ 0 & 0 & -0.001 & 0.002 & 0 \\ 0 & 0 & 0 & 0 & 0.7577I_5 \end{bmatrix}, \ \gamma = 0.0956. \tag{6.135}$$

The observer parameters are found to be

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 168.9 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 169.8 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0.7145 & 0.4050 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 1.463 & 186.9 & 0 \\ 0 & 0 & 0 & 0 & 0.4655 & 0.022 & -0.5584 & -0.0110 & -0.0553 \\ 0 & 0 & 0 & 0 & 0.0220 & -0.0106 & -0.1517 & 0.0730 & -0.0482 \\ 0 & 0 & 0 & 0 & -0.5584 & -0.1517 & 1.110 & -0.0038 & 0.4562 \\ 0 & 0 & 0 & 0 & -0.0110 & 0.0730 & -0.0038 & -0.0042 & 0.0105 \\ 0 & 0 & 0 & 0 & -0.0553 & -0.0482 & 0.4562 & 0.0105 & 0.0892 \end{bmatrix},$$

$$V = \begin{bmatrix} 0 & -168.9 & 0 & -168.9 & 0 \\ 0 & -169.8 & 0 & -169.8 & 0 \\ 0 & 0.4050 & 0 & 0 & 0 \\ 0 & 185.9 & -1 & 0 & 0 \\ 0.5345 & 0.0443 & 0.0220 & 0.0553 & 0 \\ -0.0220 & 1.121 & -0.0106 & 0.0482 & 0 \\ 0.5584 & -0.4600 & 0.8483 & -0.4562 & 0 \\ 0.0110 & -0.0147 & 0.0730 & 0.9895 & 0 \\ 0.0553 & -0.0787 & -0.0482 & -0.0892 & 1 \end{bmatrix}, \; L = 0, \; K = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{6.136}$$

$$G_l = \begin{bmatrix} 0_{4\times5} \\ 0.6599 I_5 \end{bmatrix}, \; G_n = \begin{bmatrix} 0_{4\times5} \\ 1.320 I_5 \end{bmatrix}, \; \tilde{G}_n = \begin{bmatrix} \multicolumn{2}{c}{0_{2\times2}} \\ 2.136 & 1.050 \\ 1.050 & 490.6 \end{bmatrix}.$$

The poles of Observer One are then calculated to be $\{-357.9, -187.0, -20, -0.9761 \pm j1.113, -0.6371 \pm j0.109, -0.6599, -0.6594\}$, while the poles of its sliding motion (which are also the poles for Observer Two) are $\{-358.8, -187.0, -20, -0.7114\}$. The poles for the sliding motion of Observer Two are $\{-358.8, -20\}$. The parameters ρ, $\tilde{\rho}$, and $M_{11}^{\dagger}$ are as follows:

$$\rho = \tilde{\rho} = 5, \; M_{11} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{6.137}$$

Remark 6.7 *The design of the observers (i.e. the scheme in this chapter) is complete. It will now be shown that existing fault reconstruction schemes for descriptor systems will not work for the example in* (6.124) *and* (6.126) *as follows:*

- *The system is not infinitely observable, i.e.* $\text{rank}\begin{bmatrix} E_0 \\ C_0 \end{bmatrix} = 7 < 8$*, and therefore the schemes in [5, 34, 39, 143, 167, 168] are inapplicable.*

- *From* (6.128), $\text{rank}(M_2) = 1 < 2$*. Thus, the schemes utilising a single SMO in [17, 100, 101] cannot reconstruct all components of* f_0*, and are therefore inapplicable.*

- *In the coordinates of* (6.130)*, the following partitions are obtained:*

$$\begin{bmatrix} \bar{E}_{21} \\ E_{221} \\ E_{222} \\ E_{223} \end{bmatrix} = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right], \; A_{41} = 0. \tag{6.138}$$

Therefore, it can then be shown that

$$\text{rank}\begin{bmatrix} \bar{E}_{21} \\ E_{222} \\ E_{223} \end{bmatrix} = 3 \neq r - \bar{n} + p, \; \text{rank}\begin{bmatrix} E_{221} \\ E_{222} \\ E_{223} \end{bmatrix} \neq \text{rank}\begin{bmatrix} E_{222} \\ E_{223} \end{bmatrix}, \tag{6.139}$$

that is, M1–M2 are not satisfied, and therefore the schemes in [99, 161] are also inapplicable to the example. In fact, for the cascaded SMO schemes in [99, 161] to be applicable, the eighth state of (6.125) *needs to be measurable (which also results in the system becoming infinitely observable). Even then, the schemes utilising a single SMO in [17, 100, 101] are still unable to reconstruct all components of* f *since* $\text{rank}(M_2) < \text{rank}(M)$ *in this case.* ♯

6.4.3 Simulation results

To show the effectiveness of the proposed scheme at bounding the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction, the disturbance signals are set as

$$\xi_1 = 0.2\sin\left(t + \frac{\pi}{2}\right) + 0.3, \; \xi_2 = 0.1\sin\left(0.8t + \frac{\pi}{6}\right) + 0.2. \tag{6.140}$$

The initial condition of the system is set as $\{-1, 2, 1, 5, 1, 3, 4.45, 3.225\}$, while the observers are set to have zero initial conditions. The case where the current source is disconnected is simulated, i.e. $I = 0$. The fault signals are simulated as

$$f_1 = 0.2 \sin\left(2t + \frac{\pi}{3}\right) + 0.6, \;\; \dot{f}_2 = -10 f_2 + 10\bar{f}, \tag{6.141}$$

where the signal $\bar{f}$ is given by

$$\bar{f} = \begin{cases} 0.05 & \text{for } t \leq 0.5, \\ t - 0.45 & \text{for } 0.5 < t \leq 1.2, \\ 0.75 & \text{for } 1.2 < t \leq 2, \\ -1.5t + 3.75 & \text{for } 2 < t \leq 2.3, \\ 0.3 & \text{for } t > 2.3. \end{cases} \tag{6.142}$$

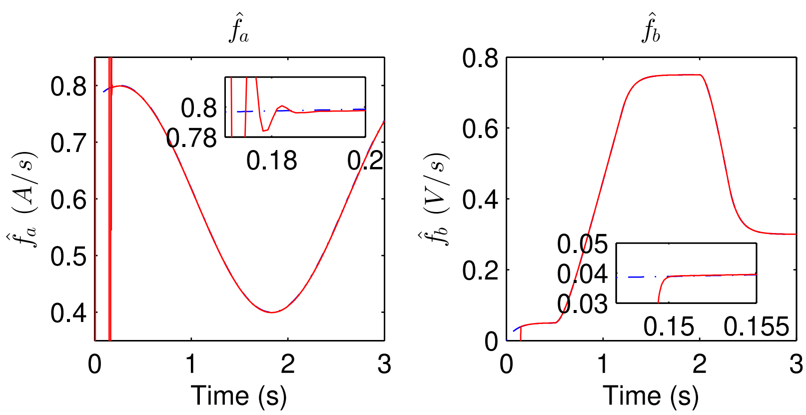

Figure 6.2: Faults (dash-dotted) and their reconstructions (solid).

Figure 6.2 shows the fault signals and their reconstructions, whereas Figure 6.3 shows the fault reconstruction errors. Notice that $\hat{f}_b$ converges onto f_b at around $t = 0.15$ s, while $\hat{f}_a$ converges onto f_a at around $t = 0.19$ s. It can be seen that fault reconstruction is achieved even in the presence of disturbances (due to the small value of γ), and that the magnitudes of the fault reconstruction errors are bounded within $\pm\gamma\|\xi_0\|$. Thus the method is verified as effective at bounding the effect of the disturbances on the fault reconstructions.

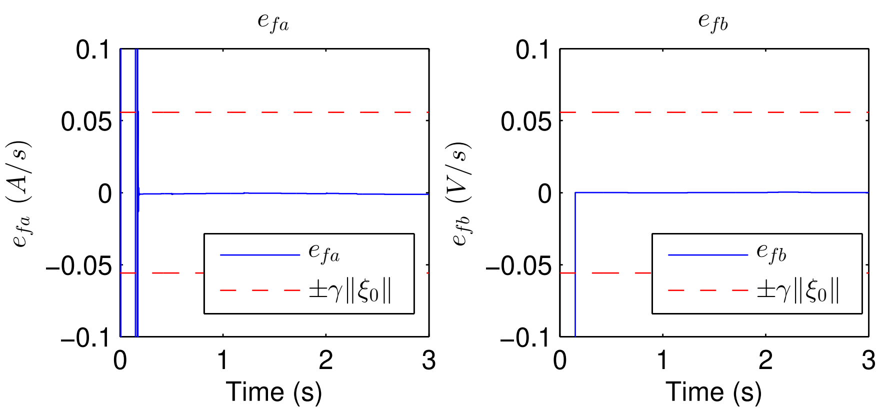

Figure 6.3: Fault reconstruction errors (solid). The dashed lines represent the upper bounds of e_f derived from LMIs (6.93)–(6.94).

6.5 Conclusion

This chapter has presented a robust fault scheme consisting of two cascaded SMOs for a class of NIODS. Certain states were treated as unknown inputs to form an infinitely observable system. A standard SMO was then applied onto this reduced-order system to reconstruct a component of the fault, and its switching term is fed into a second SMO to reconstruct the remaining faults. LMI techniques were employed to minimise the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction. The necessary and sufficient conditions for the existence of the scheme were investigated and are found to be more relaxed than those for existing fault reconstruction schemes for non-infinitely observable descriptor systems. The LMIs were formulated such that the gains for both observers are designed simultaneously, and are thus less conservative than existing methods, which design each observer separately. Finally, a simulation was carried out and its results verify the efficacy of the scheme.

Chapter 7

New results in robust fault reconstruction for NIODS

7.1 Introduction

In the previous chapters, several schemes for reconstructing faults in non-infinitely observable descriptor systems (NIODS) using sliding mode observers (SMO) have been presented. These schemes investigated in these chapters however did not fully exploit all design freedom when treating certain states as unknown inputs. In addition to treating certain states as unknown inputs, Ooi et al. [101] proposed a scheme which also re-expressed certain states as a linear combination of other states, thus reducing the number of states that are treated as unknown inputs and reducing the number of sensors potentially required. Their scheme however did not consider the effect of disturbances on the fault reconstruction. The contribution of this chapter is therefore to build on the work by Ooi et al. [101] by developing a fault reconstruction scheme for a class of NIODS such that the $\mathcal{L}_2$ gain from the external disturbances to the fault reconstruction is minimised, thereby achieving *robust fault reconstruction*.

The chapter starts off by introducing the preliminary transformations to re-express the descriptor system in a form that facilitates analysis and gain design in §7.2. Some states are re-expressed in terms of other states, and certain other states are treated as unknown inputs, thereby forming an infinitely observable reduced-order system; this was not done in the previous chapters. In §7.3, the observer by Yeu et al. [157] is applied onto the reduced-order system. The design parameters are then designed using the Bounded Real Lemma [128] such that the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction is minimised. The necessary and sufficient existence conditions for the observer are studied and presented in terms of the original system matrices. A summarised design procedure for the scheme is shown. Finally, a simulation example is performed in §7.4 to demonstrate the efficacy of the proposed scheme.

The work in this chapter has been accepted for presentation in the *American Control Conference (ACC) 2019*.

7.2 Preliminary transformations

Consider the non-infinitely observable descriptor system (NIODS):

$$E_0\dot{x}_0 = A_0x_0 + B_0u + M_0f_0 + Q_0\xi, \tag{7.1}$$

$$y = C_0x_0, \tag{7.2}$$

where $E_0, A_0 \in \mathbb{R}^{n\times n},\ B_0 \in \mathbb{R}^{n\times m},\ M_0 \in \mathbb{R}^{n\times q},\ Q_0 \in \mathbb{R}^{n\times h},\ C_0 \in \mathbb{R}^{p\times n}$, while $x_0,\ u,\ y,\ f_0$, and ξ are the states, inputs, outputs, faults, and disturbance signals, respectively. Only u and y are measurable. The fault signal f_0 represents irregular behaviour in the system which needs to be reconstructed so that information on its shape and magnitude can be extracted for purposes of timely and precise corrective action. The unknown disturbance signal ξ (which is not a fault, and could arise due to errors in modelling or from parasitic dynamics [95]) may corrupt the fault reconstruction, resulting in erroneous estimates of the fault. These inaccurate reconstructions could lead to false positives being raised, or even worse, false negatives where the effects of the faults are masked. Thus it is of interest to reconstruct f_0 while minimising the effect of ξ on its reconstruction. It is assumed that E_0 is rank-deficient, i.e. $\mathrm{rank}(E_0) = r < n$, and that M_0 and C_0 are full-column rank and full-row rank, respectively, i.e. $\mathrm{rank}(M_0) = q,\ \mathrm{rank}(C_0) = p$.

Similar to the previous chapters, system (7.1)–(7.2) is first re-expressed into a form facilitating further analysis and observer gain design using Lemma 4.1 and Proposition 7.1. Define a non-singular matrix $T_a = \begin{bmatrix} N_C^T \\ C_0 \end{bmatrix}$, where $C_0N_C = 0$, which results in

$$E_0 \mapsto E_0T_a^{-1} = \begin{bmatrix} E_{a1} & E_{a2} \end{bmatrix},\ C_0 \mapsto C_0T_a^{-1} = \begin{bmatrix} 0 & I_p \end{bmatrix}. \tag{7.3}$$

Since $\mathrm{rank}\begin{bmatrix} E_0 \\ C_0 \end{bmatrix} = \bar{n}$, it follows from (7.3) that $\mathrm{rank}(E_{a1}) = \bar{n} - p$. Using singular-value decomposition (SVD), E_{a1} can be decomposed as

$$T_b E_{a1} T_{c1} = \begin{bmatrix} 0 & I_{\bar{n}-p} \\ 0 & 0 \end{bmatrix}, \tag{7.4}$$

where T_b and T_{c1} are non-singular matrices. Define another non-singular matrix $T_c = \begin{bmatrix} T_{c1}^{-1} & 0 \\ 0 & I_p \end{bmatrix}$, and let (E_0, A_0, M_0) in the coordinates of (7.1) be transformed to have the following structures:

$$\begin{aligned} T_b E_0 (T_c T_a)^{-1} &= \begin{bmatrix} 0 & I_{\bar{n}-p} & E_1 \\ 0 & 0 & E_{c2} \end{bmatrix}, \ T_b M_0 = \begin{bmatrix} M_1 \\ M_{c2} \end{bmatrix} \\ T_b A_0 (T_c T_a)^{-1} &= \begin{bmatrix} A_1 & A_2 & A_3 \\ A_{c4} & A_{c5} & A_{c6} \end{bmatrix}. \end{aligned} \tag{7.5}$$

From $r = \text{rank}(E_0)$ and (7.5), it can be seen that $\text{rank}(E_{c2}) = r - \bar{n} + p$. Therefore there exists an invertible matrix T_{d1} such that

$$T_{d1} E_{c2} = \begin{bmatrix} E_2 \\ 0 \end{bmatrix}, \tag{7.6}$$

where $E_2 \in \mathbb{R}^{(r-\bar{n}+p)\times p}$ is full-row rank. Therefore, define an invertible matrix $T_d = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_{d1} \end{bmatrix}$, and let (A_0, M_0) in the coordinates of (7.1) be transformed to have the following structures:

$$T_d T_b A_0 (T_c T_a)^{-1} = \begin{bmatrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \\ A_7 & A_8 & A_9 \end{bmatrix}, \ T_d T_b M_0 = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix}, \tag{7.7}$$

where $A_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$, $A_5 \in \mathbb{R}^{(r-\bar{n}+p)\times(\bar{n}-p)}$, $A_9 \in \mathbb{R}^{(n-r)\times p}$. Define

$$a = \text{rank}(A_7) \leq n - \bar{n}, \tag{7.8}$$

and therefore apply SVD on A_7 as follows:

$$T_{e1} A_7 T_{f1} = \begin{bmatrix} 0 & A_{71} \\ 0 & 0 \end{bmatrix}, \tag{7.9}$$

where $A_{71} \in \mathbb{R}^{a\times a}$, T_{e1}, and T_{f1} are non-singular. Next, let

$$T_{e1} M_3 = \begin{bmatrix} M_{31} \\ M_{32} \end{bmatrix}, \tag{7.10}$$

where $M_{31} \in \mathbb{R}^{a \times q}$ and $M_{32} \in \mathbb{R}^{(n-r-a) \times q}$. Define $q_{21} = \text{rank}(M_{31})$ and $q_{22} = \text{rank}(M_{32})$, where $q_2 = q_{21} + q_{22}$. Therefore, performing SVD on M_{31} and M_{32} yields

$$\begin{aligned} T_{e2} M_{31} T_g^{-1} &= \begin{bmatrix} 0 & -I_{q21} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \updownarrow \; q_{21} \\ \updownarrow \; a - q_{21} \end{array}, \\ T_{e3} M_{32} T_g^{-1} &= \begin{bmatrix} 0 & 0 & I_{q22} \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \updownarrow \; q_{22} \\ \updownarrow \; n - r - a - q_{22} \end{array}. \end{aligned} \tag{7.11}$$

