Robust Fault and Cyber-Attack Detection in Cyber-Physical Systems

Parisa Yazdjerdi

A Thesis in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the

Requirements for the Degree of Master of Science

in Electrical and Computer Engineering

at Concordia University

Montreal, Quebec, Canada

February 2023

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis

Prepared By: Parisa Yazdjerdi

Entitled: Robust Fault and Cyber-Attack Detection in Cyber-Physical Systems

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Electrical and Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Prof. Rastko Selmic

_____ Examiner
Dr. Walter Lucia

_____ Examiner
Prof. Rastko Selmic

_____ Supervisor
Prof. Khashayar Khorasani

Approved by: _____
Prof. M. Zahangir Kabir, Graduate Program Director

_____ _____
Prof. Murad Dababi, Dean, Gina Cody School of Engineering and Computer Science

ABSTRACT

Robust Fault and Cyber-Attack Detection in Cyber-Physical Systems

Parisa Yazdjerdi, M.Sc.

Concordia Univeristy, 2023

There is a growing interest towards Cyber-Physical Systems (CPSs) due to their wide range of applications such as power systems, smart grids, aerospace, transportation, and process control systems in recent years. CPSs are more reliable than conventional systems and show higher level of performance in complex environments which make them more functional in different applications. CPS is prone to anomalies such as machine induced faults and cyber-attacks, which can destabilize the system and cause significant degradation in the system's performance or result in system failure. Hence, the study of cybersecurity challenges such as detection of anomalies and designing a resilient controller in the presence of anomalies have recently attracted significant attention. In addition, diagnosis of simultaneous cyber-attack and the physical fault is an important challenge which requires more attention in the literature. The main aim of this thesis is to design a fault and cyber-attack detection mechanism in the presence of a disturbance in CPSs and Multi-Agent Systems (MASs).

In the first chapter of this thesis, the design of fault and cyber-attack detection scheme is investigated for linear cyber-physical systems. Two different types of filters have been designed to detect concurrent fault and attacks in a cyber-physical system. It is assumed that the agent is equipped with a local controller which receives a reference signal from the command and control unit and computes the controller signal internally. Hence, the control signal is generated locally and it is not prone to cyber-attack. On the other hand, the measurement signal is

communicated over wireless network from the system to the command and control unit which is prone to cyber-attacks. Accordingly, the sensor measurement is prone to both cyber-attack and fault. It should be noted that it is assumed that there exist two communication channel for sending and receiving. In particular, the reference signal is transmitted from command and control station to the system in one channel and the other channel is used to transmit the output signal from the system to the command and control station and only the later is prone to attack. The filters are designed based on a multi-objective framework by utilizing the $\mathscr{H}_\infty$ and $\mathscr{H}_-$ formulation in a finite frequency domain. The efficiency of the proposed method in simultaneous detection of fault and cyber-attack is verified by simulating a VTOL aircraft in presence of disturbances and measurement noise. The proposed method can detect fault and cyber-attacks within a specific range of magnitudes which are within the performance of VTOL system.

The second chapter investigates the problem of the concurrent fault and cyber-attack detection in a MAS in the presence of disturbances. The MAS under healthy condition is equipped with a consensus controller based on the output information of neighbors. Hence, only the output measurement is communicated among the agents by a wireless network. The attack is assumed to be injected as a false data on the communication link between agents. To achieve this goal, a fault detection Luenberger observer is designed based on multi-objective framework by utilizing $\mathscr{H}_\infty$ and $\mathscr{H}_-$ formulation. Despite the previous chapter results, the design procedure has been done in full frequency domain rather than the finite frequency domain which results in less computational complexity but it can be used for strictly proper system (Fault and cyber-attack on the output signal). The fault detection observer is designed locally for each agent and can detect faults in the sensor and actuator. On the other hand, an Unknown Input Observer (UIO) is designed for each agent in its neighbors to detect cyber-attack which is injected as a false data on the link between agents. In particular, each agent is equipped

with a local fault detection Luenberger-based observer and a bank of UIO which detect the cyber-attack in the neighboring agents. The efficacy of the proposed method is validated by simulation of a group of UAVs model forming a multi-agent system. The main aim of attack is not to destabilize the system but to fool the monitoring station. Thus, the cyber-attack is assumed to have the same effect as fault. Accordingly, the magnitude of cyber-attack cannot be beyond the performance limitation of the UAV. The proposed method can detect cyber-attack and fault within the performance limitation of a UAV.

To my husband

Omid.

ACKNOWLEDGMENT

I would like to express my very sincere gratitude to my thesis advisors Prof. Khashayar Khorasani and Prof. Nader Meskin for their patience, motivation, enthusiasm, and immense knowledge. They are the best instructors I could have throughout my study at Concordia University. The door to Prof. Khorasani's office was always open and Prof. Meskin provides me a full support from Qatar whenever I had a trouble or a question about my research or writing my thesis. They consistently allowed this thesis to be my own work, but steered me in the right direction whenever they thought I needed it.

I would like to express my very profound gratitude to my husband Omid for his unfailing support, encouragement, and patience during my years of study and writing this thesis.

My Sincere thanks also go to my parents Roya and Mehrdad for their prayers and support during my study.

TABLE OF CONTENTS

List of Tables

List of Figures

LIST OF ACRONYMS

CPS Cyber-Physical System

MAS Multi-Agent System

UIO Unknown Input Observer

DoS Denial of Service

ZDA Zero Dynamic Attack

LMI Linear Matrix Inequality

LTI Linear Time-invariant

LPV Linear Parameter Varying

LF Low Frequency

MF Medium Frequency

HF High Frequency

IMP Internal Model Principle

VTOL Vetrical Take-Off and Landing

GKYP Generalized Kalman-Yakubovich-Popov

UAV Unmanned Aerial Vehicle

Chapter 1

INTRODUCTION

1.1 Cyber-Physical Systems

Cyber-Physical Systems (CPSs) consist of network of sensors, actuators, and embedded computers which are connected through a communication network with decision making abilities [1; 2; 3; 4]. In particular, CPSs consist of three main parts including communication layer, computational capabilities and physical components in which there exists an interaction between physical components and computational layer over wireless communication. There is a growing interest towards security of CPSs due to their wide range of applications such as power systems, smart grid, aerospace and transportation, and process control systems in recent years [5; 6; 7; 8]. CPSs are more reliable and show a higher level of performance in complex environment which makes them more functional in different applications. Figure 1.1 delineates a cyber-physical system in the presence of fault and cyber-attack on both sensor and actuator. Although cyber-physical systems add significant benefits and values to the industrial application, new challenges in ensuring the security of the cyber part of CPSs open a broad window of research named as the cybersecurity of CPSs. According to the literature, the problem of

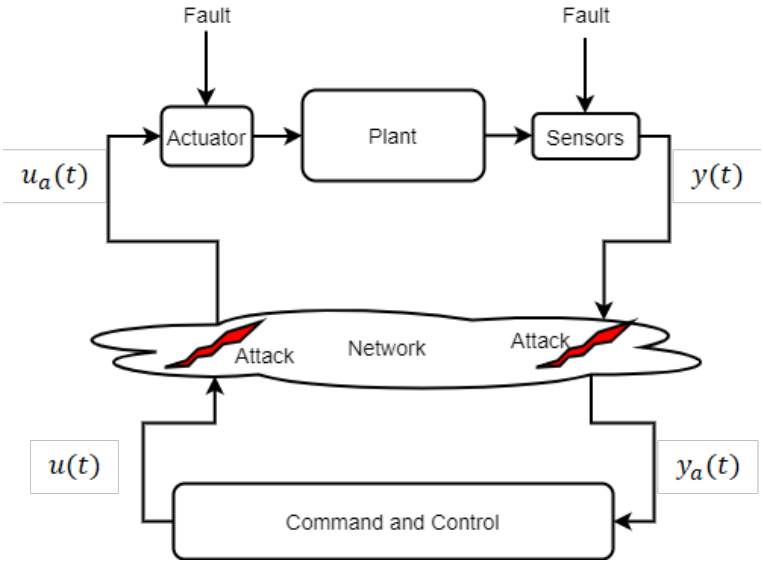


Figure 1.1: CPS in the presence of fault and cyber-attack.

cybersecurity has been studied from a computer science and information technology point of view for computer-based systems, primarily. However, it does not address the cybersecurity of several CPSs. Thus, the study of cybersecurity from the control point of view has been defined to overcome this shortage. In general, the study of cybersecurity has been covered from two different perspectives including the security of control systems from the control engineering point of view and the security of data which is mostly done by researchers in the field of computer science using encryption/ decryption methods [9; 10; 11; 12; 13; 14].

1.1.1 Fault Detection and Isolation

A wide range of studies have been done on fault detection, isolation, and diagnosis of cyber-physical systems. A fault detection and isolation approach is proposed in [15] based on state estimation technique and is validated using a water treatment system. A detection and estimation scheme is proposed in [16] for both bias and loss of effectiveness faults on actuators using adaptive threshold and sliding mode observer approach. The proposed method in this thesis is able to estimate the bias fault and loss of effectiveness of the actuator in the presence of

disturbances. A distributed fault detection method is proposed in [17] and is able to detect and isolate multiple sensor faults in the physical part of the interconnected cyber-physical system. A necessary and sufficient condition for the existence of a full order unknown input observer is provided in [18] and a robust fault detection filter for nonlinear jet engine is developed.

A fault diagnosis, detection, and isolation approach is developed in [19] based on an interactive multiple model algorithm. A real-time fault detection and isolation scheme is proposed in [20] for single and concurrent faults in the jet engines. Similarly, a multiple model based algorithm is proposed in [21] to identify, detect, and isolate sensor faults. A single dynamic observer is designed in [22] to achieve the simultaneous detection, isolation, and tracking problem. The observer is designed with the aid of $\mathscr{H}_\infty / \mathscr{H}_- / \mathscr{L}_1$ formulation. The generated residual is sensitive to fault and attenuate the effects of control signals and disturbances. The tracking goal is also achieved because the effect of disturbance and fault are minimized on the generated control signal.

A Luenberger observer is designed in the finite frequency domain for a LTI continuous-time system in the presence of fault and disturbance in [23; 24]. The problem of finite frequency fault estimation for discrete-time linear parameter varying system is investigated in the presence of an external disturbance in [25]. The sensor and actuator faults are estimated simultaneously using an auxiliary system. A fault detection observer is designed in [26] for a continuous-time networked control system which is modeled as Markov Jump systems in the presence of sensor fault and disturbance. Estimator and controller are designed simultaneously for a continuous time-invariant linear system under sensor fault and disturbance in [27]. A dynamic observer and a state feedback controller are designed by converting four finite frequency range $\mathscr{H}_\infty$ performance indices to Linear Matrix Inequalities (LMIs) using dual Generalized Kalman-Yakubovich-Popov (GKYP) lemma. A fault detection observer in the finite frequency range is

designed for Lipschitz non-linear systems in the presence of actuator faults and disturbances in [28]. The proposed method is validated by numerical simulations of a single-link flexible joint robot. A state estimator filter is designed for a T-S fuzzy system in the presence of unknown input in the finite frequency domain by utilizing $\mathscr{H}_\infty$ performance indices [29].

1.1.2 Cyber-Attack Detection and Isolation

On the other hand, the cyber part of CPSs is threatened by different types of cyber-attacks. Cyber-Attacks are mainly injected on the communication links and have different types such as Denial of Service (DoS) attack and false data injection attack [13]. Various methods have been studied in the literature to detect and isolate attacks in CPSs [30; 31; 32]. Also, cyber-attack resilient controllers have been developed in the literature to reduce the effect of the injected cyber-attack on the performance of the system [33; 34].

Undetectable and unidentifiable cyber-attacks are defined for cyber-physical system in the presence of cyber-attack and cyber-attack detection mechanism from the graph-theoretic perspective in [31]. Also, centralized and distributed filters are designed for cyber-attack detection and identification of undetectable and unidentifiable cyber-attacks. A comprehensive survey on control and cyber-attack detection of industrial systems is provided in [35] in which different types of cyber-attack on industrial CPSs are summarized. Also, various detection approaches are classified in this survey paper. A machine learning approach is developed in [32] to detect false data injection attack on the communication link of a vehicular cyber-physical system. The analysis on finite frequency cyber-attack detection approaches is provided in [36]. A comprehensive survey on security of CPSs is provided in [37] from the automatic control point of view and a broad range of cyber-attack models and defense strategies are presented for CPSs with

different applications. A brief survey on the detection of false data injection attack on sensors and actuators in CPSs has been presented in [38] for linear time-invariant and nonlinear systems.

Despite the existing approaches on designing robust mechanisms which are mainly in the full frequency range, it is worth considering that in a practical situation the faults and disturbances occur in a finite frequency range. Also, an attacker has the ability to choose a specific range of frequency, duration, and amplitude for a cyber-attack to remain stealthy. Thus, it is more efficient and less conservative to consider finite frequency analysis. It is shown in [39] that cyber-attack detection in finite frequency range is more efficient than in full frequency. Also, it should be noted that from the mathematical point of view, $\mathscr{H}_{-}$ analysis for fault/cyber-attack sensitivity analysis in the full frequency range is not possible for strictly proper systems [23]. Hence, it is important to consider fault/cyber-attack detection, isolation, and diagnosis analysis in the finite frequency range. A brief literature review on finite frequency analysis has been provided as follows.