Define an invertible matrix

$$T_{f2} = \begin{bmatrix} I_{n-\bar{n}-a} & 0 \\ 0 & (T_{e2} A_{71})^{-1} \end{bmatrix}, \tag{7.12}$$

and therefore by substituting for structures of $A_1, A_4, A_7, M_1, M_2, M_{31}, M_{32}$ from (7.7)–(7.11), the following is obtained:

$$\begin{bmatrix} I_r & 0 & 0 \\ 0 & T_{e2} & 0 \\ 0 & 0 & T_{e3} \end{bmatrix} \left[\begin{array}{c|c} A_1 & M_1 \\ \hline A_4 & M_2 \\ \hline \multirow{2}{*}{A_7} & M_{31} \\ & M_{32} \end{array}\right] \begin{bmatrix} T_{f1} T_{f2} & 0 \\ 0 & T_g^{-1} \end{bmatrix}$$
$$= \left[\begin{array}{ccc|ccc} A_{11} & A_{12} & A_{13} & M_{11} & M_{12} & M_{13} \\ \hline A_{41} & A_{42} & A_{43} & M_{21} & M_{22} & M_{23} \\ \hline 0 & I_{q21} & 0 & 0 & -I_{q21} & 0 \\ 0 & 0 & I_{a-q21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{q22} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{7.13}$$

Proposition 7.1 *Assume that* $\begin{bmatrix} A_{41} & M_{21} \end{bmatrix}$ *is full-column rank, i.e.*

$$\text{rank}\begin{bmatrix} A_{41} & M_{21} \end{bmatrix} = n - \bar{n} - a + q_1, \tag{7.14}$$

where $q_1 = q - q_{21} - q_{22}$ *(the satisfaction of* (7.14) *will be addressed later in Proposition 7.4). Therefore, there exists a non-singular matrix* T_{h1} *such that*

$$T_{h1} \begin{bmatrix} A_{41} & A_{42} & M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} 0 & A_{421} & I_{q1} & M_{221} \\ I_{n-\bar{n}-a} & A_{422} & 0 & M_{222} \\ 0 & A_{423} & 0 & M_{223} \end{bmatrix}. \tag{7.15}$$

where $M_{223} \in \mathbb{R}^{(p-q-(n-r-a-q_2))\times q_{21}}$. *Let* $\tilde{A}_m = A_{423} + M_{223}$, *and assume that* $\tilde{A}_m$ *is full-column rank, i.e.*

$$\operatorname{rank}(\tilde{A}_m) = q_{21}, \tag{7.16}$$

where the satisfaction of (7.16) *will also be addressed later in Proposition 7.4. Then there exists a set of transformations such that system* (7.1)–(7.2) *can be re-expressed as follows:*

$$\underbrace{\begin{bmatrix} 0 & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & \tilde{E}_2 \\ 0 & 0 & \bar{E}_3 \\ 0 & 0 & \tilde{E}_4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{E} \dot{x} = \underbrace{\begin{bmatrix} 0 & \bar{A}_2 & \bar{A}_3 \\ \tilde{A}_4 & \tilde{A}_5 & \tilde{A}_6 \\ \bar{A}_7 & \bar{A}_8 & \bar{A}_9 \\ 0 & \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \bar{A}_{14} & \bar{A}_{15} \\ \bar{A}_{16} & \bar{A}_{17} & \bar{A}_{18} \end{bmatrix}}_{A} x + \underbrace{\begin{bmatrix} \bar{B}_1 \\ \tilde{B}_2 \\ \bar{B}_3 \\ \tilde{B}_4 \\ \bar{B}_5 \\ \bar{B}_6 \end{bmatrix}}_{B} u + \underbrace{\begin{bmatrix} 0 \\ \tilde{M}_2 \\ 0 \\ 0 \\ 0 \\ \bar{M}_6 \end{bmatrix}}_{M} f + \underbrace{\begin{bmatrix} \bar{Q}_1 \\ \tilde{Q}_2 \\ \bar{Q}_3 \\ \tilde{Q}_4 \\ \bar{Q}_5 \\ \bar{Q}_6 \end{bmatrix}}_{Q} \xi, \tag{7.17}$$

$$y = \underbrace{\begin{bmatrix} 0 & 0 & I_p \end{bmatrix}}_{C} x, \quad x = \begin{bmatrix} x_{11} \\ x_{121} \\ x_{122} \\ \hline x_2 \\ \hline y \end{bmatrix} \begin{matrix} \updownarrow & n-\bar{n}-a \\ \updownarrow & q_{21} \\ \updownarrow & a-q_{21} \\ \updownarrow & \bar{n}-p \\ \updownarrow & p \end{matrix}, \quad f = \begin{bmatrix} f_1 \\ f_{21} \\ f_{22} \end{bmatrix} \begin{matrix} \updownarrow & q_1 \\ \updownarrow & q_{21} \\ \updownarrow & q_{22} \end{matrix}, \tag{7.18}$$

$$\begin{bmatrix} \tilde{A}_4 & \tilde{A}_5 & \tilde{M}_2 \\ \bar{A}_7 & \bar{A}_8 & 0 \\ \bar{A}_{16} & \bar{A}_{17} & \bar{M}_6 \end{bmatrix} = \left[\begin{array}{ccc|c|ccc} 0 & 0 & 0 & \tilde{A}_{51} & I_{q1} & 0 & 0 \\ 0 & I_{q21} & 0 & \tilde{A}_{52} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{A}_{53} & 0 & 0 & I_{q22} \\ \hline I_{n-\bar{n}-a} & 0 & 0 & \bar{A}_8 & 0 & 0 & 0 \\ \hline 0 & I_{q21} & 0 & \bar{A}_{171} & 0 & -I_{q21} & 0 \\ 0 & 0 & I_{a-q21} & \bar{A}_{172} & 0 & 0 & 0 \end{array}\right]. \tag{7.19}$$

where the dimensions of (E, A, C) *are conformable to the partitions of* x, *whilst the column partitions of* M *are conformable to the partitions of* f. ♯

Proof The proof consists of two portions: the first part aims to re-express certain system matrices in a form facilitating further analysis. In the coordinates of (7.15), assign the following structures:

$$T_{h1}\begin{bmatrix} A_{43} & M_{23}\end{bmatrix} = \begin{bmatrix} A_{431} & M_{231} \\ A_{432} & M_{232} \\ A_{433} & M_{233}\end{bmatrix}. \tag{7.20}$$

Furthermore, since $\tilde{A}_m$ is full-column rank, there exists a non-singular matrix T_{i1} such that

$$T_{i1}\begin{bmatrix} \tilde{A}_m & A_{433} & M_{233}\end{bmatrix} = \begin{array}{c} \begin{bmatrix} I_{q21} & A_{4331} & M_{2331} \\ 0 & A_{4332} & M_{2332}\end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; q_{21} \\ \updownarrow \; p-q-(n-r-a-q_{22}) \end{array}. \tag{7.21}$$

In the following (and remaining) portion of the proof, the transformations to re-express the system are defined and applied. Define the non-singular matrices

$$\begin{gathered} T_e = \left[\begin{array}{c|cc} I_r & 0 & 0 \\ \hline 0 & T_{e2} & 0 \\ 0 & 0 & T_{e3}\end{array}\right]\begin{bmatrix} I_r & 0 \\ 0 & T_{e1}\end{bmatrix},\ T_f = \begin{bmatrix} (T_{f1}T_{f2})^{-1} & 0 \\ 0 & I_{\bar{n}}\end{bmatrix}, \\ T_h = \begin{bmatrix} I_{\bar{n}-p} & 0 & 0 \\ 0 & T_{h1} & 0 \\ 0 & 0 & I_{n-r}\end{bmatrix},\ T_i = \begin{bmatrix} I_{n-p+q1-a} & 0 & 0 \\ 0 & T_{i1} & 0 \\ 0 & 0 & I_{n-r}\end{bmatrix} T_{i2},\ T_j = \begin{bmatrix} T_{j1} & T_{j2} \\ 0 & T_{j4}\end{bmatrix}, \end{gathered} \tag{7.22}$$

where the partitions $T_{i1},\ T_{j1},\ T_{j2}$, and T_{j4} are defined as

$$T_{i2} = \begin{bmatrix} I_{n-p+q1-a} & 0 & 0 & 0 \\ 0 & I_{r+p+a-q1-n} & M_{223} & 0 \\ 0 & 0 & I_{q21} & 0 \\ 0 & 0 & 0 & I_{n-r-q21} \end{bmatrix},\ T_{j4} = \begin{bmatrix} 0 & 0 & 0 & I_{n-r-a-q22} \\ I_{q21} & 0 & 0 & 0 \\ 0 & I_{a-q21} & 0 & 0 \end{bmatrix},$$

$$T_{j1} = \begin{bmatrix} I_{\bar{n}-p} & -M_{11} & -A_{11} & -X_1 & 0 \\ 0 & I_{q-q2} & 0 & -X_2 & 0 \\ 0 & 0 & 0 & I_{q21} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n-\bar{n}-a} & -X_3 & 0 \\ 0 & 0 & 0 & 0 & I_{p-q-(n-r-a-q22)} \end{bmatrix}, \tag{7.23}$$

$$T_{j2} = \begin{bmatrix} T_{j21} & T_{j22} & T_{j23} & 0 \\ M_{221} & X_2A_{4331} - A_{431} & X_2M_{2331} - M_{231} & 0 \\ 0 & -A_{4331} & -M_{2331} & 0 \\ 0 & 0 & I_{q22} & 0 \\ M_{222} & X_3A_{4332} - A_{432} & X_3M_{2332} - M_{232} & 0 \\ 0 & -A_{4332} & -M_{2332} & 0 \end{bmatrix},$$

where the elements $X_1,\ X_2,\ X_3,\ T_{j21},\ T_{j22}$, and T_{j23} are

$$X_1 = -\left(M_{12} + A_{12} - M_{11}X_2 - A_{11}X_3\right),\ X_2 = A_{421} + M_{221},\ X_3 = A_{422} + M_{222},$$
$$T_{j21} = M_{12} - M_{11}M_{221} - A_{11}M_{222},\ T_{j22} = -A_{13} + M_{11}A_{431} + A_{11}A_{432} - X_1A_{4331},$$
$$T_{j23} = -M_{13} + M_{11}M_{231} + A_{11}M_{232} - X_1M_{2331}. \tag{7.24}$$

Therefore, apply the state equation transformation $T_{pre} = T_jT_iT_hT_eT_dT_b$, the state transformation $T_{aft} = T_aT_cT_f$, and the fault transformation T_g such that

$$\begin{gathered} E_0 \mapsto T_{pre}E_0T_{aft}^{-1},\ A_0 \mapsto T_{pre}A_0T_{aft}^{-1},\ B_0 \mapsto T_{pre}B_0,\ M_0 \mapsto T_{pre}M_0T_e^{-1}, \\ Q_0 \mapsto T_{pre}Q_0,\ C_0 \mapsto C_0T_{aft}^{-1},\ x_0 \mapsto x = T_{aft}x_0,\ f_0 \mapsto f = T_gf_0. \end{gathered} \tag{7.25}$$

It can be seen from the structures in (7.5)–(7.7), (7.13), (7.15), and (7.21) that the system in the coordinates of (7.25) is identical to the structures given in (7.17)–(7.19), thus completing the proof. ■

Note that (7.14) and (7.16) are necessary conditions for the structures in (7.15) and (7.21) to be obtained. Next, define an invertible matrix T, where

$$T = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_p \\ 0 & \bar{T}_p \end{bmatrix}, \; T_p = \begin{bmatrix} I_{q+n-\bar{n}-a} & 0 & 0 & 0 \\ 0 & T_{51} & T_{52} & 0 \end{bmatrix}. \tag{7.26}$$

where $T_p \in \mathbb{R}^{p\times(n-\bar{n}+p)}$ is full-row rank and $\bar{T}_p \in \mathbb{R}^{(n-\bar{n})\times(n-\bar{n}+p)}$. The matrix T_p represents design freedom to be exploited, while $\bar{T}_p$ is solely to make T full rank. One suitable choice of $\bar{T}_p$ would be the transpose of the right null-space of T_p, i.e. $T_p\bar{T}_p^T = 0$. Next, let T_{51} and T_{52} have the following dimensions: $T_{51} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(p-q-(n-r-a-q_{22}))}$, $T_{52} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(n-r-a-q_{22})}$. Apply the state equation transformation T onto the system in the coordinates of (7.17)–(7.19) to obtain

$$\begin{bmatrix} 0 & I_{\bar{n}-p} & \bar{E}_1 \\ 0 & 0 & \tilde{E}_2 \\ 0 & 0 & \bar{E}_3 \\ 0 & 0 & \bar{E}_4 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{T}_p\check{E} \end{bmatrix} \dot{x} = \begin{bmatrix} 0 & \bar{A}_2 & \bar{A}_3 \\ \tilde{A}_4 & \tilde{A}_5 & \tilde{A}_6 \\ \bar{A}_7 & \bar{A}_8 & \bar{A}_9 \\ 0 & \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{16} & \bar{A}_{17} & \bar{A}_{18} \\ \bar{T}_p\check{A}_1 & \bar{T}_p\check{A}_2 & \bar{T}_p\check{A}_3 \end{bmatrix} x + \begin{bmatrix} \bar{B}_1 \\ \tilde{B}_2 \\ \bar{B}_3 \\ \bar{B}_4 \\ \bar{B}_6 \\ \bar{T}_p\check{B} \end{bmatrix} u + \begin{bmatrix} 0 \\ \tilde{M}_2 \\ 0 \\ 0 \\ \bar{M}_6 \\ \bar{T}_p\check{M} \end{bmatrix} f + \begin{bmatrix} \bar{Q}_1 \\ \tilde{Q}_2 \\ \bar{Q}_3 \\ \bar{Q}_4 \\ \bar{Q}_6 \\ \bar{T}_p\check{Q} \end{bmatrix} \xi, \tag{7.27}$$

$$y = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} x. \tag{7.28}$$

By substituting for the structures of $\tilde{A}_4$, $\bar{A}_7$, $\bar{A}_{16}$, $\tilde{M}_2$, and $\bar{M}_6$ from (7.19) system (7.27)–(7.28) can be expanded to become:

$$
\left[\begin{array}{ccc|c|c}
0 & 0 & 0 & I_{\bar{n}-p} & \bar{E}_1 \\ \hline
0 & 0 & 0 & 0 & \tilde{E}_{21} \\
0 & 0 & 0 & 0 & \tilde{E}_{22} \\
0 & 0 & 0 & 0 & \tilde{E}_{23} \\ \hline
0 & 0 & 0 & 0 & \bar{E}_3 \\
0 & 0 & 0 & 0 & \bar{E}_4 \\ \hline
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & \bar{T}_p\check{E}
\end{array}\right]
\left[\begin{array}{c}
\dot{x}_{11} \\ \dot{x}_{121} \\ \dot{x}_{122} \\ \hline \dot{x}_2 \\ \hline \dot{y}
\end{array}\right]
=
\left[\begin{array}{ccc|c|c}
0 & 0 & 0 & \bar{A}_2 & \bar{A}_3 \\ \hline
0 & 0 & 0 & \tilde{A}_{51} & \tilde{A}_{61} \\
0 & I_{q21} & 0 & \tilde{A}_{52} & \tilde{A}_{62} \\
0 & 0 & 0 & \tilde{A}_{53} & \tilde{A}_{63} \\ \hline
I_{n-\bar{n}-a} & 0 & 0 & \bar{A}_8 & \bar{A}_9 \\
0 & 0 & 0 & \bar{A}_{11} & \bar{A}_{12} \\ \hline
0 & I_{q21} & 0 & \bar{A}_{171} & \bar{A}_{181} \\
0 & 0 & I_{a-q21} & \bar{A}_{172} & \bar{A}_{182} \\ \hline
\bar{T}_p\check{A}_{11} & \bar{T}_p\check{A}_{12} & \bar{T}_p\check{A}_{13} & \bar{T}_p\check{A}_2 & \bar{T}_p\check{A}_3
\end{array}\right]
\left[\begin{array}{c}
x_{11} \\ x_{121} \\ x_{122} \\ \hline x_2 \\ \hline y
\end{array}\right]
$$

$$
+\left[\begin{array}{c}
\bar{B}_1 \\ \hline \tilde{B}_{21} \\ \tilde{B}_{22} \\ \tilde{B}_{23} \\ \hline \bar{B}_3 \\ \bar{B}_4 \\ \hline \bar{B}_{61} \\ \bar{B}_{62} \\ \hline \bar{T}_p\check{B}
\end{array}\right] u +
\left[\begin{array}{ccc}
0 & 0 & 0 \\ \hline
I_{q1} & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & I_{q22} \\ \hline
0 & 0 & 0 \\
0 & 0 & 0 \\ \hline
0 & -I_{q21} & 0 \\
0 & 0 & 0 \\ \hline
\bar{T}_p\check{M}_1 & \bar{T}_p\check{M}_2 & \bar{T}_p\check{M}_3
\end{array}\right]
\left[\begin{array}{c} f_1 \\ f_{21} \\ f_{22} \end{array}\right] +
\left[\begin{array}{c}
\bar{Q}_1 \\ \hline \tilde{Q}_{21} \\ \tilde{Q}_{22} \\ \tilde{Q}_{23} \\ \hline \bar{Q}_3 \\ \bar{Q}_4 \\ \hline \bar{Q}_{61} \\ \bar{Q}_{62} \\ \hline \bar{T}_p\check{Q}
\end{array}\right] \xi, \tag{7.29}
$$

$$
y = \left[\begin{array}{ccc|c|c} 0 & 0 & 0 & 0 & I_p \end{array}\right]
\left[\begin{array}{c}
x_{11} \\ x_{121} \\ x_{122} \\ \hline x_2 \\ \hline y
\end{array}\right]. \tag{7.30}
$$