The sensor and actuator cyber-attack detection and estimation scheme for cyber-physical systems with discrete time-invariant linear dynamic is studied in [40]. Characterization of undetectable and indistinguishable cyber-attacks are defined for CPSs and sufficient conditions related to detectable and distinguishable cyber-attacks are obtained where both sensor and actuator attacks are considered. A Luenberger observer is designed to detect low frequency cyber-attack using $\mathscr{H}_{\infty}$ analysis. Attack detection filter is designed in [36] to detect sensor attacks in different frequency ranges. According to the existing literature, it is clear that the finite frequency analysis is mainly done for fault diagnosis approaches. There exist very few studies for cyber-attack detection, isolation, and estimation in this domain [40; 36]. Also, the detection and isolation of fault and cyber-attacks in finite frequency ranges is not studied in the

literature.

It is worth considering that very limited attention to detection and isolation of concurrent fault and attack on the cyber-physical systems has been studied. A data-driven approach is proposed in [41] to distinguish cyber-attacks from faults in a smart grid system. A machine learning-based method is developed in [42] to distinguish attacks from faults. A fault and attack detection and isolation scheme is provided in [38] by placing one filter at the command and control and one on the plant side which increases the communication load. A simultaneous stochastic fault and attack detection and isolation method is developed in [43] in which a fault detection filter is designed to detect and isolate stochastic faults and a mixed coding and message authentication approach is developed to detect and isolate cyber-attacks. Table 1.1 summarizes the existing literature on designing observers in both finite and full frequency for different types of systems. According to this table, the existence of a fault, attack, and disturbance is omitted from the literature which is solved in this thesis. Also, it can be observed that the importance of detection and isolation of fault and attack in MASs requires more attention.

1.2 Multi-Agent Systems

Multi-Agent systems are mainly composed of two or more intelligent interacting agents aimed to solve problems which are not easily solvable by an individual agent. A wide range of research has been done on MASs due to the high demand in a variety of applications such as smart grids, smart manufacturing, and intelligent transportation systems. Control of MAS are mainly done by the information sharing of the neighbors through wireless communication channel which are prone to cyber-attacks. Cyber-Attack on the communication links between agents can cause instability, performance degradation, and destruction depending on the aim of the attacker. Hence, safety and security of the multi-agent systems is an important issue to be

Table 1.1: *Literature Review on Finite Frequency Analysis*

Ref	Freq. domain	Anomaly	Sys. Type	Objective	Single /MAS
[23]	MF	Fault & Dist	Cont. LTI	Designing Luenberger observer for fault detection	Single
[24]	MF	Fault	Cont. LTI	Designing Luenberger observer for fault detection	Single
[25]	All	Fault & Dist.	Cont. LPV	Fault estimation	Single
[40]	LF	Attack.	Disc.	Attack detection and estimation(Luenberger)	Single
[26]	LF	Fault & Dist.	Cont.	Fault detection	Single
[44]	LF	Fault & Dist.	Cont. LPV	Fault detection and isolation	MAS
[28]	LF	Fault & Dist.	Cont. Lipstchiz nonlinear	Fault detection	Single
[36]	All	Attack & Dist.	Cont.	Attack detection	Single
[27]	LF,HF	Fault & Dist.	Cont.	Observer-based controller design	Single
[29]	MF	Unknown Input	Cont. T-S fuzzy	State estimator observer	Single

studied.

One of the primarily research topics in the context of multi-agent systems is designing the controller to achieve a desired task such as reaching a formation or consensus. Extensive studies can be found in the literature on different consensus protocols, communication constraints and mechanisms of MASs [45; 46; 47].

1.2.1 Fault Detection and Isolation

Similar to CPSs, the physical safety of MASs is among the most important research topics in this domain. Most recent strategies in designing a fault-tolerant controller in MASs are provided in [48]. Several robust control methods are developed in the literature to detect the fault and achieve control objectives for multi-agent systems [49; 50]. A fault detection and isolation filter is designed in [51] for a heterogeneous linear multi-agent system. It is observed that the

physical safety of MASs has been improved due to a wide range of studies on fault detection, estimation, diagnosis, and fault-tolerant control. An unknown input observer is designed in [44] to detect and isolate finite frequency range fault on the sensor and actuator signal for the class of linear heterogeneous parameter varying multi-agent systems.

An observer-based distributed fault detection and isolation based on UIO is proposed in [52] for a second-order MASs. The fault is modeled on the state of the system and the relative information between the neighbors is considered to design the UIO. The faulty agent is then isolated from the network. A distributed fault detection scheme for interconnected second-order systems is proposed in [53] using a bank of UIOs. The interaction among agents is considered by a distributed control law. The fault is considered on the system dynamics and the feasibility of the proposed scheme is studied with respect to local measurements. The fault agent is removed from the network in the mitigation procedure. The problem of fault detection and isolation in a large interconnected system in the presence of model uncertainties is addressed in [54]. In [55], the problem of simultaneous sensor and actuator fault estimation is addressed and an output-feedback-based fault-tolerant controller is designed for a Markovian jump system. Moreover, several robust control methods are developed in the literature to detect the fault and achieve control objectives for multi agent systems [49; 50].

An event-triggered distributed scheme is proposed in [56] for simultaneous fault detection and tracking control for MASs. A module is designed for each agent to generate a residual signal for fault detection and a control input to achieve the tracking goal. The data transmission among agents is considered to be event-triggered by utilizing a dynamic triggering rule which results in less amount of data transmission. A robust fault detection scheme is proposed in [57] for higher-order MASs in the presence of disturbance by using UIO. A partitioning method is proposed if the UIO decoupling condition is not satisfied. The problem of fault detection

multi-agent systems has been addressed in [58]. A bank of reduced order UIOs is designed in each agent to detect the fault in the neighboring agents. For the cases where the matching condition is not satisfied, the UIO is partitioned into two parts such that the effect of one part on the residual signal can be removed and the effect of the second part can be attenuated on the residual signal by H_∞ performance index

1.2.2 Cyber-Attack Detection and Isolation

The existing approaches on detection and isolation of cyber-attacks in the multi-agent framework is as follows. A cyber-attack detection method based on stability analysis and loss of effective watermarking signal is designed in [59] for multi-agent systems under a replay attack on the sensor measurement. It is shown that the replay attack is not detectable by any of the agents except when the watermarking control strategy is applied. Optimization methods are used to design a watermarking signal with minimum loss of effectiveness on the performance of the system. It is shown that watermarking signal is better to be communicated within agents for the best performance however, it requires more communicational bandwidth which is one of the main limitations in the field of MASs.

An observer-based approach is proposed in [60] for cyber-attack detection in a network of UAVs in a formation flying setup. The attack is injected on each agent and on the communication link between the agents. The compromised agent is identified in the network of UAVs based on the generated residual signal. A cyber-attack detection method based on an ellipsoidal set membership filtering approach is proposed in [61] for a leader-follower-based multi-agent system in the presence of bounded noise and quantization effect. Denial of Service (DoS) and False Data Injection attack are injected on the sensor measurement of each agent. It is assumed that the leader is cyber-attack free and based on the known bounds of noise and quantization

effect, prediction and estimation sets are calculated. The lack of intersection between the two sets results in cyber-attack detection in the system. It is worth considering that the assumptions on noise bound and cyber-attack free leader are not realistic in engineering disciplines.

An anomaly detection and identification approach is described in [62] for multi-agent systems subject to physical fault and false data injection attack . The cyber-attack is injected on the communication link between agents. An independent detector is designed for each agent and a bank of cooperative detector is designed for each agent to detect and identify anomalies in the presence of disturbances. Moreover, the cooperative detector is enhanced using the betweenness centrality of edges which provides the quantitative measure of an edge to be under cyber-attack, i.e. probability of an edge to be attacked. This approach enhanced the performance of detection and identification which is shown in the experimental results of this paper. A simultaneous stochastic fault and cyber-attack detection and isolation method is developed in [43] in which a fault detection filter is designed to detect and isolate stochastic faults and a mixed coding and message authentication approach is developed to detect and isolate cyber-attacks.

A distributed covert attack detection mechanism is developed in [63] for an interconnected MASs. In the proposed method, each subsystem is equipped with two observers that estimate the states using different information. The first observer defined as a distributed observer is designed based on the local model of its corresponding subsystem, local sensor information, and the communicated data from neighbors. The second observer is an unknown input observer that uses local information and measurements to estimate the states and it is called decentralized. The attack is then detected by comparing the estimated states of the two proposed observers. An intrusion detection technique is developed in [64] to detect cyber-attacks in a linear multi-agent system using an unknown input observer. In this paper, the problem of consensus is solved in

the presence of a node attack. A filter is designed for each agent to estimate the states of its neighbors. The attack is detected by comparing the estimated states with the received information from the corresponding node. It is worth mentioning that this method can only detect and isolate an attack when it is applied to a single node. A method is proposed in [65] to detect and isolate compromised agents in a linear MAS. The sufficient condition is provided based on the communication graph topology of MAS. A sufficiently connected graph is required to detect the misbehaving node. Also, the conditions on designing a stealthy attack are provided for not sufficiently connected systems.

In order to detect an attack in a MAS, an algorithm based on a support vector machine is developed in [66]. Each agent in this framework is equipped with data processing and decision making approaches to detect attacks by peer-to-peer communication with other agents. A replay attack detection technique is proposed in [67] for a group of vehicles as MAS with cooperative cruise control. The decentralized detection mechanism is developed using noisy control signal methodology and cross-correlator. A noisy authentication signal injected by the leader vehicle and an auxiliary model of the MASs is described which is used to generate a signal which can be compared by the under-attack vehicle data. The cross-correlation method is used to generate a residual signal that identifies the presence of an attack. An approach is proposed in [68] to detect DoS attacks in MASs presented as connected vehicles. The DoS attack is injected by occupying the communication channel with the fake request from the attacker and does not let the valid request from the user to be transmitted. The effect of DoS attack is modeled as a time delay in the mathematical model of the system. The proposed method consists of observers design by sliding mode theory and adaptive observer approaches.

There exist an extensive study of cyber-attack detection and isolation of micro-grids. Mostly, the authors refer to micro-girds as a special type of multi-agent system. The authors in [69]

designed a distributed cyber-attack detection in micro-girds. A false data injection attack is considered on the communication channel between distributed generation units. The cyber-attack detection is done using an UIO to estimate the state of neighbors locally. A detection and isolation scheme is proposed in [70] for false data injection attack detection and isolation in smart grids using unknown input observers. An adaptive threshold is designed to improve the detection performance against disturbance. A distributed cyber-attack detection scheme is proposed in [71] for a class of linear time-invariant systems with the application to micro-girds. The proposed detection scheme consists of a Luenberger observer and a bank of UIOs. The detectability analysis has been done and it is delineated that some classes of attack can be detected by exploiting both detection modules.

1.2.3 Security Control

Several approaches have been investigated in the literature to design secure controller in the presence of various types of attacks in multi-agent systems. Designing a resilient controller is investigated in the literature by many researchers as follows. A cyber-attack resilient controller is proposed in [72] for both discrete and continuous-time MAS with a leader-follower scheme under strategic cyber-attack on the communication graph. The cyber-attack is modeled as a random Markov process. If necessary conditions on cyber-attack frequency and cyber-attack duration are met, the mean square exponential consensus tracking for MAS is achieved using a stochastic secure control strategy. The stability analysis is done using Lyapunov's method. Nevertheless, it is assumed that the mathematical description is noise and disturbance free which is not realistic. A resilient control framework is described in [73] for distributed multi-agent systems under Internal Model Principle (IMP) and non IMP-based cyber-attacks using graph theoretic approach. It is worth considering that IMP-based cyber-attack depends on

the knowledge of system dynamics. The cyber-attack is injected into sensor measurement or control input and modeled as an additive term to the actuator or measurement signal. The self-belief and trustworthiness of the information received by each agent is calculated using a Kullback-Libeler (KL) divergence-based criterion in which each agent detects its neighbors' misbehavior. The self-belief is continuously updated and communicated with other neighbors to inform them about the correctness of information which results in cyber-attack detection. Moreover, each agent forms a belief based on the neighbor's belief in the case that its self-belief value is low. In addition, the self-belief values which are communicated between each agent are used in the control protocol to mitigate cyber-attacks.

A resilient controller based on leader-follower graph is designed in [74] for a linear MAS that consists of a leader, group of followers, and group of malicious followers which may or may not be under attack. A criterion is proposed based on the known maximum number of under-attack agents such that each healthy follower can detect and filter out the anomaly and uses the healthy information from neighboring agents to update its own state. In particular, the control strategy is designed to ensure that the state of healthy followers converges to the leaders state if the communication graph is leader-follower graph. A resilient controller is designed for multi-agent systems modeled as a single integrator in the presence of DoS attack on the communication link between agents in [75]. The author considers a formation control problem for a group of agents based on their own and single neighbor's information. Each agent is able to correct its performance according to its own and one of its neighbors information.

The design of an event-triggered resilient controller is also studied in the literature. A resilient event-triggered controller is designed in [76] for a multi-agent system under intermittently random DoS (IRDoS) cyber-attack to reach mean square consensus. The cyber-attack is injected into the communication link between agents. The proposed distributed controller is re-

silient to a specific IRDoS attack with a specific duration and success probability rate. It should be considered that most of the case studies in recent publications are very simple numerical examples and is not applicable on real engineering systems. A resilient autonomous controller is designed for a leader-follower MAS in the presence of the attack on sensors and actuators in [77]. In order to avoid propagation of attack on other components and to attenuate the effect of the injected attack on the agent, a distributed observer-based $\mathscr{H}_\infty$ controller is designed. The trustworthiness of each agent is derived using a reinforcement learning algorithm that results in the confidence value of that agent. The low confidence value of one agent indicate that the data received from that agent is compromised and should be neglected by the neighbors.