It can be seen from (7.29)–(7.30) that $\begin{bmatrix} x_{121} \\ x_{122} \end{bmatrix}$ can be re-expressed as a linear combination of $x_2,\ y,\ u,\ f$, and ξ as follows:

$$
\begin{bmatrix} x_{121} \\ x_{122} \end{bmatrix} = -\begin{bmatrix} \bar{A}_{171} & \bar{A}_{181} \\ \bar{A}_{172} & \bar{A}_{182} \end{bmatrix}\begin{bmatrix} x_2 \\ y \end{bmatrix} - \begin{bmatrix} \bar{B}_{61} \\ \bar{B}_{62} \end{bmatrix} u + \begin{bmatrix} 0 & I_{q21} & 0 \\ 0 & 0 & 0 \end{bmatrix} f - \begin{bmatrix} \bar{Q}_{61} \\ \bar{Q}_{62} \end{bmatrix} \xi. \tag{7.31}
$$

Using the re-expression (7.31), system (7.29)–(7.30) can be re-expressed as follows:

$$\left[\begin{array}{c|c|c} 0 & I_{\bar{n}-p} & \bar{E}_1 \\ \hline 0 & 0 & \bar{E}_{21} \\ 0 & 0 & \bar{E}_{22} \\ 0 & 0 & \bar{E}_{23} \\ \hline 0 & 0 & \bar{E}_3 \\ 0 & 0 & \bar{E}_4 \\ \hline 0 & 0 & \bar{T}_p\check{E} \end{array}\right] \begin{bmatrix} \dot{x}_{11} \\ \dot{x}_2 \\ \dot{y} \end{bmatrix} = \left[\begin{array}{c|c|c} 0 & \bar{A}_2 & \bar{A}_3 \\ \hline 0 & \bar{A}_{51} & \bar{A}_{61} \\ 0 & \bar{A}_{52} & \bar{A}_{62} \\ 0 & \bar{A}_{53} & \bar{A}_{63} \\ \hline I_{n-\bar{n}-a} & \bar{A}_8 & \bar{A}_9 \\ 0 & \bar{A}_{11} & \bar{A}_{12} \\ \hline \bar{T}_p\check{A}_{11} & \bar{T}_p\check{A}_{12} & \bar{T}_p\check{A}_{13} \end{array}\right] \begin{bmatrix} x_{11} \\ x_2 \\ y \end{bmatrix} + \left[\begin{array}{c} \bar{B}_1 \\ \hline \bar{B}_{21} \\ \bar{B}_{22} \\ \bar{B}_{23} \\ \hline \bar{B}_3 \\ \bar{B}_4 \\ \hline \bar{T}_p\check{B} \end{array}\right] u$$

$$+ \left[\begin{array}{ccc} 0 & 0 & 0 \\ \hline I_{q1} & 0 & 0 \\ 0 & I_{q21} & 0 \\ 0 & 0 & I_{q22} \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline \bar{T}_p\check{M}_1 & \bar{T}_p\check{M}_2 & \bar{T}_p\check{M}_3 \end{array}\right] \begin{bmatrix} f_1 \\ f_{21} \\ f_{22} \end{bmatrix} + \left[\begin{array}{c} \bar{Q}_1 \\ \hline \bar{Q}_{21} \\ \bar{Q}_{22} \\ \bar{Q}_{23} \\ \hline \bar{Q}_3 \\ \bar{Q}_4 \\ \hline \bar{T}_p\check{Q} \end{array}\right] \xi, \tag{7.32}$$

$$y = \begin{bmatrix} 0 & 0 & I_p \end{bmatrix} \begin{bmatrix} x_{11} \\ x_2 \\ y \end{bmatrix}. \tag{7.33}$$

Then treat x_{11} as an unknown input, and define

$$\bar{x} = \begin{bmatrix} x_2 \\ y \end{bmatrix}, \ \bar{f} = \begin{bmatrix} f \\ x_1 \end{bmatrix}. \tag{7.34}$$

System (7.32)–(7.33) can then be re-expressed as the following reduced-order system:

$$\left[\begin{array}{c|c} I_{\bar{n}-p} & \bar{E}_1 \\ \hline 0 & \bar{E}_{21} \\ 0 & \bar{E}_{22} \\ 0 & \bar{E}_{23} \\ 0 & \bar{E}_3 \\ 0 & \bar{E}_4 \end{array}\right] \dot{\bar{x}} = \left[\begin{array}{c|c} \bar{A}_2 & \bar{A}_3 \\ \hline \bar{A}_{51} & \bar{A}_{61} \\ \bar{A}_{52} & \bar{A}_{62} \\ \bar{A}_{53} & \bar{A}_{63} \\ \bar{A}_8 & \bar{A}_9 \\ \bar{A}_{11} & \bar{A}_{12} \end{array}\right] \bar{x} + \left[\begin{array}{c} \bar{B}_1 \\ \hline \bar{B}_{21} \\ \bar{B}_{22} \\ \bar{B}_{23} \\ \bar{B}_3 \\ \bar{B}_4 \end{array}\right] u + \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ \hline I_{q1} & 0 & 0 & 0 \\ 0 & I_{q21} & 0 & 0 \\ 0 & 0 & I_{q22} & 0 \\ 0 & 0 & 0 & I_{n-\bar{n}-a} \\ 0 & 0 & 0 & 0 \end{array}\right] \bar{f} + \left[\begin{array}{c} \bar{Q}_1 \\ \hline \bar{Q}_{21} \\ \bar{Q}_{22} \\ \bar{Q}_{23} \\ \bar{Q}_3 \\ \bar{Q}_4 \end{array}\right] \xi, \tag{7.35}$$

$$y = \begin{bmatrix} 0 & I_p \end{bmatrix} \bar{x}. \tag{7.36}$$

For ease of analysis, denote the partitions for (7.35)–(7.36) as follows:

$$\underbrace{\begin{bmatrix} I_{\bar{n}-p} & \bar{E}_1 \\ 0 & \bar{E}_2 \end{bmatrix}}_{\bar{E}} \dot{\bar{x}} = \underbrace{\begin{bmatrix} \bar{A}_2 & \bar{A}_3 \\ \bar{A}_5 & \bar{A}_6 \end{bmatrix}}_{\bar{A}} \bar{x} + \underbrace{\begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix}}_{\bar{B}} u + \underbrace{\begin{bmatrix} 0 & 0 \\ \bar{M}_2 & \bar{A}_4 \end{bmatrix}}_{\bar{M}} \bar{f} + \underbrace{\begin{bmatrix} \bar{Q}_1 \\ \bar{Q}_2 \end{bmatrix}}_{\bar{Q}} \xi, \tag{7.37}$$

$$y = \underbrace{\begin{bmatrix} 0 & I_p \end{bmatrix}}_{\bar{C}} \bar{x}, \; \begin{bmatrix} \bar{A}_{a5} & \bar{Q}_{a2} \end{bmatrix} = \left[\begin{array}{c|c} \bar{A}_{51} & \bar{Q}_{21} \\ \bar{A}_{52} & \bar{Q}_{22} \\ \bar{A}_{53} & \bar{Q}_{23} \end{array}\right]. \tag{7.38}$$

Notice that $\operatorname{rank} \begin{bmatrix} \bar{E} \\ \bar{C} \end{bmatrix} = \bar{n}$ (which is full-column rank) and the output of the reduced-order system (7.37)–(7.38) is a measurable signal y. Thus the observer by Yeu et al. [157] can be designed based on $\left(\bar{E}, \bar{A}, \bar{B}, \bar{M}, \bar{Q}, \bar{C}\right)$ and driven by u and y, and applied onto system (7.37)–(7.38) to reconstruct $\bar{f}$, and therefore reconstruct f_0.

7.3 Observer formulation

The observer by Yeu et al. for system (7.37)–(7.38) has the following structure:

$$\dot{z} = \left(R\bar{A} - G_l\bar{C}\right) z - R\bar{B}u - \left(G\left(I_p - \bar{C}V\right) + R\bar{A}V\right) y - G_n\nu, \tag{7.39}$$

$$\hat{\bar{x}} = Vy - z, \tag{7.40}$$

$$\nu = -\rho \frac{e_y}{\|e_y\|}, \tag{7.41}$$

$$e_y = \bar{C}\hat{\bar{x}} - y, \; G_l = \begin{bmatrix} G_{l1} \\ G_{l2} \end{bmatrix}, \; G_n = \begin{bmatrix} 0 \\ G_{n2} \end{bmatrix}, \tag{7.42}$$

where $G_{l1} \in \mathbb{R}^{(\bar{n}-p)\times p}$, $|G_{n2}| \neq 0$, $R \in \mathbb{R}^{\bar{n}\times\bar{n}}$ and $V \in \mathbb{R}^{\bar{n}\times p}$. Pre-multiply (7.37) with R and add $V\dot{y}$ to both sides to obtain

$$R\bar{E}\dot{\bar{x}} + V\dot{y} = \left(R\bar{E} + V\bar{C}\right)\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\dot{y}. \tag{7.43}$$

Next, suppose $R\bar{E} + V\bar{C} = I_{\bar{n}}$. Equation (7.43) becomes

$$\dot{\bar{x}} = R\bar{A}\bar{x} + R\bar{B}u + R\bar{M}\bar{f} + R\bar{Q}\xi + V\bar{C}\dot{\bar{x}}. \tag{7.44}$$

Corollary 7.1 *The matrices R, V from* (7.39)–(7.40) *will have the following structures:*

$$R = \begin{bmatrix} I_{\bar{n}-p} & R_2 \\ 0 & R_4 \end{bmatrix}, \; V = \begin{bmatrix} -\left(\bar{E}_1 + R_2\bar{E}_2\right) \\ I_p - R_4\bar{E}_2 \end{bmatrix} \tag{7.45}$$

where $R_4 \in \mathbb{R}^{p\times p}$, $|R_4| \neq 0$. ♯

Proof Since $\mathrm{rank}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = \bar{n}$, then the matrices R and V can be chosen such that

$$\begin{bmatrix}R & V\end{bmatrix}\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix} = I_{\bar{n}}, \tag{7.46}$$

that is, $\begin{bmatrix}R & V\end{bmatrix}$ is chosen to be the Moore-Penrose inverse of $\begin{bmatrix}\bar{E}\\ \bar{C}\end{bmatrix}$. Partition the matrices R and V generally as follows:

$$\begin{bmatrix}R & V\end{bmatrix} = \left[\begin{array}{cc|c} R_1 & R_2 & V_1\\ R_3 & R_4 & V_2\end{array}\right], \tag{7.47}$$

where $R_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$ and $R_4 \in \mathbb{R}^{p\times p}$. By substituting $\bar{E}$ and $\bar{C}$ from (7.37)–(7.38) and R and V from (7.47) into (7.46), it can be seen that R and V would take the forms given in (7.45). ■

Substitute z from (7.40) into (7.39) to obtain the following analytical structure for the observer:

$$\begin{aligned}\left(V\dot{y} - \dot{\hat{\bar{x}}}\right) &= \left(R\bar{A} - G_l\bar{C}\right)\left(Vy - \hat{\bar{x}}\right) - R\bar{B}u - \left(G_l\left(I_p - \bar{C}V\right) + R\bar{A}V\right)y - G_n\nu\\ \dot{\hat{\bar{x}}} &= \left(R\bar{A} - G_l\bar{C}\right)\hat{\bar{x}} + R\bar{B}u + G_l\bar{C}\bar{x} + G_n\nu + V\bar{C}\dot{\bar{x}}.\end{aligned} \tag{7.48}$$

Define the state estimation error for the observer as

$$e = \hat{\bar{x}} - \bar{x} = \begin{bmatrix}e_2\\ e_y\end{bmatrix}\begin{array}{l}\updownarrow\ \ \bar{n}-p\\ \updownarrow\ \ p\end{array}. \tag{7.49}$$

Hence by subtracting (7.44) from (7.48), the error equation for the observer (which characterises its performance) is given by

$$\dot{e} = \left(R\bar{A} - G_l\bar{C}\right)e - R\bar{M}f - R\bar{Q}\xi + G_n\nu. \tag{7.50}$$

Let R_2 have the structure

$$R_2 = \begin{bmatrix}0 & R_{23}\end{bmatrix}, \tag{7.51}$$

where $R_{23} \in \mathbb{R}^{(\bar{n}-p)\times(p-q-(n-\bar{n}-a))}$. By substituting the structures of $\bar{A}, \bar{M}, \bar{Q}, \bar{C}$ from (7.37)–(7.38), G_l and G_n from (7.42), and R from (7.45) and (7.51) into (7.50), the following error system is obtained:

$$\begin{aligned}\begin{bmatrix}\dot{e}_2\\ \dot{e}_y\end{bmatrix} &= \begin{bmatrix}\bar{A}_2 + R_{23}\bar{A}_{11} & \bar{A}_3 + R_{23}\bar{A}_{12} - G_{l1}\\ R_4\bar{A}_5 & R_4\bar{A}_6 - G_{l2}\end{bmatrix}\begin{bmatrix}e_2\\ e_y\end{bmatrix} + \begin{bmatrix}0\\ G_{n2}\end{bmatrix}\nu\\ &\quad - \begin{bmatrix}0 & 0\\ R_4\bar{M}_2 & R_4\bar{A}_4\end{bmatrix}\bar{f} - \begin{bmatrix}\bar{Q}_1 + R_{23}\bar{Q}_4\\ R_4\bar{Q}_2\end{bmatrix}\xi.\end{aligned} \tag{7.52}$$

7.3.1 Convergence of the observer

Proposition 7.2 *Suppose there exists a matrix $P = P^T > 0$ that satisfies*

$$P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0, \tag{7.53}$$

where $P = \begin{bmatrix}P_1 & 0\\ 0 & G_{n2}^{-1}\end{bmatrix}$, $P_1 \in \mathbb{R}^{(\bar{n}-p)\times(\bar{n}-p)}$. *If ρ in* (7.41) *is chosen as follows:*

$$\begin{aligned}\rho &> \|G_{n2}^{-1}R_4\bar{A}_5\|\frac{2\mu_1\beta}{\mu_0} + \|G_{n2}^{-1}R_4\begin{bmatrix}\bar{M}_2 & \bar{A}_4\end{bmatrix}\|\alpha + \|G_{n2}^{-1}R_4\bar{Q}_2\|\beta + \eta,\\ \mu_0 &= -\lambda_{\max}\left(P\mathbb{A} + \mathbb{A}^T P\right) > 0,\ \mu_1 = \|PR\bar{Q}\|,\ \eta \in \mathbb{R}^+,\\ \alpha &\geq \|f\|_{\max} + \|x_{11}\|_{\max},\ \beta \geq \|\xi_{\max}\|,\end{aligned} \tag{7.54}$$

then sliding motion $(e_y, \dot{e}_y = 0)$ takes place in finite time. ♯

Proof The error system (7.52) has the same form as that of the Edwards-Spurgeon SMO for a system represented by the quadruple $(R\bar{A}, R\bar{M}, R\bar{Q}, \bar{C})$, and so the proof of convergence is adapted from [128] as follows: define a candidate Lyapunov function $W = e^T P e > 0$, and differentiate W with respect to time to obtain

$$\begin{aligned}\dot{W} &= e^T\left(P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P\right)e - 2e^T PR\bar{M}\bar{f} - 2e^T PR\bar{Q}\xi + 2e_y^T\nu\\ &\leq -\mu_0\|e\|^2 + 2\|e\|\mu_1\beta - 2\|e_y\|\left(\rho - \|G_{n2}^{-1}\bar{C}R\bar{M}\|\alpha\right).\end{aligned} \tag{7.55}$$

By setting $\rho > \|G_{n2}^{-1}\bar{C}R\bar{M}\|\alpha$, (7.55) becomes

$$\dot{W} \leq \|e\|\left(-\mu_0\|e\| + 2\mu_1\beta\right), \tag{7.56}$$

By using the same arguments in Proposition 4.2, it can be shown that the magnitude of e would be bounded $\left(\|e\| \le \frac{2\mu_1\beta}{\mu_0}\right)$ in finite time. The next (and remaining) portion of the proof aims to show how sliding motion $(e_y, \dot{e}_y = 0)$ is achieved.

Define another candidate Lyapunov function $W_y = e_y^T G_{n2}^{-1} e_y > 0$. Differentiating it with respect to time yields

$$\begin{aligned}
\dot{W}_y &= e_y^T \left(G_{n2}^{-1}\left(R_4\bar{A}_6 - G_{l2}\right) + \left(R_4\bar{A}_6 - G_{l2}\right)^T G_{n2}^{-1}\right) e_y \\
&\quad + 2e_y^T \left(G_{n2}^{-1}R_4\bar{A}_5 e_2 - G_{n2}^{-1}R_4\begin{bmatrix}\bar{M}_2 & \bar{A}_4\end{bmatrix}\bar{f} - G_{n2}^{-1}R_4\bar{Q}_2\xi\right) \\
&\le -2\|e_y\|\left(\rho - \|G_{n2}^{-1}R_4\bar{A}_5\|\frac{2\mu_1\beta}{\mu_0} - \|G_{n2}^{-1}R_2\begin{bmatrix}\bar{M}_2 & \bar{A}_4\end{bmatrix}\|\alpha - \|G_{n2}^{-1}R_4\bar{Q}_2\|\beta\right).
\end{aligned} \tag{7.57}$$

Note that

$$\|e_y\|^2 = \left(\sqrt{G_{n2}^{-1}}e_y\right)^T G_{n2}\left(\sqrt{G_{n2}^{-1}}e_y\right) \ge \lambda_{\min}\left(G_{n2}\right)\|\sqrt{G_{n2}^{-1}}e_y\|^2 = \lambda_{\min}\left(G_{n2}\right)W_y. \tag{7.58}$$

Hence by setting ρ such that (7.54) is satisfied (which also satisfies $\rho > \|G_{n2}^{-1}\bar{C}R\bar{M}\|\alpha$), (7.58) becomes

$$\dot{W}_y \le -2\eta\|e_y\| \le \sqrt{\lambda_{\min}\left(G_{n2}\right)}\sqrt{W_y}, \tag{7.59}$$

which is the *reachability condition* [128] resulting in $e_y = 0$ in finite time, and a sliding motion taking place on the surface $\mathcal{S} = \{e : \bar{C}e = 0\}$, thus proving the proposition. ■

After sliding motion $(e_y, \dot{e}_y = 0)$ occurs, (7.52) becomes

$$\dot{e}_2 = \left(\bar{A}_2 + R_{23}\bar{A}_{11}\right)e_2 - \left(\bar{Q}_1 + R_{23}\bar{Q}_4\right)\xi, \tag{7.60}$$

$$0 = R_4\left(\bar{A}_5 e_2 - \begin{bmatrix}\bar{M}_2 & \bar{A}_4\end{bmatrix}\bar{f} - R_4\bar{Q}_2\xi\right) + G_{n2}\nu. \tag{7.61}$$

7.3.2 Robustly reconstructing the fault

Define the measurable signal $\bar{\nu} = R_4 G_{n2}^{-1}\nu$. By substituting the structures of $\bar{A}_5$, $\bar{M}_2$, $\bar{A}_4$, and $\bar{Q}_2$ from (7.35) and (7.38), the following expression is obtained:

$$\bar{\nu} = -\begin{bmatrix}\bar{A}_{a5} \\ \bar{A}_8 \\ \bar{A}_{11}\end{bmatrix}e_2 + \begin{bmatrix}I_q & 0 \\ 0 & I_{n-\bar{n}-a} \\ 0 & 0\end{bmatrix}\begin{bmatrix}f \\ x_{11}\end{bmatrix} + \begin{bmatrix}\bar{Q}_{a2} \\ \bar{Q}_3 \\ \bar{Q}_4\end{bmatrix}\xi. \tag{7.62}$$

Define the fault reconstruction signal as

$$\hat{f} = L\bar{\nu}, \tag{7.63}$$

where $L = \begin{bmatrix} I_q & 0 & L_0 \end{bmatrix}$, where $L_0 \in \mathbb{R}^{q\times(p-q-n+\bar{n}+a)}$ represents design freedom. Premultiply (7.62) with L to obtain

$$\hat{f} = -(\bar{A}_{a5} + L_0\bar{A}_{11})e_2 + (\bar{Q}_{a2} + L_0\bar{Q}_4)\xi + f, \tag{7.64}$$

which estimates f completely. Note that x_{11} does not appear and thus has no effect on $\hat{f}$, but e_2 and ξ affect $\hat{f}$; this issue will be addressed later in §7.3.2 when the observer is designed to minimise their effects on $\hat{f}$.

Define the fault reconstruction error

$$e_f = \hat{f} - f. \tag{7.65}$$

Rearranging (7.60)–(7.61) using (7.64)–(7.65) yields

$$\dot{e}_2 = \underbrace{\left(\bar{A}_2 + R_{23}\bar{A}_{11}\right)}_{\breve{A}} e_2 + \underbrace{\left(-\bar{Q}_1 - R_{23}\bar{Q}_4\right)}_{\breve{B}} \xi, \tag{7.66}$$

$$e_f = \underbrace{\left(-\bar{A}_{a5} - L_0\bar{A}_{11}\right)}_{\breve{C}} e_2 + \underbrace{\left(\bar{Q}_{a2} + L_0\bar{Q}_4\right)}_{\breve{D}} \xi, \tag{7.67}$$

which is a state-space system that shows how the disturbance ξ affects the fault reconstruction error e_f. The goal now is to design the observer such that the $\mathcal{L}_2$ gain from ξ to e_f is minimised, by choice of L_0 and R_{23}.

Remark 7.1 *If* $\xi = 0$ *(i.e. system* (7.1)–(7.2) *is not affected by external disturbances),* (7.66)–(7.67) *becomes*

$$\dot{e}_2 = \breve{A}e_2, \tag{7.68}$$

$$e_f = \breve{C}e_2, \tag{7.69}$$

which implies $e_f \to 0$, which in turn implies $\hat{f} \to f$. Thus, in the absence of disturbances, the proposed scheme is able to asymptotically reconstruct the faults. ♯

7.3.3 Observer design for robust fault reconstruction

The effect of ξ on e_f is minimised using the Bounded Real Lemma, which has been described in Lemma 4.2. By applying the Bounded Real Lemma onto error system (7.66)–(7.67), the $\mathcal{L}_2$ gain from ξ to e_f would not exceed the positive scalar γ if there exists a matrix $P_1 = P_1^T > 0$ such that

$$\begin{bmatrix} P_1\breve{A} + \breve{A}^T P_1 & P_1\breve{B} & \breve{C}^T \\ \breve{B}^T P_1 & -\gamma I_h & \breve{D}^T \\ \breve{C} & \breve{D} & -\gamma I_q \end{bmatrix} < 0. \tag{7.70}$$

Thus, the objective is to find the solution for $\gamma,\ P_1,\ L_0$, and R_{23} that minimises γ subject to inequality (7.70), while also satisfying

$$P\left(R\bar{A} - G_l\bar{C}\right) + \left(R\bar{A} - G_l\bar{C}\right)^T P < 0, \tag{7.71}$$

which is required to guarantee sliding motion in the observer (refer to Proposition 7.2). The choice of G_l is not unique; the only requirement is that (7.71) is satisfied. In this paper, G_l is designed using a modified version of the method in [128], which is described in Proposition 7.3 below.

Proposition 7.3 *Define $\bar{D} \in \mathbb{R}^{p\times p}$ and $\gamma_0 \in \mathbb{R}^+$ as user-defined variables. Then define*

$$J = P_1 R_{23},\ \bar{F} = \begin{bmatrix} \breve{C} & F_1 \end{bmatrix}, \tag{7.72}$$

where $F_1 \in \mathbb{R}^{q\times p}$. Suppose there exists matrices $P,\ R_{23},\ L_0,\ F_1,\ J$, and γ that satisfy the following inequalities:

$$\begin{bmatrix} P(R\bar{A}) + (R\bar{A})^T P - \gamma_0 \bar{C}^T(\bar{D}\bar{D}^T)^{-1}\bar{C} & -P(R\bar{Q}) & \bar{F}^T \\ -(R\bar{Q})^T P & -\gamma_0 I_h & \breve{D}^T \\ \bar{F} & \breve{D} & -\gamma_0 I_q \end{bmatrix} < 0, \tag{7.73}$$

$$\begin{bmatrix} P_1\bar{A}_2 + J\bar{A}_{11} + (P_1\bar{A}_2 + J\bar{A}_{11})^T & -P_1\breve{B} & \breve{C}^T \\ -\breve{B}^T P_1 & -\gamma I_h & \breve{D}^T \\ \breve{C} & \breve{D} & -\gamma I_q \end{bmatrix} < 0. \tag{7.74}$$

If the observer parameters are chosen as

$$R_{23} = P_1^{-1} J,\ G_l = \frac{\gamma_0}{2} P^{-1}\bar{C}^T(\bar{D}\bar{D}^T)^{-1}, \tag{7.75}$$

then (7.71) *is satisfied, and $\|e_f\| \leq \gamma\|\xi\|$.* ♯

Proof From the structure of G_l in (7.75), the following can be inferred:

$$PG_l\bar{C} + (G_l\bar{C})^T P = \gamma_0 \bar{C}^T (\bar{D}\bar{D}^T)^{-1} \bar{C}. \tag{7.76}$$

that is, the top-left element of LMI (7.73) can be re-expressed as

$$P(R\bar{A} - G_l\bar{C}) + (R\bar{A} - G_l\bar{C})^T P. \tag{7.77}$$

The structure of LMI (7.73) and equation (7.77) imply (7.71) is satisfied. Furthermore, by substituting R_{23} from (7.75) into LMI (7.74), it can be seen that (7.74) has the same form as (7.70), which describes the Bounded Real Lemma bounding the $\mathcal{L}_2$ gain from ξ to e_f. Thus, the proof is complete. ■

The solution γ to the LMI pair (7.73)–(7.74) would be the upper bound of the $\mathcal{L}_2$ gain from ξ to e_f. As long as there exists an upper bound on $\|\xi\|$, then there would also exist an upper bound on e_f, given by $\|e_f\| \leq \gamma\|\xi\|$.

7.3.4 Existence conditions

Theorem 7.1 *The proposed SMO scheme can reconstruct f_0 for system* (7.1)–(7.2) *if and only if the following conditions hold:*

Q1. $\mathrm{rank}\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} = q + n - \bar{n},$

Q2. $p \geq q + n - \bar{n} - a.$

The following condition is necessary, and also sufficient if $p - q - n + \bar{n} + a > 0$*:*

Q3. $\mathrm{rank}\begin{bmatrix} sE_0 - A_0 & M_0 \\ C_0 & 0 \end{bmatrix} = n + q \;\forall\; s \in \mathbb{C}^+.$

If $p - q - n + \bar{n} + a = 0$*, the sufficient condition is that*

Q4. The eigenvalues of $\bar{A}_2$ (as defined in (7.17)*) are stable.* ♯

Proof The remainder of this subsection forms the constructive proof for Theorem 7.1.

In the analysis of the observer scheme in the prior sub-sections, several assumptions were made and are listed as follows:

R1. $\begin{bmatrix} A_{41} & A_{51} \end{bmatrix}$ is full-column rank in (7.14) so that T_{h1} in (7.15) exists.

R2. $\tilde{A}_m$ in (7.16) is full-column rank so that T_{i1} in (7.21) exists.

R3. Matrices T_{51} and T_{52} in (7.26) exist such that they have the dimensions $T_{51} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(p-q-(n-r-a-q_{22}))}$ and $T_{52} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(n-r-a-q_{22})}$.

R4. Error system (7.60)–(7.61) (and therefore (7.66)–(7.67)) can be made stable.

Therefore it is of interest to re-express these conditions in terms of the original system matrices so that it is easier for the designer to check if the proposed scheme is applicable from the outset.

Proposition 7.4 *Condition Q1 is necessary and sufficient for the satisfaction of R1–R2.* $\sharp$

Proof Pre-multiply (M_0, A_0) with $T_e T_d T_b$ and post-multiply M_0 and A_0 with $(T_a T_c T_f)^{-1}$ and T_g^{-1}, respectively in Proposition 7.1 to obtain

$$\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} = \left[\begin{array}{ccc|ccc} M_{21} & M_{22} & M_{23} & A_{41} & A_{42} & A_{43} \\ 0 & -I_{q21} & 0 & 0 & I_{q21} & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{a-q21} \\ 0 & 0 & I_{q2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{7.78}$$

Since $\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} \in \mathbb{R}^{(n-\bar{n}+p)\times(n-\bar{n}+q)}$ from (7.5),

$$\text{rank}\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} = \min\{n-\bar{n}+p, n-\bar{n}+q\} \le n-\bar{n}+q. \tag{7.79}$$

Equations (7.78)–(7.79) imply

$$\text{rank}\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} = \text{rank}\underbrace{\begin{bmatrix} M_{21} & M_{22} & A_{41} & A_{42} \\ 0 & -I_{q21} & 0 & I_{q21} \end{bmatrix}}_{M_{ac}} + a - q_{21} + q_{22}. \tag{7.80}$$

The matrix M_{ac} (and therefore (7.80)) could be further simplified as follows:

$$\begin{aligned}\operatorname{rank}(M_{ac}) &= \operatorname{rank}\begin{bmatrix} M_{21} & M_{22} & A_{41} & A_{42} \\ 0 & -I_{q21} & 0 & I_{q21} \end{bmatrix}\begin{bmatrix} I_{q1} & 0 & 0 & 0 \\ 0 & I_{q21} & 0 & 0 \\ 0 & 0 & I_{n-\bar{n}-a} & 0 \\ 0 & I_{q21} & 0 & I_{q21} \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} & A_{42} \\ 0 & 0 & 0 & I_{q21} \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} \end{bmatrix} + q_{21}. \end{aligned} \tag{7.81}$$

Therefore, (7.79)–(7.81) imply

$$\operatorname{rank}\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} \end{bmatrix} \le n - \bar{n} - a + q - q_{22}. \tag{7.82}$$

Conditions R1–R2 imply $\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} \end{bmatrix}$ is full-column rank, i.e.

$$\operatorname{rank}\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} \end{bmatrix} = n - \bar{n} - a + q - q_{22}. \tag{7.83}$$

To show that Q1 is necessary, suppose Q1 is not satisfied, i.e.

$$\operatorname{rank}\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} < q + n - \bar{n}. \tag{7.84}$$

By comparing (7.84) with (7.79), (7.84) would imply

$$\operatorname{rank}\begin{bmatrix} M_{21} & M_{22}+A_{42} & A_{41} \end{bmatrix} < n - \bar{n} - a + q - q_{22}, \tag{7.85}$$

that is, (7.83) is not satisfied. This would in turn imply that R1–R2 cannot both be satisfied (it is possible either R1 or R2 is still satisfied, but never both), and it would be impossible to attain the structures in both (7.15) and (7.21), thus showing the necessity of Q1. On the other hand, if Q1 is satisfied, then it is straightforward to see from (7.15) and (7.21) that T_{h1} and T_{i1} would exist such that

$$T_{h1}\begin{bmatrix} M_{21} & A_{41} \end{bmatrix} = \begin{bmatrix} I_{q1+n-\bar{n}-a} \\ 0 \end{bmatrix}, \; T_{i1}\left(\tilde{A}_m\right) = \begin{bmatrix} I_{q1} \\ 0 \end{bmatrix}, \tag{7.86}$$

i.e. R1–R2 are satisfied, showing the sufficiency of Q1 and therefore proving the proposition. ■

Proposition 7.5 *Condition R3 is satisfied if and only if Q2 is satisfied.* $\sharp$