Designing a secure consensus controller is attracting the considerable interest of researchers according to the literature. A secure consensus controller is designed for linear MASs in the presence of a DoS attack in [78]. The conditions on the frequency and duration of the DoS attack are derived using decay rate for different attack modes to ensure that the proposed state feedback controller and observer-based controller guarantee the consensus in the presence of DoS attack. The mean square consensus of a multi-agent system is studied under a false data injection attack on the communication link between agents in [79]. An estimator is designed for each agent to estimate its neighbor's states. By comparing the estimated and received state from other agents, the cyber-attack is detected. To obtain the security of the system in the presence of cyber-attack, the estimated value of the state is used for the next iteration. A secure mean square consensus distributed control is designed for a leader-follower based multi-agent systems under random cyber-attack in [80]. An observer is designed for each agent to estimate the state of the system. It is shown that under sufficient conditions on cyber-attack frequency and duration the consensus is achieved and stability analysis is done using Lyapunov theory. The cyber-attack is injected internally between observer and controller.

The problem of the mean square bounded synchronization of multi-agent systems with a leader-follower scheme under deception attack on the controller to actuator signal is investigated in [81]. The system model consists of an impulsive input which is designed using the pinning strategy to secure the network in the presence of a cyber-attack. Sufficient conditions are provided on cyber-attack success probability, coupling strength, impulsive intervals, and pinning matrix to guarantee secure mean square bounded synchronization of the system. A similar strategy is provided in [82] in the presence of false data injection attack on the sensor to controller channel. The main shortcoming in this article is that the effect of noise and disturbance is not considered. Also, despite the existing nonlinearities in the system description, cyber-attacks are treated as in linear time-invariant systems. A secure consensus controller is developed in [83] for a class of linear leader-follower MAS in the presence of periodic DoS attacks. The sufficient condition to achieve secure consensus is derived with the aid of the Lyapunov function theory.

It is worth considering that many authors in the literature investigate designing an event-triggering consensus mechanism mainly in the presence of DoS attacks. A distribute event-triggered consensus-based controller is designed for linear MASs in the presence of a DoS attack in [84]. In this paper, sufficient conditions on frequency and range of attack are derived using a switching framework to ensure secure consensus in the presence of an attack. The stability of the proposed controller is verified using the Lyapunov function and linear matrix inequality. The problem of designing a distributed secure consensus control for a stochastic linear MAS is investigated in [85] in the presence of white Gaussian noise and two types of attack. Two types of attacks are defined in this paper as connectivity-maintained and connectivity-broken attacks targeting the graph topology. This problem is formulated as a switching system and sufficient conditions are derived to ensure the secure consensus. The problem of secure

consensus of linear MASs in the presence of periodic DoS attack has been investigated in [86]. The MAS in the presence of periodic DoS attacks is formulated as a switched system with time delays. Based on the consensus error dynamic sufficient conditions are derived to ensure exponential stability of the switched system with time delay. Also, an optimal sleep-time for the event-triggering mechanism is calculated.

The author in [87] designed an event-triggering secure consensus controller for a MAS in the presence of an aperiodic DoS attack. The triggering condition is derived based on the control input signal rather than state measurements. Then, sufficient conditions on frequency and duration of attack are discussed to assure secure consensus with the aid of the event-triggering topology. An event-triggering mechanism is provided in [88] to assure secure consensus of linear multi-agent systems in the presence of a DoS attack. Based on the discussion on the duration and frequency of the DoS attack, event-triggering conditions are derived and it has been shown that the system reaches consensus exponentially in the presence of a DoS attack for both leaderless and leader-follower topology. An event-triggering consensus controller is designed for a multi-agent system with a stochastic discrete-time model under deception cyber-attack and lossy sensors in [89]. The cyber-attack is injected into the sensor of the agent. An observer is designed to estimate states to be used in the feedback law to achieve the consensus. The deception cyber-attack is injected internally on the sensor measurement with the same representative as a fault.

Design of intelligent cyber-attack is among the most crucial research field to be considered in the literature because it is highly important for a defender to anticipate the novel cyber-attacks and observe the damages on the system in order to design a better defense mechanism. A novel cyber-attack is designed in [90] and is injected on the communication link. In contrast to other papers in the literature, the controllability of the system is studied from an adversary's

point of view. A linear time-invariant MAS with an output feedback controller is considered under the directed graph. First, the conditions under which a subset of agents are controllable are studied. Also, the conditions under which the rest of the agents are controllable by an adversary as followers of the first set are analyzed. However, it is assumed that the Laplacian matrix should be diagonalizable which is a restrictive assumption. It is shown that the adversary is not able to inject zero dynamic attacks simultaneously with other cyber-attacks [91].

According to the existing studies, some potential research direction for multi-agent systems can be counted as follows:

- Design of cyber-attack detection and secure consensus control in the presence of different types of cyber-attacks.
- Simultaneous detection and isolation of cyber-attack from fault requires more attention in the literature.
- Reducing the detection time of cyber-attacks is important to be studied (Finite-time detection).

Moreover, it is observed that an extensive study is done on MASs with linear dynamic models. Different approaches are used to develop detection, mitigation, and resilient controller in the presence of an cyber-attack. Besides cyber-attacks, it should be considered that CPSs and multi-agent systems are prone to faults. Thus, it is very important to consider the presence of both fault and cyber-attack in the analysis. To the best of my knowledge, the problem of designing fault and cyber-attack detection and classification mechanism has limited focus in the literature despite its importance. Also, in the practical situation, it is crucial to consider effect of noise and disturbance in analysis. In addition, mitigation of fault and cyber-attack on both linear and nonlinear systems is considered as a potential research field.

1.3 Thesis Contribution

The main contributions of this thesis are as follows:

- Simultaneous fault and cyber-attack detection in cyber-physical systems in the finite frequency domain. To the best of our knowledge, this is the first time in the literature to study this problem. There exists a wide range of studies on fault detection and isolation on CPSs in finite frequency and few papers on cyber-attack detection and isolation. However, as discussed earlier in this report, simultaneous fault and cyber-attack in CPSs can occur and it is one of the challenges to detect these anomalies.

- The concurrent attack and fault can be detected simultaneously while in [42] the attack can be detected in fault-free case and fault can be detected in attack-free case only. This is a limiting assumption and simply ignores the possibility of concurrent occurrence of anomalies.

- The effect of disturbances have not been considered in the design procedure of attack/fault detection mechanisms in both full and finite frequency analysis [15; 31; 24; 40]. It is worth mentioning that some authors considered the effects of disturbance in fault or cyber-attack detection schemes. However, since there is limited research on concurrent or simultaneous anomalies, the robustness of the proposed filters are improved with respect to the literature.

- Designing a simultaneous fault and cyber-attack detection for MASs has not been studied in the literature. The literature mainly focuses on either fault detection and isolation [52; 53] or cyber-attack detection [69; 60]. Hence, considering both aforementioned anomalies in a system is considered as a contribution.

- The monitoring system is designed locally in the proposed framework and it only requires the output information of the neighboring agent. It is more efficient in terms of data transmission among the agents. Also, in the cyber-attack detection methods proposed in the literature, the input to the neighboring agent is either transmitted or assumed to be known [62] while in this approach this assumption is relaxed.

1.4 Thesis Organization

The remainder of this thesis is organized as follows. A review of CPS and MAS in the presence of fault and attack is provided in Section 2. In addition, different types of attack have been studied in Section 2.1. Finite Frequency analysis such as the definition of GKYP lemma is provided in Section 2.2. Section 2.3 describes few lemmas which are used in mathematical derivations of next chapters. Moreover, graph theory and consensus protocols are described in Sections 2.4 and 2.5, respectively.

In Chapter 3, Robust fault and cyber-attack detection mechanism is designed for cyber-physical systems. Section 3.1 provides a system description of a continuous-time linear cyber-physical system in the presence of fault and cyber-attack. The problem formulation for the design of two different types of filters has been proposed in Section 3.2. Sections 3.3 and 3.4 provide the detailed derivation of the proposed Luenberger-based observer and Normal filter, respectively. The simulation results to verify the effectiveness of the proposed methods are presented in Section 3.6.

A robust fault and cyber-attack detection and classification scheme is proposed in Chapter 4 for multi-agent systems. The state space model of a multi-agent system in the presence of the fault, disturbance, and cyber-attack is presented in Section 4.1. Section 4.2 describes the problem formulation for the design of a Luenberger-based fault detection and UIO-based attack

detection filter. The detailed derivation of the proposed observers is described in Section 4.3. Finally, the efficacy of the proposed method is verified by simulation on a state space lateral model of a UAV in Section 4.5.

Finally, Chapter 5 provides a future direction and general conclusion of the methods proposed in this thesis.

Chapter 2

BACKGROUND INFORMATION

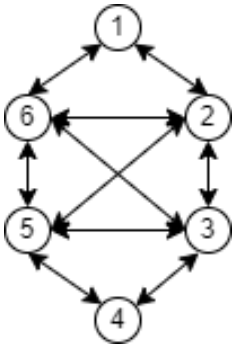
2.1 System Description

The linear continuous time-invariant state space model of a cyber-physical system in the presence of fault, attack, and disturbance is as follows:

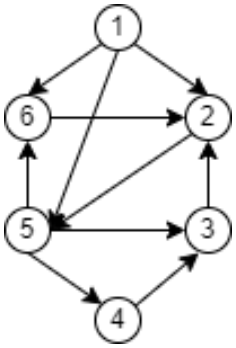
$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + B_f f_a(t) + B_a a_a(t) + B_d d(t) \\ y(t) &= Cx(t) + D_f f_s(t) + D_a a_s(t) + D_d d(t) + w, \end{aligned} \tag{2.1}$$

where $x(t)$, $u(t)$, and $y(t)$, denote state, input, and output of the system. Signals $f_a(t)$ and $f_s(t)$ present fault on actuators and sensors, respectively. Attack signals on actuators and sensors are represented as $a_a(t)$ and $a_s(t)$, respectively. $d(t)$ presents disturbance signal, and measurement noise is presented as w. The matrices A, B, B_f, B_d, B_a, C, D_f, D_a, and D_d are known with appropriate dimension. As mentioned earlier, MASs are categorized as a specific type of CPSs. It is observed in the literature that similar approaches have been utilized by authors to design defense mechanisms in both fields of MASs and CPSs.

The linear homogeneous continuous-time state space representation of a MASs with N agents



(a) Undirected graph.



(b) Directed graph.

Figure 2.1: Communication topology for MASs.

is represented as follows:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + Bu_i(t) + B_f f_{ai}(t) + B_a a_{ai}(t) + B_d d_i(t) \\ y_i(t) &= Cx_i(t) + D_f f_{si}(t) + D_a a_{si}(t) + D_d d_i(t) + w_i \end{aligned} \tag{2.2}$$

where $i = 1,2,...N$, state of the agent i is represented as x_i , u_i is the input to agent i and the output of agent i is shown as y_i. Matrices A, B, B_a, B_d, B_f, C, D, D_a, D_d, and D_f are identical and known for all agents. w_i is a white random Gaussian noise on sensor measurement i.
The communication topology in MASs can be categorized into directed or undirected depending on the application and aim of the system. Undirected graph refers to bidirectional vertex between agents as shown in Figure 2.1a and directed graph which is delineated in Figure 2.1b refers to directed edges between agents. It is worth mentioning that in the literature both the directed and undirected network topology are considered.

Different types of cyber-attacks are considered on the multi-agent system presented in (2.2) namely, False data injection, Denial of Service (DoS), deception, and replay attack. The cyber-attacks are mainly modeled on sensor measurement, actuator signal, and communication links between agents. Different types of existing cyber-attack are well described in [13] which provides a comprehensive review on cybersecurity of networked control systems and cyber-

physical systems. Different types of cyber-attacks and their characteristics, existing security mechanisms, and their objectives are well studied in this reference and the motivation to design cyber-attack detection mechanisms and secure controllers are highlighted.

The most well-known cyber-attacks will be discussed in this section namely False-Data Injection, DoS, Replay, Zero dynamics, and Covert cyber-attack. Figure 2.2 shows the positioning of these five cyber-attacks in cyber-attack space. The attack space consists of three dimensions as follows: **System knowledge** can be used by an adversary to design stealthy cyber-attacks which are harder to detect and have severe impacts on the system. By **Disclosure resources**, the adversary can obtain sensitive information about system operation during the cyber-attack. It is worth considering that resource disclosure cannot affect the performance of the system (i.e. eavesdropping). Finally, an adversary can degrade the performance of the system by **Resource disruption** by violating the integrity or availability of data (i.e. DoS attack). It is obvious from Figure 2.2 that a covert cyber-attack requires full knowledge of the system while a DoS attack can be injected without the knowledge of the system. In addition, replay attacks can be injected by eavesdropping and cause resource disruption regardless of system knowledge. Moreover, the five aforementioned cyber-attacks are positioned in the impact space as shown in Figure 2.3. This Figure shows that zero dynamics attack interact with the dynamic of the system itself while a DoS attack compromise the actuator or sensor signals. False data injection is mainly injected on the sensor measurements while a replay attack can compromise both sensor and actuator.