Proof Matrices T_{51} and T_{52} have been defined in (7.26) such that they have the dimensions $T_{51} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(p-q-(n-r-a-q_{22}))}$ and $T_{52} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(n-r-a-q_{22})}$. To prove that Q2 is necessary, suppose Q2 is not satisfied, i.e.

$$p < q + n - \bar{n} - a. \tag{7.87}$$

It is straightforward to see that T_{51} and T_{52} would not exist if this is the case (and therefore R3 is not satisfied), proving the necessity of Q2. To show that Q2 is sufficient, recall from (7.11) and (7.21) that

$$\begin{aligned} T_{e3}M_{32}T_g^{-1} &= \begin{array}{c} \begin{bmatrix} 0 & I_{q22} \\ 0 & 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; q_{22} \\ \updownarrow \; n-r-a-q_{22} \end{array}, \\ T_{i1}\tilde{A}_m &= \begin{array}{c} \begin{bmatrix} I_{q21} \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow \; q_{21} \\ \updownarrow \; p-q-(n-r-a-q_{22}) \end{array}. \end{aligned} \tag{7.88}$$

Equation (7.88) implies

$$n-r-a-q_{22} \geq 0, \; p-q-(n-r-a-q_{22}) \geq 0. \tag{7.89}$$

Therefore, if Q2 is satisfied (i.e. $p \geq q+n-\bar{n}-a$), then T_{51} and T_{52} would have the dimensions $T_{51} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(p-q-(n-r-a-q_{22}))}$ and $T_{52} \in \mathbb{R}^{(p-q-n+\bar{n}+a)\times(n-r-a-q_{22})}$ (i.e. R3 is satisfied), thus completing the proof. ■

Proposition 7.6 *For R4 to be satisfied, Q3 is necessary. If* $p-q-n+\bar{n}+a > 0$*, then Q3 is also sufficient; otherwise (for* $p-q-n+\bar{n}+a = 0$*), Q4 is sufficient.* $\sharp$

Proof Expand the left-hand side of Q3 using (7.17)–(7.18) to get

$$\underbrace{\operatorname{rank}\begin{bmatrix} sE_0 - A_0 & M_0 \\ C_0 & 0 \end{bmatrix}}_{R(s)} = \operatorname{rank}\left[\begin{array}{cccc|c} 0 & sI_{\bar{n}-p} - \bar{A}_2 & s\bar{E}_1 - \bar{A}_3 & 0 \\ -\tilde{A}_4 & -\tilde{A}_5 & s\tilde{E}_2 - \tilde{A}_6 & \tilde{M}_2 \\ -\bar{A}_7 & -\bar{A}_8 & s\bar{E}_3 - \bar{A}_9 & 0 \\ 0 & -\tilde{A}_{11} & s\tilde{E}_4 - \tilde{A}_{12} & 0 \\ 0 & -\bar{A}_{14} & -\bar{A}_{15} & 0 \\ -\bar{A}_{16} & -\bar{A}_{17} & -\bar{A}_{18} & \bar{M}_6 \\ \hline 0 & 0 & I_p & 0 \end{array}\right]. \tag{7.90}$$

where $R(s)$ is the Rosenbrock matrix of (E_0, A_0, M_0, C_0), and the values of s that make $R(s)$ lose rank are the zeros of the system (E_0, A_0, M_0, C_0) [109]. Further expanding and simplifying $R(s)$ using (7.19) yields

$$\operatorname{rank}\left(R(s)\right) = \operatorname{rank}\underbrace{\left[\begin{array}{ccc|c|ccc} 0 & 0 & 0 & sI_{\bar{n}-p} - \bar{A}_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\tilde{A}_{51} & I_{q1} & 0 & 0 \\ 0 & -I_{q21} & 0 & -\tilde{A}_{52} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\tilde{A}_{53} & 0 & 0 & I_{q22} \\ \hline -I_{n-\bar{n}-a} & 0 & 0 & -\bar{A}_8 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\tilde{A}_{11} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\tilde{A}_{14} & 0 & 0 & 0 \\ \hline 0 & -I_{q21} & 0 & -\tilde{A}_{171} & 0 & -I_{q21} & 0 \\ 0 & 0 & -I_{a-q21} & -\tilde{A}_{172} & 0 & 0 & 0 \end{array}\right]}_{R_2(s)} + p. \tag{7.91}$$

Simplifying $R_2(s)$ then gives the following:

$$\operatorname{rank}\left(R_2(s)\right) = \operatorname{rank}\left[\begin{array}{c} sI_{\bar{n}-p} - \bar{A}_2 \\ \hline \tilde{A}_{11} \\ \bar{A}_{14} \end{array}\right] + q + n - \bar{n} = \operatorname{rank}\underbrace{\begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ \bar{A}_{a11} \end{bmatrix}}_{R_3(s)} + q + n - \bar{n}. \tag{7.92}$$

From the Popov-Hautus-Rosenbrock (PHR) rank test [60], if the values of s that make $R_3(s)$ lose rank (i.e. the unobservable modes of $\bar{A}_2$) are stable, the pair $\left(\bar{A}_2, \bar{A}_{a11}\right)$ is said to be *detectable* - hence Q3 can be recast as: $\left(\bar{A}_2, \bar{A}_{a11}\right)$ is *detectable*.

<u>Proof of Necessity:</u>

Recall that for error state equation (7.60) to be stable, it is required that $\breve{A}$ is stable, which implies $\lambda\left(\bar{A}_2 + R_{23}\bar{A}_{11}\right) < 0$, i.e. $\left(\bar{A}_2, \bar{A}_{11}\right)$ is *detectable*. From (7.19) and (7.26)–(7.27),

$$\bar{A}_{11} = T_{51}\tilde{A}_{11} + T_{52}\bar{A}_{14} = \underbrace{\begin{bmatrix} T_{51} & T_{52} \end{bmatrix}}_{T_5}\bar{A}_{a11}. \tag{7.93}$$

Equation (7.93) implies that the detectability of $\left(\bar{A}_2, \bar{A}_{11}\right)$ depends on T_5, which constitutes design freedom. Hence, the requirement is recast as: matrix T_5 exists such that a

is detectable. From the PHR rank test, if the values of s that make the following matrix $R_4(s)$ lose rank are stable, then $\left(\bar{A}_2, \bar{A}_{11}\right)$ is said to be *detectable*, whereby

$$R_4(s) = \begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ \bar{A}_{11} \end{bmatrix} = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_5 \end{bmatrix} \underbrace{\begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ \bar{A}_{a11} \end{bmatrix}}_{R_3(s)}. \tag{7.94}$$

From (7.94) it follows that $\operatorname{rank}(R_4(s)) \leq \operatorname{rank}(R_3(s))$. Therefore if a value of s makes $R_3(s)$ lose rank, it would also make $R_4(s)$ lose rank, and hence the zeros of (E_0, A_0, M_0, C_0) are also the unobservable modes of $\left(\bar{A}_2, \bar{A}_{11}\right)$. This therefore shows that Q3 is necessary for $\left(\bar{A}_2, \bar{A}_{11}\right)$ to be detectable (i.e. for R4 to be satisfied).

<u>Proof of Sufficiency:</u>

For the case where $p - q - n + \bar{n} + a > 0$, let Z be a matrix containing the generalised right-eigenvectors of $\bar{A}_2$; therefore $Z^{-1}\bar{A}_2 Z$ is a matrix in the *Jordan canonical form*, where the diagonal elements are the real parts of the eigenvalues of $\bar{A}_2$ [98]. Pre-multiply $R_4(s)$ from (7.94) with $\begin{bmatrix} Z^{-1} & 0 \\ 0 & I_{p-q-n+\bar{n}+a} \end{bmatrix}$ and post-multiply with Z to obtain

$$\begin{bmatrix} Z^{-1} & 0 \\ 0 & I_{p-q-n+\bar{n}+a} \end{bmatrix} \begin{bmatrix} sI_{\bar{n}-p} - \bar{A}_2 \\ \bar{A}_{11} \end{bmatrix} Z = \begin{bmatrix} I_{\bar{n}-p} & 0 \\ 0 & T_5 \end{bmatrix} \begin{bmatrix} Z^{-1}\left(sI_{\bar{n}-p} - \bar{A}_2\right) Z \\ \bar{A}_{a11} Z \end{bmatrix}. \tag{7.95}$$

A zero of (E_0, A_0, M_0, C_0), which is an unobservable mode of $\left(Z^{-1}\bar{A}_2 Z, \bar{A}_{11} Z\right)$, will therefore appear as an element of $Z^{-1}\bar{A}_2 Z$ where its corresponding column in $\bar{A}_{11} Z$ is zero. If Q3 is satisfied, however, the columns within $\bar{A}_{a11} Z$ corresponding to elements within $Z^{-1}\bar{A}_2 Z$ would be non-zero. Recall that T_5 is design freedom; therefore, a single row within T_5 can be chosen such that $\bar{A}_{11} Z$ has non-zero elements at the columns corresponding to the diagonals of $Z^{-1}\bar{A}_2 Z$ that are unstable. Thus, the sufficiency of Q4 when $p - q - n + \bar{n} + a > 0$ is proven.

When $p - q - n + \bar{n} + a = 0,\ T_p$ in (7.26) becomes

$$T_p = \begin{bmatrix} I_{q+n-\bar{n}-a} & 0 \end{bmatrix}. \tag{7.96}$$

Since T_5 does not exist, then $\breve{A}$ in (7.66) becomes $\bar{A}_2$. Hence it is sufficient that the eigenvalues of $\bar{A}_2$ are stable (i.e. Q4 is satisfied) for error system (7.60)–(7.61) (and therefore (7.66)–(7.67)) to be made stable (i.e. R4 is satisfied). ■

Thus, Theorem 7.1 is proven. The following summarises the strengths of the proposed scheme over previous works:

- The work presented in this chapter supersedes the work presented in [17, 100] and Chapter 4 as the scheme in this chapter treats less states as unknown inputs (by re-expressing some as a linear combination of other states instead), thereby reducing the number of sensors potentially required for the scheme.
- Since the work presented in Chapters 5 and 6 build on the technique developed in Chapter 4, the scheme presented in this chapter could also further reduce the number of sensors potentially required by the schemes in Chapters 5 and 6.
- The proposed scheme is able to reconstruct faults robustly against disturbances. This improves on previous findings in [99, 100, 101, 161], which did not consider the effect of external disturbances on the fault reconstruction.

7.3.5 Design procedure

The design procedure for the proposed observer scheme can be summarised as follows:

1. Determine T_a, T_b, T_c, T_d from (7.3)–(7.4) and (7.6), and apply the state equation transformation T_dT_b and the state transformation T_cT_a.
2. Check that Q1–Q3 (and Q4 if necessary) given in Theorem 7.1 hold for the system. If they do not hold, do not continue as the observer scheme is not applicable.
3. Determine T_e, T_f, T_g, T_h, T_i from (7.22)–(7.24). Apply the state equation transformation T_{pre}, the state transformation T_{aft}, and the fault transformation T_g according to (7.25).
4. Set values for T_{51} and T_{52} in (7.26) such that the pair $\left(\bar{A}_2, \bar{A}_{11}\right)$ is detectable.
5. Derive the reduced-order system (7.37)–(7.38) using (7.31)–(7.36).

6. Choose R_4 such that R_4 is invertible, and $R_4\bar{E}_2 + V_2 = I_p$.

7. Choose values for LMI design parameters $\bar{D}, \gamma_0$.

8. Use an LMI solver to determine P, L_0, J, γ from (7.73)–(7.74). Then calculate R_{23}, G_l from (7.75), and G_n from (7.42) and (7.53).

9. Set ρ to satisfy (7.54).

10. Reconstruct f according to (7.63).

7.4 Simulation example

The efficacy of the proposed scheme is demonstrated through the following example: consider a modified version of the chemical mixing tank in [157] described by the following dynamical model:

$$\dot{c}_3 = -0.375c_3 - 0.0667q_3 + 0.1q_1, \tag{7.97}$$

$$0 = -q_3 + q_1, \tag{7.98}$$

$$\dot{c}_5 = 0.3c_3 + 0.0533q_3 - 0.5c_5 - 0.04q_5 + 0.02q_4, \tag{7.99}$$

$$0 = q_3 - q_5 + q_4, \tag{7.100}$$

where q_1 is the flow rate of the influent into the first tank, c_3 and q_3 are the concentration and flow rate of the influent from the first tank into the second tank, respectively, q_4 is the flow rate of influent from another pipe into the second tank, and c_5 and q_5 are the concentration and flow rate of the effluent from the second tank, respectively. In this example, the influents from external sources are treated as faults, i.e.

$$f_a = q_1, \; f_b = q_4, \tag{7.101}$$

where f_a and f_b are the faults.

7.4.1 System formulation

The system matrices (E_0, A_0, M_0) in the framework of (7.1)–(7.2) are

$$\begin{aligned} E_0 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad M_0 = \begin{bmatrix} 0.1 & 0 \\ 1 & 0 \\ 0 & 0.02 \\ 0 & 1 \end{bmatrix}, \\ A_0 &= \begin{bmatrix} -0.375 & -0.0667 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0.3 & 0.0533 & -0.5 & -0.04 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \end{aligned} \tag{7.102}$$

for the system variables

$$x_0 = \begin{bmatrix} \text{concentration, } c_3 \text{ (mol/l)} \\ \text{flow rate, } q_3 \text{ (l/s)} \\ \text{concentration, } c_5 \text{ (mol/l)} \\ \text{flow rate, } q_5 \text{ (l/s)} \end{bmatrix}, \quad f_0 = \begin{bmatrix} f_a \text{ (l/s)} \\ f_b \text{ (l/s)} \end{bmatrix}. \tag{7.103}$$

To showcase the robustness of the observer, the pipes connecting the tanks are assumed to leak, causing the flow-rates to fluctuate by ξ_1 and ξ_2. Hence define the disturbance signal $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$, and assume that measurements are only available for c_5 and q_5, so Q_0 and C_0 have the form

$$Q_0 = \begin{bmatrix} 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.104}$$

7.4.2 Observer design

To ease readability, the steps outlined in §7.3.5 will be referred to in the following design of the observer scheme.

Step 1: To obtain the system in the form (7.17)–(7.19), apply the state equation transformation TdT_b and the state transformation T_cT_a, where

$$T_dT_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; T_cT_a = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.105}$$

Steps 2–3: The following partitions are obtained:

$$M_{c2} = \begin{bmatrix} 1 & 0 \\ 0 & 0.02 \\ 0 & 1 \end{bmatrix}, \; A_{c4} = \begin{bmatrix} 1 \\ -0.0533 \\ -1 \end{bmatrix}, \; A_7 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{7.106}$$

From (7.106), the following can then be verified:

$$\operatorname{rank}\begin{bmatrix} M_{c2} & A_{c4} \end{bmatrix} = 3 \; (= q + n - \bar{n}), \; p = 3 > 2 \; (= q + n - \bar{n} - a), \tag{7.107}$$

which shows that Q1–Q3 hold for the system, and the transformations introduced in Proposition 7.1 exist for the system in (7.102) and (7.104). Since $p - q - n + \bar{n} + a = 0$, the satisfaction of Q4 would need to be verified. Apply the state equation transformation T_{pre}, the state transformation T_{aft}, and the fault transformation T_g, where

$$T_{pre} = \begin{bmatrix} 1 & -0.1 & -1 & 0.02 \\ 0 & 0 & -30.03 & 0.6006 \\ 0 & 0.7071 & 0 & 0.7071 \\ 0 & 0.5 & 0 & -0.5 \end{bmatrix}, \; T_{aft} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; T_g = \begin{bmatrix} -0.5 & 0.5 \\ 0.7071 & 0.7071 \end{bmatrix}. \tag{7.108}$$

Note that because $n - \bar{n} - a + q_1 = 0$, A_{41} and M_{21} (and therefore T_{h1}) do not exist, and $T_h = I_4$. The value of $\bar{A}_2$ is found to be

$$\bar{A}_2 = -0.675 < 0, \tag{7.109}$$

thus showing Q4 is also satisfied, thereby guaranteeing the existence of the observer scheme proposed in this chapter. The proposed observer scheme is now designed.