2.1.1 DoS Attack

Denial of service attack is defined as network jamming when the communication packets transmitting sensor or actuator signal stops. In the absence of data, the last received data will be

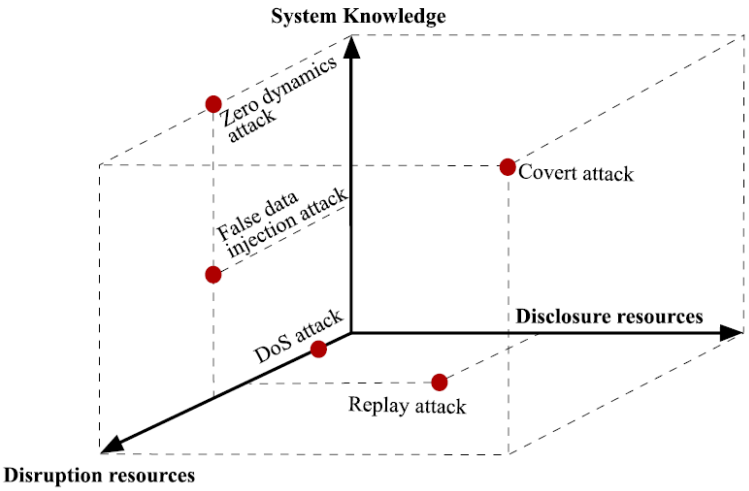


Figure 2.2: Attacks position in the cyber-attack space [13].

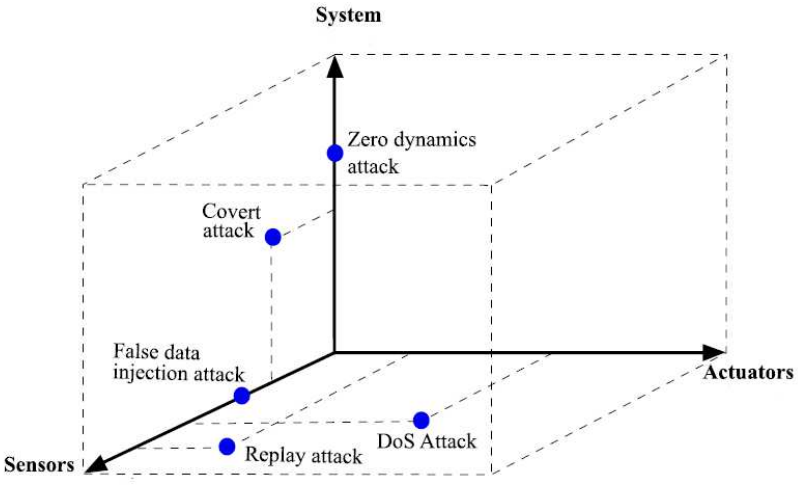


Figure 2.3: Attacks position in the impact space [13].

considered by mechanism. It should be noted that a DoS attack can be easily detected and is not stealthy. However, it can be considered as poor network connectivity. The implementation of this cyber-attack does not require any a priori knowledge of the system nor the disclosure resource. Preventing the data transmission affects the availability of the resources. This cyber-attack can be injected into the communication link among two agents, between controller and actuator, and on the link between sensor measurement and observer or controller depending on the mechanism.

2.1.2 False Data Injection Attack

False data injection attack usually refers to adding a bias term to the sensor or actuator signal such that it is undetectable while it has the maximum impact on the stability or degrades the performance of the system depending on the adversarial aim. A false data injection attack requires disruption capabilities to manipulate the sensor measurements or controller to actuator signals depending on the attackers goal. Also, in order to remain stealthy, the adversary requires the closed-loop system knowledge and existing anomaly detectors. False data injection is one of the most well-known and easy to implement cyber-attacks on multi-agent systems. The attacker may compromise the communication link between two agents or among agents and command and control block.

2.1.3 Replay Attack

Replay attack consists of two stages; in the first stage the adversary eavesdrops and records the data from time t_0 to t_k, then he/she re-plays the recorded signal from t_{k+1} until the end of the cyber-attack. In other words, this cyber-attack replaces the recorded data with real-time data on either or both sensor to controller or controller to actuator data. Replay attacks are mainly

injected into the security cameras or movies such that the recorded video will replay to hide the theft and fool the human operator. Injecting replay attack may be detectable or undetectable by anomaly detectors depending on the condition and application of the system. The attacker requires to disclose resources and disrupt the sensor and actuator channels by replacing the recorded data.

2.1.4 Covert Attack

In the covert attack, an adversary compromises both the actuator and output measurement signal to ensure stealthiness and a low possibility of being detected. Depending on the attacker's goal, he/she can manipulate the output measurement which corresponds to the under cyber-attack actuator signal. It is worth mentioning that in covert attack the full knowledge of the system is required in addition to the ability to compromise input and output channels to each agent. This cyber-attack mainly occurs between agents and the command and control subsystem and not on the communication channel between agents because it is usually assumed that agents transmit their output measurement among their neighbors.

2.1.5 Zero Dynamics Attack

The adversary injects zero dynamics attacks (ZDAs) by exploiting the linearity of the cyber-attack and exciting the zero dynamics of the plant. In particular, in ZDAs the adversary manipulates the input signal such that the output of the system be zero with the assumption that he/she has the full knowledge of the system. Those nonzero input signals which result in output to be zero are called zero dynamics of the plant and they are a subset of invariant zeros [1]. The invariant zeros can be calculated as the values $\lambda \in \mathbb{C}$ that results in matrix P being rank deficient

where matrices A, B, and C are given in Equation (2.2) and

$$P = \begin{bmatrix} \lambda I - A & -B \\ C & 0 \end{bmatrix}. \tag{2.3}$$

For a continuous-time system, those zeros in the left half plane are called the minimum phase (stable) while the unstable zeros are called the non-minimum phase. The zero dynamics can be obtained by solving the following equality:

$$\begin{bmatrix} \lambda I - A & -B \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.4}$$

where x and u are nonzero states and input to be calculated. Hence, with the nonzero state and control input, the output of the system is zero. Exciting the non-minimum phase zeros can result in a major damage to the system and controller and cause actuator saturation in the system.

2.2 Finite-Frequency Analysis

The generalized Kalman-Yakubovich-Popov (GKYP) lemma is an extension to the standard KYP lemma which characterizes the frequency domain inequalities problems and converts them into convex LMIs in finite frequency. Consider a linear time-invariant system

$$\begin{aligned} \dot{x}(t) = &\quad Ax(t) + Bu(t), \\ y(t) = &\quad Cx(t) + Du(t), \end{aligned} \tag{2.5}$$

Table 2.1: Ω *and* Ξ *Different Frequency Ranges*

	LF	MF	HF
Ω	$\lvert\omega\rvert \le \omega_l$	$\omega_1 \le \omega \le \omega_2$	$\lvert\omega\rvert \ge \omega_h$
Ξ	$\begin{bmatrix} -Q & P \\ P & \omega_l^2 Q \end{bmatrix}$	$\begin{bmatrix} -Q & P + j\omega_c Q \\ P - j\omega_c Q & -\omega_l\omega_2 Q \end{bmatrix}$	$\begin{bmatrix} Q & P \\ P & -\omega_h^2 Q \end{bmatrix}$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^u$, and $y(t) \in \mathbb{R}^y$ are state, control input, and output of the system, respectively with the transfer function matrix $G(s)$. For a given system (2.5) and its corresponding transfer function, the Generalized KYP lemma is as follows [92]:

Let a symmetric matrix $\Pi \in \mathbb{R}^{(n+n_y)\times(n+n_y)}$ be given, the following statements are equivalent:

(i) The finite frequency inequality

$$\begin{bmatrix} G(j\omega) \\ I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} G(j\omega) \\ I \end{bmatrix} < 0, \quad \forall \omega \in \Omega \tag{2.6}$$

(ii) There exist Hermitian matrices $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{n\times n}$ satisfying $Q > 0$, and

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} C & D \\ 0 & I \end{bmatrix} < 0, \tag{2.7}$$

where Ω and Ξ are defined in Table 2.1 for continuous-time system in which $\omega_c = 1/2(\omega_1 + \omega_2)$.