Step 4: Since $p - q - n + \bar{n} + a = 0$, T_5 does not exist, and therefore T in (7.26) becomes

$$T = I_4. \tag{7.110}$$

Step 5: Furthermore, since $a - q_{21} = 0$, x_{122} in (7.18) does not exist. Thus, x_{121} is re-expressed using (7.31) as follows:

$$x_{121} = \left[\begin{array}{c|cc} 0 & 0 & -0.05 \end{array}\right] \begin{bmatrix} x_2 \\ y \end{bmatrix} + \begin{bmatrix} 0.5 & -0.5 \end{bmatrix} \xi. \tag{7.111}$$

By treating x_{11} as an unknown input and using (7.35)–(7.36), the following reduced-order system is formed:

$$\begin{gathered} \bar{E} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -30.03 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \bar{A} = \begin{bmatrix} -0.675 & 0.5 & 0.02 \\ -9.009 & 15.01 & 0.1006 \\ 0 & 0 & -0.7071 \end{bmatrix}, \ \bar{M} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ \bar{Q} = \begin{bmatrix} 0.1 & -0.02 \\ 0.5 & -1.100 \\ -0.7071 & -0.7071 \end{bmatrix}, \ \bar{C} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{gathered} \tag{7.112}$$

Step 6: Since $p - q - n + \bar{n} + a = 0$, R_{23} in (7.51) does not exist, and therefore R_4 and V are chosen such that

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.0333 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ V = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{7.113}$$

Steps 7–8: The LMI design parameters in Proposition 7.3 are chosen as $\bar{D} = I_2$ and $\gamma_0 = 2$. By using the LMI Control Toolbox within MATLAB on LMIs (7.73)–(7.74), the following values of P and γ are obtained:

$$P = \begin{bmatrix} 122.5 & 0 & 0 \\ 0 & 704.0 & 19.88 \\ 0 & 19.88 & 2.612 \end{bmatrix}, \ \gamma = 1.546. \tag{7.114}$$

The observer parameters are found to be:

$$G_l = G_n = \begin{bmatrix} 0 & 0 \\ 0.0018 & -0.0138 \\ -0.0138 & 0.4876 \end{bmatrix}, \ L = I_2, \tag{7.115}$$

where L_0 does not exist since $p - q - n + \bar{n} + a = 0$. The poles of the observer are found to be $\{-0.1914, -0.9855, -1.194\}$, while the pole of its sliding motion is $\{-0.675\}$.

Step 9: The parameter ρ is set as

$$\rho = 50. \tag{7.116}$$

Remark 7.2 *The design of the observer (i.e. the scheme in this chapter) is complete. It will now be shown that existing fault reconstruction schemes for descriptor systems will not work for the example in* (7.102) *and* (7.104) *as follows:*

- *System* (7.102) *and* (7.104) *is not infinitely observable, and hence the schemes in [58, 104] are not applicable.*
- *The SMO schemes in [99, 100, 101, 161] and Chapter 5 do not consider robustness, and therefore are not applicable.*
- *Since* $p = q$*, E3 is not satisfied, and therefore the scheme presented in Chapter 4 [17] is not applicable.*
- *Since* $\mathrm{rank}(M_2) = 2 > 2p - n = 0$*, N2 is not satisfied, and the cascaded SMO scheme presented in Chapter 6 is not applicable.* ♯

7.4.3 Simulation results

To showcase the effectiveness of the proposed scheme, two scenarios were simulated. The first scenario is disturbance-free $(\xi = 0)$ to show the efficacy of the fault reconstruction. In the second scenario, the disturbance signals are set to be non-zero to show the effectiveness of the approach at bounding the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction as follows:

$$\xi_1 = 0.1\sin(0.1t) + 0.12,\ \xi_2 = 0.05\sin\left(0.2t + \frac{\pi}{2}\right) + 0.07. \tag{7.117}$$

In both scenarios for the simulation, the initial condition of the system was set as $x(0) = \{2, 0.5398, 3, 0.8198\}$, while the initial condition of the observer was set at zero. The fault signals were simulated as

$$\begin{aligned} f_a &= 0.3\sin\left(0.3t + \frac{\pi}{3}\right) + 0.4 + 0.15u(t-35), \\ f_b &= 0.2\sin\left(0.2t + \frac{5\pi}{6}\right) + 0.3 - 0.05u(t-50), \end{aligned} \tag{7.118}$$

where $u(t-a)$ is the Heaviside unit step function.

Remark 7.3 *Since f_a and f_b are both discontinuous and time-varying, the schemes by [102, 155] are not applicable. The scheme in [102] considers only constant faults (where $\dot{f} = 0$), which is clearly not satisfied in this case. The scheme in [155] assumes the fault is continuous, which is also not satisfied due to the presence of the Heaviside step function in the fault signals.* ♯

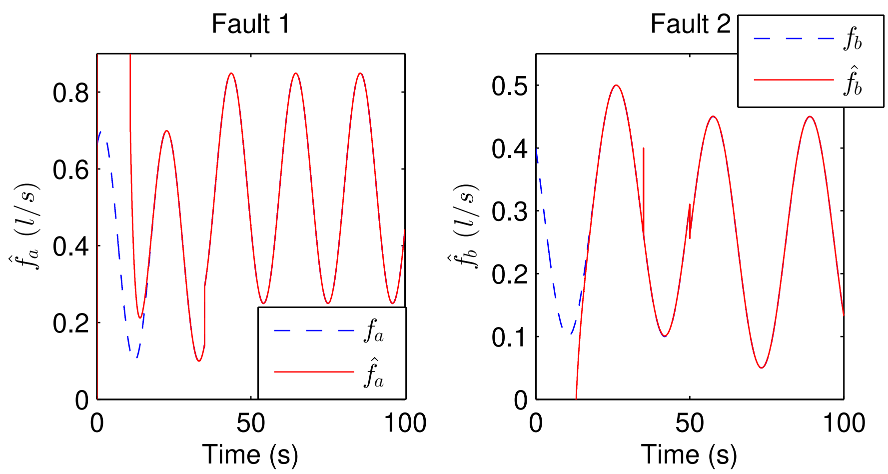

Figure 7.1: Faults (dash-dotted) and their reconstructions (solid) in the disturbance-free scenario.

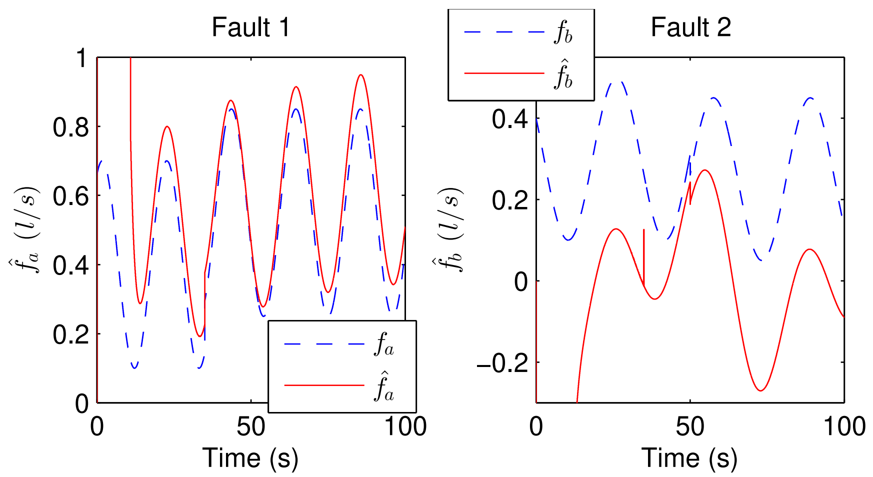

Figure 7.2: Faults (dash-dotted) and their reconstructions (solid) in the scenario with non-zero disturbances.

Figure 7.1 shows the fault reconstructions in the disturbance-free scenario. It can be seen that asymptotic fault reconstruction is achieved, verifying the effectiveness of

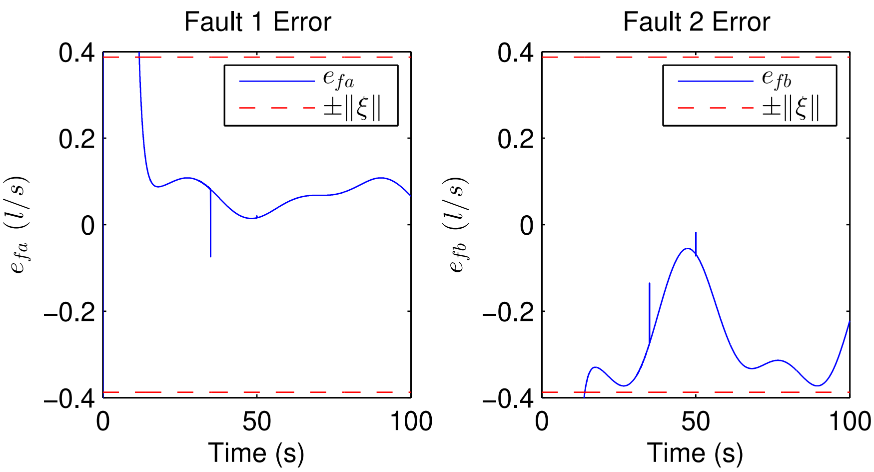

Figure 7.3: Fault reconstruction errors (solid) in the scenario with non-zero disturbances. The dashed lines represent the upper bounds of e_f derived from LMIs (7.73)–(7.74).

the proposed method of fault reconstruction as per Remark 7.1. Figure 7.2 shows the reconstruction of the faults in the scenario with non-zero disturbances. The fault reconstructions are visibly affected, but Figure 7.3 shows that the magnitudes of the fault reconstruction errors are bounded within $\pm\|\xi\|$, thus validating that the method is effective at bounding the effect of the disturbances on the fault reconstructions.

7.5 Conclusion

This chapter has presented a robust fault reconstruction scheme for a class of NIODS using SMOs, which improves on previous work that required a larger number of measurable outputs, or did not consider disturbances and their effects on the fault reconstruction. This was accomplished by first re-expressing certain states as a linear combination of other states, and then treating some other states as unknown inputs. A standard SMO scheme was then applied onto the reduced-order system to reconstruct the fault signal. The necessary and sufficient conditions for the existence of the scheme were investigated and presented in terms of the original system matrices. LMI techniques were used to minimise the effect of disturbances on the fault reconstruction, and the design procedure for the observer scheme was shown. Finally, a simulation was carried out, and the results verify the efficacy of the scheme.

Chapter 8

Conclusion and recommendations for future research

This chapter concludes the thesis and summarises its main contributions. Possible future work along the lines of the presented work are also recommended for consideration.

Chapter 3 discussed the sliding mode observer (SMO) by Yeu et al. [157] is presented. The SMO by Yeu et al. is able to estimate states and faults for descriptor systems, but requires the system to be *infinitely observable*. Two methods to re-express non-infinitely observable descriptor systems (NIODS) as infinitely observable reduced-order systems were then presented [100, 101]. In both methods, certain states were treated as unknown inputs; in the second method, some states are also treated as a linear combination of other states. After this re-expression, the SMO by Yeu et al. could be applied onto the system to estimate the states and faults of the original descriptor system. Finally, a simulation example utilising the second method was shown.

Chapter 4 presented a robust fault reconstruction scheme for a class of NIODS using SMOs. This work improves on the scheme in [100] by incorporating robust fault reconstruction via design of the observer gains. Certain states were removed and treated as unknown inputs, thereby formulating an infinitely-observable reduced-order system. The observer by Yeu et al. [157] was then applied onto the reduced-order system to reconstruct the fault signal. The necessary and sufficient conditions guaranteeing the existence of the scheme were presented in terms of the original system matrices. LMI techniques were utilised to minimise the effect of disturbances on the fault reconstruction, and a summarised design procedure for the observer scheme was shown. A simulation using a chemical mixing tank model was performed, and the results verified the efficacy of the scheme.

Chapter 5 improved on the scheme developed in chapter 4 by utilising two SMOs in cascade to perform state and fault estimation for a class of NIODS. Certain states were treated as unknown inputs to formulate an infinitely observable reduced-order system. The observer by Yeu et al. [157] was applied onto this reduced-order system to estimate the states and some components of the fault in finite time. The switching term is found to be the output of an analytical (non-descriptor) state-space system treating the remaining unestimated components of the fault as unknown inputs. Thus the switching term was fed into a second SMO (which is an Edwards-Spurgeon SMO [8]) to estimate the remaining faults in finite time. The necessary and sufficient conditions for the existence of the scheme were presented in terms of the original system matrices; these conditions are found to be more relaxed than those for existing state and fault estimation schemes for NIODS. Finally, a simulation was carried out, and the results verify the efficacy of the scheme.

Chapter 6 built on the findings in chapter 5 by performing robust fault reconstruction for a class of NIODS utilising two SMOs in cascade. The approach presented in chapter 5 was used to reconstruct the fault. LMI techniques were used to design the gains of the observers such that the $\mathcal{L}_2$ gain from the disturbances to the fault reconstruction is minimised. The necessary and sufficient conditions for the existence of the scheme were investigated and are found to be more relaxed than those for existing fault reconstruction schemes for NIODS. The LMIs were also formulated such that the gains for both observers are designed simultaneously, and are thus less conservative than existing methods, which design each observer separately. Finally, a simulation was carried out and its results verify the efficacy of the scheme.

Chapter 7 then improved on the work in the previous chapters by first re-expressing certain states as a linear combination of other states, before treating certain other states as unknown inputs to formulate an infinitely observable descriptor system. This technique would potentially reduce the number of measurable outputs, since the preceding work may treat a larger number of states as unknown inputs compared to when this method of reformulation is used. The chapter also improved on the work by [101] by also considering robust fault reconstruction via design of the observer gains. The observer by Yeu et al. [157] was then applied onto the reduced-order system to reconstruct the fault signal. The necessary and sufficient conditions for the existence of the scheme were investigated and presented in terms of the original system matrices. LMI techniques were used to minimise the effect of disturbances on the fault reconstruction, and the design procedure for the observer scheme was shown. Finally, a simulation was carried out, and the results verify the efficacy of the scheme.

8.1 Recommendations for Future Work

This thesis presented investigations into state estimation and fault reconstruction. The systems being considered however are linear, which greatly limits the applicability of the observer schemes. Practical systems are often non-linear and operate at conditions that are significantly different from their initial conditions. These large excursions due to non-linearities may cause the system to generate false alarms, or even mask the effects of a fault occurring altogether [28, 149, 150]. Thus a possible extension of current work would be to consider the effect of non-linearities on state estimation and fault reconstruction in non-infinitely observable descriptor systems. The LMI techniques by [128] could be deployed to minimise the effects of the non-linearities on the state estimation and fault reconstruction errors.

Another consideration for the systems under study would be the inclusion of time-delays in the system. *Aftereffects*, where the past performance of the system affects the current behaviour (which gives rise to the term 'time-delay'), are prevalent across many different fields [106]. These time-delays give rise to different considerations, such as the stability of the system [50]. There have been investigations into observers for time-delay systems [96, 131], but current schemes for descriptor systems with time-delays require the system to be infinitely observable [167]. Hence the reformulation technique utilised in this thesis can be used to further extend the applicability of observers into non-infinitely observable descriptor systems (NIODS).

The scheme presented in Chapter 5 is for full-state estimation. For systems with a large number of states (i.e. high-dimensional systems), however, estimating the entire state vector could be computationally demanding and impractical. Furthermore, in many cases, information regarding only a subset of the state vector is required; not all states need to be estimated. Functional observers are commonly used to estimate a part of the state vector (as opposed to the entire state vector) [71, 110, 132]. Existing results for functional observation in descriptor systems however require the system to be infinitely observable [73]. Thus one possible future development would be to extend the use of the reformulation technique in this thesis into the functional estimation of states for NIODS.

Tan and Edwards [129] have developed a robust fault reconstruction scheme for state-space systems using multiple SMOs in cascade. By increasing the number of SMOs in cascade (up until the number of observers equals the number of states), it is found that the matching condition can be relaxed even further, thereby increasing the applicability of the scheme. These findings are however only applicable to state-space systems. Furthermore, current investigations for the use of SMOs in cascade for descriptor systems have only considered at most two SMOs in cascade (the schemes presented in Chapters 5 and 6, [99, 161]). Therefore, a potential line of inquiry would be studying the use of multiple SMOs in cascade for NIODS.

There exist a class of systems known as *singularly perturbed systems*, where the states of the system evolve on different time-scales, resulting in a *slow* and a *fast subsystem* [10, 77, 145]. The implications on the mathematical model of the system would be that the time-derivative of certain states are multiplied with a small parameter $\varepsilon \ll 1$, resulting in a *quasi-infinitely observable descriptor system*. Therefore, a possible line of future investigation would be to apply the reformulation technique in this thesis onto singularly-perturbed systems and to minimise the effects of ε on the state estimates and fault reconstructions using LMI techniques introduced in [128].

Finally, the findings in this thesis have been verified through relatively simple examples. Practical engineering systems are however more complex and contain intricacies absent in the examples used. Therefore, a case study applying the presented schemes onto practical systems to demonstrate applicability could be a point for future development.

References

[1] O.M. Aamo. Leak detection, size estimation and localization in pipe flows. *IEEE Transactions on Automatic Control*, 61(1):246–251, 2016.

[2] A. Al-Ghanimi, J. Zheng, and Z. Man. A fast non-singular terminal sliding mode control based on perturbation estimation for piezoelectric actuators systems. *International Journal of Control*, 90(3):480–491, 2017.

[3] H.K. Alaei and A. Yazdizadeh. Robust output disturbance, actuator, and sensor faults reconstruction using $\mathcal{H}_\infty$ sliding mode descriptor observer for uncertain nonlinear boiler system. *International Journal of Control, Automation and Systems*, 16(3):1271–1281, 2018.

[4] M. Alma, H.S. Ali, and M. Darouach. Adaptive oberver design for linear descriptor systems. In *Proceedings of the American Control Conference (ACC)*, pages 5144–5149, Milwaukee, WI, 2018. IEEE.

[5] M. Alma and M. Darouach. Adaptive observers design for a class of linear descriptor systems. *Automatica*, 50(2):578–583, February 2014.

[6] H. Alwi and C. Edwards. Fault tolerant control of a civil aircraft using a sliding mode based scheme. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 1011–1016, Seville, December 2005. IEEE.

[7] H. Alwi and C. Edwards. Fault detection and fault-tolerant control of a civil aircraft using a sliding-mode-based scheme. *IEEE Transactions on Control Systems Technology*, 16(3):499–510, 2008.

[8] H. Alwi, C. Edwards, and C.P. Tan. *Fault detection and fault-tolerant control using sliding modes*. Advances in Industrial Control. Springer, London, 2011.

[9] L. Besnard, Y.B. Shtessel, and B. Landrum. Quadrotor vehicle control via sliding mode controller driven by sliding mode disturbance observer. *Journal of the Franklin Institute*, 349(2):658–684, 2012.

[10] Z. Bougatef, N. Abdelkrim, A. Tellili, and M.N. Abdelkrim. Fault diagnosis and accommodation for singularly perturbed time-delayed systems: descriptor approach. In *Proceedings of the 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering*, pages 86–92, Monastir, 2017.

[11] B. Boulkroune, S. Halabi, and A. Zemouche. Fault detection filter for a class of nonlinear descriptor systems. *International Journal of Control*, 86(2):253–262, February 2013.

[12] M. Buciakowski, M. Witczak, V. Puig, D. Rotondo, F. Nejjari, and J. Korbicz. A bounded-error approach to simultaneous state and actuator fault estimation for a class of nonlinear systems. *Journal of Process Control*, 52:14–25, 2017.

[13] B. Cantó, C. Coll, and E. Sánchez. Positive $n-$periodic descriptor control systems. *Systems & Control Letters*, 53(5):407–414, 2004.

[14] C. Canudas de Wit and J.J.E. Slotine. Sliding observers for robot manipulators. *Automatica*, 27(5):859–864, September 1991.

[15] A. Chakrabarty, M.J. Corless, G.T. Buzzard, S.H. Żak, and A.E. Rundell. State and unknown input observers for nonlinear systems with bounded exogenous inputs. *IEEE Transactions on Automatic Control*, 62(11):5497–5510, 2017.

[16] A. Chakrabarty, A.E. Rundell, S.H. Żak, F. Zhu, and G.T. Buzzard. Unknown input estimation for nonlinear systems using sliding mode observers and smooth window functions. *SIAM Journal on Control and Optimization*, 56(5):3619–3641, 2018.

[17] J.C.L. Chan, C.P. Tan, and H. Trinh. Robust fault reconstruction for a class of infinitely unobservable descriptor systems. *International Journal of Systems Science*, 48(8):1646–1655, 2017.

[18] J.C.L. Chan, C.P. Tan, H. Trinh, and M.A.S. Kamal. State and fault estimation for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade. *Journal of the Franklin Institute*, 356(5):3010–3029, 2019.

[19] J.C.L. Chan, C.P. Tan, H. Trinh, M.A.S. Kamal, and Y.S. Chiew. Robust fault reconstruction for a class of non-infinitely observable descriptor systems using two sliding mode observers in cascade. *Applied Mathematics and Computation*, 350:78–92, 2019.

[20] F. Chen and M.W. Dunnigan. Comparative study of a sliding-mode observer and Kalman filters for full state estimation in an induction machine. *IEE Proceedings Electric Power Applications*, 149(1):53–64, January 2002.

[21] M. Chen, P. Shi, and C. Lim. Robust constrained control for MIMO nonlinear systems based on disturbance observer. *IEEE Transactions on Automatic Control*, 60(12):3281–3286, 2015.

[22] W. Chen, A. Fekih, and Z. Mao. Fault detection, estimation/reconstruction, and fault-tolerant control: theory and practice. *Mathematical Problems in Engineering*, 2016:2, 2016. Art. no. 5293171.

[23] X. Chen, W. Shen, Z. Cao, and A. Kapoor. A novel approach for state of charge estimation based on adaptive switching gain sliding mode observer in electric vehicles. *Journal of Power Sources*, 246:667–678, 2014.

[24] Y. Chen and G. Duan. Conditions for C-controllability and C-observability of rectangular descriptor systems. In *Proceedings of the 6th World Congress on Intelligent Control and Automation*, pages 629–630, Dalian, 2006.

[25] Y. Chen, Z. Wang, W. Qian, and F.E. Alsaadi. Asynchronous observer-based $\mathcal{H}_\infty$ control for switched stochastic systems with mixed delays under quantization and packet dropouts. *Nonlinear Analysis: Hybrid Systems*, 27:225–238, 2018.

[26] M. Chilali and P. Gahinet. $\mathcal{H}_\infty$ design with pole placement constraints: an LMI approach. *IEEE Transactions on Automatic Control*, 41(3):358–367, 1996.

[27] H. Cho, S.J. Yoo, and S. Kwak. State observer based sensor less control using Lyapunov's method for boost converters. *IET Power Electronics*, 8(1):11–19, 2015.

[28] W.S. Chua, C.P. Tan, M. Aldeen, and S. Saha. A robust fault estimation scheme for a class of nonlinear systems. *Asian Journal of Control*, 19(2):779–804, 2017.

[29] D. Cobb. Controllability, observability, and duality in singular systems. *IEEE Transactions on Automatic Control*, 29(12):1076–1082, December 1984.

[30] C. Coll, M.J. Fullana, and E. Sánchez. Some invariants of discrete-time descriptor systems. *Applied Mathematics and Computation*, 127(2):277–287, 2002.

[31] M. Corless and J.A.Y. Tu. State and input estimation for a class of uncertain systems. *Automatica*, 34(6):757–764, 1998.

[32] L. Dan, Y. Ying, and Z. Zhengen. Observer-based fault estimation for nonlinear singularly perturbed systems. In *Proceedings of the 34th Chinese Control Conference (CCC)*, pages 6351–6355, Hangzhou, 2015. IEEE.

[33] A.P. Dani, Z. Kan, N.R. Fischer, and W.E. Dixon. Structure estimation of a moving object using a moving camera: an unknown input observer approach. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 5005–5010, Orlando, FL, 2011. IEEE.

[34] M. Darouach. Observers and observer-based control for descriptor systems revisited. *IEEE Transactions on Automatic Control*, 59(5):1367–1373, 2014.

[35] M. Darouach, F. Amato, and M. Alma. Functional observers design for descriptor systems via lmi: continuous and discrete-time cases. *Automatica*, 86:216–219, 2017.

[36] M. Darouach, M. Zasadzinski, and S.J. Xu. Full-order observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, 39(3):606–609, 1994.

[37] A. Di Giorgio, A. Pietrabissa, F.D. Priscoli, and A. Isidori. Robust protection scheme against cyber-physical attacks in power systems. *IET Control Theory & Applications*, 12(13):1792–1801, 2018.

[38] X. Ding, X. Liu, and S. Zhong. Delay-independent criteria for exponential admissibility of switched descriptor systems. *Applied Mathematics and Computation*, 228:432–445, 2014.

[39] D. Du. Fault detection for discrete-time linear systems based on descriptor observer approach. *Applied Mathematics and Computation*, 293:575–585, 2017.

[40] N.H. Du, N.C. Liem, C.J. Chyan, and S.W. Lin. Lyapunov stability of quasilinear implicit dynamic equations on time scales. *Journal of Inequalities and Applications*, 2011(1):1–27, 2011. Art. no. 979705.

[41] G.R. Duan. *Analysis and design of descriptor linear systems*. Advances in Mechanics and Mathematics. Springer, New York, 2010.

[42] C. Edwards. A comparison of sliding mode and unknown input observers for fault reconstruction. In *Proceedings of the 43rd IEEE Conference on Decision and Control*, pages 5279–5284, Nassau, December 2004. IEEE.

[43] C. Edwards, H. Alwi, and C.P. Tan. Sliding mode methods for fault detection and fault tolerant control with application to aerospace systems. *International Journal of Applied Mathematics and Computer Science*, 22(1):109–124, 2012.

[44] C. Edwards and S.K. Spurgeon. On the development of discontinuous observers. *International Journal of Control*, 59(5):1211–1229, May 1994.

[45] C. Edwards and S.K. Spurgeon. *Sliding mode control: theory and applications*. Taylor & Francis Ltd, London, 1998.

[46] C. Edwards, S.K. Spurgeon, and R.J. Patton. Sliding mode observers for fault detection and isolation. *Automatica*, 36(4):541–553, April 2000.

[47] C. Edwards and C.P. Tan. A comparison of sliding mode and unknown input observers for fault reconstruction. *European Journal of Control*, 12(3):245–260, January 2006.

[48] V. Estrada-Manzo, Z. Lendek, and T.M. Guerra. Generalized lmi observer design for discrete-time nonlinear descriptor models. *Neurocomputing*, 182:210–220, 2016.

[49] M.M. Fahmy and J. O'Reilly. Observers for descriptor systems. *International Journal of Control*, 49(6):2013–2028, June 1989.

[50] E. Fridman. Tutorial on Lyapunov-based methods for time-delay systems. *European Journal of Control*, 20(6):271–283, 2014.

[51] L. Fridman, J. Moreno, and R. Iriarte, editors. *Sliding Modes after the First Decade of the 21st Century*, volume 412 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag Berlin Heidelberg, Berlin, 2012.

[52] P.M. Gahinet, A. Nemirovskii, A.J. Laub, and M. Chilali. *LMI Control Toolbox User's Guide*. The MathWorks, Inc., Natick, MA, 1995.

[53] G. Gandolfo. *Economic Dynamics: Methods and Models*. North-Holland, Amsterdam, 1985.

[54] Z. Gao and S. Ding. Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems. *Automatica*, 43(5):912–920, May 2007.

[55] Z. Gao and S.X. Ding. Fault reconstruction for Lipschitz nonlinear descriptor systems via linear matrix inequality approach. *Circuits, Systems & Signal Processing*, 27(3):295–308, 2008.

[56] H. Gholami and T. Binazadeh. Observer-based $\mathcal{H}_\infty$ finite-time controller for time-delay nonlinear one-sided lipschitz systems with exogenous disturbances. *Journal of Vibration and Control*, 2018.

[57] W.J. Hahn, A.J. DiLellio, and J.S. Dyer. *Market-calibrated Forecasts for Natural Gas Prices*. Energy Institute, The University of Texas at Austin, Austin, TX, 2016. White Paper UTEI/2016-07-1.

[58] H. Hamdi, M. Rodrigues, C. Mechmeche, and N. Benhadj Braiek. Observer-based fault diagnosis for time-delay LPV descriptor systems. *IFAC-PapersOnLine*, 51(24):1179–1184, 2018.

[59] K. Han, J. Feng, Y. Li, and S. Li. Reduced-order simultaneous state and fault estimator based fault tolerant preview control for discrete-time linear time-invariant systems. *IET Control Theory & Applications*, 12(11):1601–1610, 2018.

[60] M.L.J. Hautus. Controllability and observability conditions of linear autonomous systems. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series A, Mathematical sciences*, 72:443–448, 1969.

[61] M. Hou and P.C. Müller. Design of a class of Luenberger observers for descriptor systems. *IEEE Transactions on Automatic Control*, 40(1):133–136, January 1995.

[62] M. Hou and P.C. Müller. Causal observability of descriptor systems. *IEEE Transactions on Automatic Control*, 44(1):158–163, January 1999.

[63] M. Hou and P.C. Müller. Observer design for descriptor systems. *IEEE Transactions on Automatic Control*, 44(1):164–168, 1999.

[64] L. Hu, Z. Wang, Q.-L. Han, and X. Liu. State estimation under false data injection attacks: security analysis and system protection. *Automatica*, 87:176–183, 2018.

[65] S. Hui and S.H. Żak. Stress estimation using unknown input observer. In *Proceedings of the 2013 American Control Conference*, pages 259–264, Washington, DC, 2013. IEEE.

[66] R. Isermann and P. Ballé. Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5), 1997.

[67] A. Isidori. *Nonlinear Control Systems*. Communications and Control Engineering. Springer-Verlag London, London, 1995.

[68] A. Jafari, H. Singh, J. Karunanayaka, H. Ryu, J. Chong, and A.C. Etoundi. Lyapunov observer/controller for stable haptic interatction. In *Proceedings of the 2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 1–6, Auckland, 2018.

[69] A. Katriniok and D. Abel. Adaptive EKF-based vehicle state estimation with online assessment of local observability. *IEEE Transactions on Control System Technology*, 24(4):1368–1381, 2016.

[70] S. Kawai and N. Hori. Mapping discrete-time models for descriptor-systems with consistent initial conditions. *Transactions of the Canadian Society for Mechanical Engineering*, 40(1):59–77, 2016.

[71] C.Y. Kee, C.P. Tan, K.Y. Ng, and H. Trinh. New results in robust functional state estimation using two sliding mode observers in cascade. *International Journal of Robust and Nonlinear Control*, 24(15):2079 – 2097, February 2014.

[72] D. Koenig, B. Marx, and S. Varrier. Filtering and fault estimation of descriptor switched systems. *Automatica*, 63:116–121, 2016.

[73] J. Lan and R.J. Patton. Robust fault-tolerant control based on a functional observer for linear descriptor systems. *IFAC-PapersOnLine*, 48(14):138–143, 2015.

[74] J. Lan and R.J. Patton. A decoupling approach to integrated fault-tolerant control for linear systems with unmatched non-differentiable faults. *Automatica*, 89:290–299, 2018.

[75] X. Li, C.K. Ahn, D. Lu, and S. Guo. Robust simultaneous fault estimation and non-fragile output feedback fault-tolerant control for Markovian jump systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pages 1–8, 2018.

[76] Y. Li, Y. Tan, R. Dong, and H. Li. State estimation of macromotion positioning tables based on switching Kalman filter. *IEEE Transactions on Control Systems Technology*, 25(3):1076–1083, 2017.

[77] D. Liu, Y. Yang, and Z. Zhao. Observer-based fault estimation for nonlinear singular perturbed systems. In *Proceedings of the 25th Chinese Control Conference*, pages 6351–6355, Guiyang, 2015.

[78] M. Liu and P. Shi. Sensor fault estimation and tolerant control for Itô stochastic systems with a descriptor sliding mode approach. *Automatica*, 49(5):1242–1250, 2013.

[79] Y. Liu, Z. Wang, and D. Zhou. State estimation and fault reconstruction with integral measurements under partially decoupled disturbances. *IET Control Theory & Applications*, 12(10):1520–1526, 2018.

[80] D.G. Luenberger. Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8(2):74–80, April 1964.

[81] D.G. Luenberger. Dynamic equations in descriptor form. *IEEE Transactions on Automatic Control*, 22(3):312–321, June 1977.

[82] J. Marzat, H. Piet-Lahanier, F. Damongeot, and E. Walter. Model-based fault diagnosis for aerospace systems: a survey. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 226(10):1329–1360, 2012.

[83] K. Mathiassen, D. Dall'Alba, R. Muradore, P. Fiorini, and O.J. Elle. Robust real-time needle tracking in 2-D ultrasound images using statistical filtering. *IEEE Transactions on Control Systems Technology*, 25(3):966–978, 2017.

[84] D.Q. Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50(12):2967–2986, 2014.

[85] C. Mellucci, P.P. Menon, C. Edwards, and A. Ferrara. Second-order sliding mode

observers for fault reconstruction in power networks. *IET Control & Applications*, 11(16):2772–2782, 2017.

[86] M. Mousavi, M. Rahnavard, and S. Haddad. Observers based fault reconstruction schemes using terminal sliding modes. *International Journal of Control*, pages 1–8, 2018.

[87] M. Mousavi, M. Rahnavard, M.R.H. Yazdi, and M. Ayati. On the development of terminal sliding mode observers. In *Proceedings of the Iranian Conference on Electrical Engineering (ICEE)*, pages 951–956, Mashhad, 2018.

[88] M. Muehlebach and S. Trimpe. Distributed event-based state estimation for networked systems: an LMI approach. *IEEE Transactions on Automatic Control*, 63(1):269–276, 2018.

[89] P.C. Müller and M. Hou. On the observer design for descriptor systems. *IEEE Transactions on Automatic Control*, 38(11):1666–1671, December 1993.

[90] A.M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, and J. Ragot. State estimation of two-time scale multiple models. Application to wastewater treatment plant. *Control Engineering Practice*, 19(11):1354–1362, November 2011.

[91] K.Y. Ng, C.P. Tan, C. Edwards, and Y.C. Kuang. New results in robust actuator fault reconstruction for linear uncertain systems using sliding mode observers. *International Journal of Robust and Nonlinear Control*, 17(14):1294–1319, September 2007.

[92] K.Y. Ng, C.P. Tan, Z. Man, and R. Akmeliawati. New results in disturbance decoupled fault reconstruction in linear uncertain systems using two sliding mode observers in cascade. *International Journal of Control, Automation and Systems*, 8(3):506–518, 2010.

[93] K.Y. Ng, C.P. Tan, and D. Oetomo. Disturbance decoupled fault reconstruction using cascaded sliding mode observers. *Automatica*, 48(5):794–799, May 2012.

[94] K.Y. Ng, C.P. Tan, and D. Oetomo. Enhanced fault reconstruction using cascaded sliding mode observers. In *Proceedings of the 12th International Workshop on Variable Structure Systems*, pages 208–213, Mumbai, 2012. IEEE.

[95] C.M. Nguyen, P.N. Pathirana, and H. Trinh. Robust observer design for uncertain one-sided Lipschitz systems with disturbances. *International Journal of Robust and Nonlinear Control*, 28(4):1366–1380, 2018.

[96] M.C. Nguyen, H. Trinh, and P.T. Nam. Linear functional observers with guaranteed $\epsilon-$convergence for discrete time-delay systems with input/output disturbances. *International Journal of Systems Science*, 47(13):3193–3205, 2016.

[97] R. Nikhoukhah, S.L. Campbell, and F. Delebecque. Observer design for general linear time-invariant systems. *Automatica*, 34(5):575–583, 1998.

[98] P.J. Olver and C. Shakiban. *Applied Linear Algebra*, volume 215 of *Undergraduate Texts in Mathematics*. Springer, Cham, Cham, 2018.

[99] J.H.T. Ooi, C.P. Tan, W.-S. Chua, and X. Wang. State and unknown input estimation for a class of infinitely unobservable descriptor systems using two observers in cascade. *Journal of the Franklin Institute*, 354(18):8374–8397, 2017.

[100] J.H.T. Ooi, C.P. Tan, and K.Y. Ng. State and fault estimation for infinitely unobservable descriptor systems using sliding mode observers. *Asian Journal of Control*, 59(4):1458–1461, 2015.

[101] J.H.T. Ooi, C.P. Tan, S.G. Nurzaman, and K.Y. Ng. A sliding mode observer for infinitely unobservable descriptor systems. *IEEE Transactions on Automatic Control*, 62(7):3580–3587, 2017.

[102] G. Osorio-Gordillo, C. Astorga-Zaragoza, A. Pérez Estrada, R. Vargas-Méndez, M. Darouach, and L. Boutat-Baddas. Fault estimation for descriptor linear systems based on generalised dynamic observer. *International Journal of Systems Science*, 49(11):2398–2409, 2018.

[103] R.J. Patton. Fault-Tolerant Control. In J. Baillieul and T. Samad, editors, *Encyclopedia of Systems and Control*, pages 422–428. Springer London, London, 2015.

[104] G. Phanomchoeng, A. Zemouche, W. Jeon, R. Rajamani, and F. Mazenc. $\mathcal{H}_\infty$ observer for descriptor nonlinear systems with nonlinear output equations. In *Proceedings of the American Control Conference*, pages 5952–5956, Milwaukee, WI, 2018. IEEE.

[105] K.F. Prochazka, H. Eduardo, and S.R. Klein. Integrated fault-tolerant control of an over-actuated aircraft using optimal control allocation and robust sliding mode observers. In *Proceedings of the IEEE Conference on Control Technology and Applications (CCTA)*, pages 171–178, Copenhagen, 2018. IEEE.

[106] J.-P. Richard. Time-delay systems: an overview of some recent advances and open problems. *Automatica*, 39(10):1667–1694, 2003.

[107] G. Rinaldi, M. Cucuzzella, and A. Ferrara. Sliding mode observers for a network of thermal and hydroelectric power plants. *Automatica*, 98:51–57, 2018.

[108] G. Rinaldi, P.P. Menon, C. Edwards, and A. Ferrara. Sliding mode based dynamic state estimation for synchronous generators in power systems. *IEEE Control Systems Letters*, 2(4):785–790, 2018.

[109] H.H. Rosenbrock. *State space and multivariable theory*. John-Wiley, New York, 1970.

[110] F. Rotella and I. Zambettakis. A note on functional observability. *IEEE Transactions on Automatic Control*, 61(10):3197–3202, 2016.

[111] L. Sadelli, M. Fruchard, and A. Ferreira. 2d observer-based control of a vascular microrobot. *IEEE Transactions on Automatic Control*, 62(5):2194–2206, 2017.

[112] M. Saif and Y. Guan. A new approach to robust fault detection and identification. *IEEE Transactions on Aerospace and Electronic Systems*, 29(3):685–695, July 1993.

[113] F. Shi and R.J. Patton. Fault estimation and active fault tolerant control for linear parameter varying descriptor systems. *International Journal of Robust and Nonlinear Control*, 25(5):689–706, March 2015.

[114] D.N. Shields. Observers for descriptor systems. *International Journal of Control*, 55(1):249–256, 1992.

[115] Y. Shoukry, P. Nuzzo, A. Puggelli, A.L. Sangiovanni-Vincentelli, S.A. Seshia, and P. Tabuada. Secure state estimation for cyber-physical systems under sensor attacks: a satisfiability modulo theory approach. *IEEE Transactions on Automatic Control*, 62(10):4917–4932, 2017.

[116] Y. Shtessel, C. Edwards, L. Fridman, and A. Levant. *Sliding Mode Control and Observation*. Birkhäuser, Basel, 2014.

[117] J. Sjöberg. *Descriptor Systems and Control Theory*. Division of Automatic Control, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2005. Report no. LiTH-ISY-R-2688.

[118] J.J.E. Slotine, J.K. Hedrick, and E.A. Misawa. On sliding observers for nonlinear systems. *ASME Journal of Dynamic Systems, Measurement and Control*, 109(3):245–252, September 1987.

[119] J.J.E. Slotine, J.K. Hendrick, and E.A. Misawa. Nonlinear state estimation using sliding observers. In *Proceedings of the 24th IEEE Conference on Decision and Control*, pages 332–339, Athens, 1986. IEEE.

[120] Y.-D. Song, Y. Lu, and Z.-X. Gan. Descriptor sliding mode approach for fault/noise reconstruction and fault-tolerant control of nonlinear uncertain systems. *Information Sciences*, 367–368:194–208, 2016.

[121] Sarah K. Spurgeon. Sliding mode observers: a survey. *International Journal of Systems Science*, 39(8):751–764, August 2008.

[122] X. Su, Q. Zhang, and H. Shi. Infinite controllability and infinite observability of generalized dynamical systems. *Dynamics of Continuous, Discrete and Impulsive Systems. Series A: Mathematical Analysis*, 11(5–6):825–834, 2004.

[123] C. Tan and X. Zhang. Observability of linear time-invariant descriptor systems with a derivative in the output. *Journal of Applied Mathematics and Computing*, 33(1-2):239–250, May 2010.

[124] C.P. Tan. *Sliding mode observers for fault detection and isolation*. PhD thesis, Engineering Department, University of Leicester, Leicester, 2002.

[125] C.P. Tan, F. Crusca, and M. Aldeen. Extended results on robust state estimation and fault detection. *Automatica*, 44(8):2027–2033, 2008.

[126] C.P. Tan and C. Edwards. Reconstruction of sensor faults using a secondary sliding mode observer. In *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 579–584, Orlando, FL, 2001.

[127] C.P. Tan and C. Edwards. Sliding mode observers for detection and reconstruction of sensor faults. *Automatica*, 38(10):1815–1821, October 2002.

[128] C.P. Tan and C. Edwards. Sliding mode observers for robust detection and reconstruction of actuator and sensor faults. *International Journal of Robust and Nonlinear Control*, 13(5):443–463, April 2003.

[129] C.P. Tan and C. Edwards. Robust fault reconstruction in uncertain linear systems using multiple sliding mode observers in cascade. *IEEE Transactions on Automatic Control*, 55(4):855–867, April 2010.

[130] C.P. Tan, X. Yu, and Z. Man. Terminal sliding mode observers for class of nonlinear systems. *Automatica*, 46(8):1401–1404, 2010.

[131] H.M. Tran, H. Trinh, and P.T. Nam. Functional observer-based fault detection of time-delay systems via an LMI approach. In *Proceedings of the 5th Australian Control Conference*, pages 194–199, Gold Coast, 2015. IEEE.

[132] H. Trinh and T. Fernando. *Functional Observers for Dynamical Systems*, volume 420 of *Lecture Notes in Control and Information Sciences*. Springer Berlin Heidelberg, Berlin, 2012.

[133] V.I. Utkin. Variable structure systems with sliding modes. *IEEE Transactions on Automatic Control*, 22(2):212–222, April 1977.

[134] V.I. Utkin. *Sliding modes in control and optimization*. Communications and Control Engineering Series. Springer-Verlag, Berlin, 1992.

[135] S.L. Varricchio, N. Martins, L.T.G. Lima, and S. Carneiro Jr. Studying harmonic problems using a descriptor system approach. In *Proceedings of the International Conference on Power Systems Transients*, pages 641–646, Budapest, 1999.

[136] K.C. Veluvolu and Y.C. Soh. Fault reconstruction and state estimation with sliding mode observers for lipschitz non-linear systems. *IET Control Theory & Applications*, 5(11):1255–1263, 2011.

[137] G. Verghese, B. Levy, and T. Kailath. A generalized state-space for singular systems. *IEEE Transactions on Automatic Control*, 26(4):811–831, August 1981.

[138] B.L. Walcott, M.J. Corless, and S.H. Żak. Comparative study of non-linear state-observation techniques. *International Journal of Control*, 45(6):2109–2132, 1987.

[139] B.L. Walcott and S.H. Żak. State observation of nonlinear uncertain dynamical systems. *IEEE Transactions on Automatic Control*, 32(2):166–170, February 1987.

[140] B.L. Walcott and S.H. Żak. Combined observer-controller synthesis for uncertain dynamical systems with applications. *IEEE Transactions on Systems, Man, and Cybernetics*, 18(1):88–104, 1988.

[141] W.Q. Wang and Y. Zou. Analysis of impulsive modes and Luenberger observers for descriptor systems. *Systems & Control Letters*, 44(5):347–353, December 2001.

[142] X. Wang, C.P. Tan, and D. Zhou. A novel sliding mode observer for state and fault estimation in systems not satisfying matching and minimum phase conditions. *Automatica*, 79:290–295, 2017.

[143] Z. Wang, M. Rodrigues, D. Theilliol, and Y. Shen. Actuator fault estimation observer design for discrete-time linear parameter-varying descriptor systems. *International Journal of Adaptive Control and Signal Processing*, 29(2):242–258, 2015.

[144] J. Wei, Y. Wu, and J. Dong. Actuator and sensor faults estimation for discrete-time descriptor linear parameter-varying systems in finite frequency domain. *International Journal of Systems Science*, 49(7):1572–1585, 2018.

[145] J. Xu, P. Shi, C.-C. Lim, and C. Cai. A descriptor-system approach for finite-frequency $\mathcal{H}_\infty$ control of singularly perturbed systems. *Information Sciences*, 370–371:79–91, 2016.

[146] Q. Xu, Z. Jing, and S. Hu. Stability analysis of nonlinear dynamic system with linear observer for a multilink flexible manipulator. *International Journal of Non-Linear Mechanics*, 103:27–36, 2018.

[147] X.-G. Yan and C. Edwards. Fault estimation for single output nonlinear systems using an adaptive sliding mode observer. *IFAC Proceedings Volumes*, 41(2):1896–1901, 2008.

[148] X.-G. Yan, S.K. Spurgeon, and C. Edwards. *Variable Structure Control of Complex Systems*. Communications and Control Engineering. Springer, Cham, 2017.

[149] X.G. Yan and C. Edwards. Nonlinear robust fault reconstruction and estimation using a sliding mode observer. *Automatica*, 43(9):1605–1614, September 2007.

[150] X.G. Yan and C. Edwards. Adaptive sliding-mode-observer-based fault reconstruction for nonlinear systems with parametric uncertainties. *IEEE Transactions on Industrial Electronics*, 55(11):4029–4036, 2008.

[151] C. Yang, Q. Kong, and Q. Zhang. Observer design for a class of nonlinear descriptor systems. *Journal of the Franklin Institute*, 59(5):1284–1297, 2013.

[152] C. Yang, Q. Zhang, J.H. Chou, and F. Yin. $\mathcal{H}_\infty$ observer design for descriptor systems with slope-restricted nonlinearities. *Asian Journal of Control*, 59(4):1133–1140, 2012.

[153] C. Yang, Q. Zhang, and L. Zhou. *Stability analysis and design for nonlinear singular systems*. Lecture Notes in Control and Information Sciences. Springer, Berlin, 2012.

[154] J. Yang, F. Zhu, X. Wang, and X. Bu. Robust sliding-mode observer-based sensor fault estimation, actuator fault detection and isolation for uncertain nonlinear systems. *International Journal of Control, Automation and Systems*, 13(5):1037–1046, 2015.

[155] X. Yao, J.H. Park, L. Wu, and L. Guo. Disturbance-observer-based composite hierarchical anti-disturbance control for singular markovian jump systems. *IEEE Transactions on Automatic Control*, 2018.

[156] T.K. Yeu and S. Kawaji. Sliding mode observer based fault detection and isolation in descriptor systems. In *Proceedings of the American Control Conference*, volume 6, pages 4543–4548, Anchorage, AK, May 2002. IEEE.

[157] T.K. Yeu, H.S. Kim, and S. Kawaji. Fault detection, isolation, and reconstruction for descriptor systems. *Asian Journal of Control*, 7(4):356–367, October 2005.

[158] G. Ying, H. Ming, and Y. Dan. Fault reconstruction for linear descriptor systems using PD observer in finite frequency domain. In *Proceedings of the 24th Chinese Control and Decision Conference (CCDC)*, pages 2281–2286, Taiyuan, 2012. IEEE.

[159] E. Yip and R. Sincovec. Solvability, controllability, and observability of continuous descriptor systems. *IEEE Transactions on Automatic Control*, 26(3):702–707, June 1981.

[160] J. Yu and Z. Liu. Fault reconstruction based on sliding mode observer for linear descriptor systems. In *Proceedings of the Asian Control Conference*, pages 1132–1137, Hong Kong, August 2009. IEEE.

[161] J. Yu, G. Sun, and H.R. Karimi. Fault-reconstruction-based cascaded sliding mode observers for descriptor linear systems. *Mathematical Problems in Engineering*, 2012:1–20, 2012. Art. no. 623426.

[162] P. Yu, Y. Shtessel, and C. Edwards. Continuous higher order sliding mode control with adaptation of air breathing hypersonic missile. *International Journal of Adaptive Control and Signal Processing*, 30(8–10):1099–1117, 2016.

[163] M. Zerrougui, L. Boutat-Baddas, and M. Darouach. Observers design for a class of continuous time nonlinear singular systems. *Journal of Control Science and Engineering*, 2011:1–8, 2011. Art. no. 967571.

[164] C. Zhang, J. He, L. Jia, C. Xu, and Y. Xiao. Virtual line-shafting control for permanent magnet synchronous motor systems using sliding-mode observer. *IET Control Theory & Applications*, 9(3):456–464, 2015.

[165] Q. Zhang, F. Liu, C. Fan, and W. Xie. Fuzzy numbers intuitionistic fuzzy descriptor systems. *Information Sciences*, 469:44–59, 2018.

[166] Y.C. Zhang, Z.M. Zhao, T. Lu, L.Q. Yuan, W. Xu, and J.G. Zhu. A comparative study of Luenberger observer, sliding mode observer and extended Kalman filter for sensorless vector control of induction motor drives. In *Proceedings of the Energy Conversion Congress and Exposition*, pages 2466–2473, San Jose, CA, September 2009. IEEE.

[167] G. Zheng and F.J. Bejarano. Observer design for linear singular time-delay systems. *Automatica*, 80:1–9, 2017.

[168] G. Zheng, D. Boutat, and H. Wang. A nonlinear Luenberger-like observer for nonlinear singular systems. *Automatica*, 86:11–17, 2017.