***Definition* 1** *The finite frequency $\mathscr{H}_-$ of a transfer function $G(j\omega)$ is defined as follows:*

$$||G(j\omega)||_- := \inf_{\omega} \sigma_{\min}(G(j\omega)), \quad \forall \omega \in \Omega \tag{2.8}$$

Definition 2 *The finite frequency $\mathscr{H}_\infty$ of a transfer function $G(j\omega)$ is defined as follows:*

$$||G(j\omega)||_\infty := \sup_\omega \sigma_{\max}(G(j\omega)), \quad \forall \omega \in \Omega \tag{2.9}$$

Lemma 1 *($\mathscr{H}_-$ condition [23])* Consider system (2.5) and its corresponding transfer function matrix. Given symmetric matrix $\Pi = \begin{bmatrix} -I & 0 \\ 0 & \beta^2 I \end{bmatrix}$, and a frequency bound Ω, the following statements are equivalent:

$$\begin{aligned} &\text{(i)} \ \ \sigma_{\min}(G(j\omega)) > \beta, \quad \forall \omega \in \Omega \\ &\text{(ii)} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} C & D \\ 0 & I \end{bmatrix} < 0. \end{aligned} \tag{2.10}$$

Remark: It is worth considering that $\mathscr{H}_-$ in the full frequency domain is a special case of Lemma 1. According to Theorem 3 in [93], the followng two statements are equivalent considering system (2.5) and its corresponding transfer function matrix.

$$\begin{aligned} &\text{(i)} \ \ \sigma_{\min}(G(j\omega)) > \beta, \quad \forall \omega \in [0, \infty) \\ &\text{(ii)} \begin{bmatrix} PA + A^{\mathrm{T}}P + C^{\mathrm{T}}C & PB + C^{\mathrm{T}}D \\ * & D^{\mathrm{T}}D - \beta^2 I \end{bmatrix} > 0. \end{aligned} \tag{2.11}$$

For $D^{\mathrm{T}}D - \beta^2 I > 0$, D should be full rank matrix which is not applicable because matrix D can have any value. Accordingly, it is trivial that for any strictly proper system in which $D = 0$, $\mathscr{H}_-$ analysis in full frequency domain cannot be obtained [23]. In other words, it is worth considering that it is not feasible to design $\mathscr{H}_-/\mathscr{H}_\infty$ fault detection observer over the entire frequency range for a system when the sensor fault distribution matrix is not of full column

rank. In this situation, the $\mathscr{H}_-$index performance of the error system is always zero.

***Lemma* 2** *($\mathscr{H}_\infty$ condition [23])* Consider system (2.5) and its corresponding transfer function matrix. Given symetric matrix $\Pi = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix}$, and a frequency bound Ω, the following statements are equivalent:

$$\begin{aligned} &\text{(i) } \sigma_{\max}(G(j\omega)) < \gamma, \quad \forall \omega \in \Omega \\ &\text{(ii) } \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} C & D \\ 0 & I \end{bmatrix} < 0. \end{aligned} \tag{2.12}$$

According to Definition 1 and Lemma 1, larger value of β results in higher sensitivity of residual signal to the output. For instance, considering a system with the attack as an input, and the residual signal as an output, larger the value of β results in better detection of the attack. Moreover, based on the Definition 2 and Lemma 2, the smaller value of γ implies more attenuation of the effect of output on the residual signal. For example, if a fault signal is considered as an input to a system and the residual signal as an output, the smaller the value of γ ensures the better attenuation of the effects of fault on the residual signal.

2.3 General Lemmas

The following lemmas are used in the design procedure of the proposed detection filters.

***Lemma* 3** *(Projection Lemma)* Let $\Theta \in \mathbb{R}^{m \times m}$, U, and V with column dimension m be given. There exists a matrix F satisfying:

$$U^{\mathrm{T}} F V + V^{\mathrm{T}} F U + \Theta < 0, \tag{2.13}$$

if and only if $N_U^{\mathrm{T}}\Theta N_U < 0$ and $N_V^{\mathrm{T}}\Theta N_V < 0$, where N_U and N_V are arbitrary matrices with, columns form a basis of the null space of U and V, respectively.

***Lemma* 4** *(Finsler's Lemma)* Let $\zeta \in \mathbb{R}^n$, $P \in \mathbb{R}^{n\times n}$, and $H \in \mathbb{R}^{n\times m}$. Let $H^{\perp}$ be any matrix such that $H^{\perp}H = 0$. The following statements are equivalent.

1. $\zeta^{\star}P\zeta < 0, \ \ \forall H^{\star}\zeta = 0, \ \ \zeta \neq 0$

2. $H^{\perp}PH^{\perp\star} < 0,$

3. $\exists X \in \mathbb{R}^{m\times n} : P + HX + X^{\star}H^{\star} < 0.$

***Lemma* 5 ([23])** Defining the following matrices,

$$\mathscr{A} = \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}, \quad \mathscr{B} = \begin{bmatrix} -C^{\mathrm{T}} \\ -D^{\mathrm{T}} \end{bmatrix},$$
$$\mathscr{L} = \begin{bmatrix} I & 0 \end{bmatrix},$$

and for the state error dynamics equation corresponding to system (2.5) and Luenberger observer:

$$\begin{aligned} \dot{\tilde{x}}(t) &= A\tilde{x}(t) + Bu(t) + L(y(t) - \tilde{y}(t)) \\ \tilde{y}(t) &= C\tilde{x}(t) + Du(t) \end{aligned} \tag{2.14}$$

is defined as follows:

$$\begin{aligned} \dot{e}(t) &= \bar{A}e(t) + \bar{B}u(t), \\ r(t) &= Ce(t) + Du(t), \end{aligned} \tag{2.15}$$

where $\bar{A} = A - LC$, $\bar{B} = B - LD$, and L is the Luenberger gain observer. Let $R \in \mathbb{R}^{n\times(2n+n_u+n_y)}$, $\Pi \in \mathbb{R}^{(n_y+n_u)\times(n_y+n_u)}$ be given. The following inequality

$$\begin{bmatrix} \bar{A} & \bar{B} \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} \bar{A} & \bar{B} \\ I & 0 \end{bmatrix} + \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} C & D \\ 0 & I \end{bmatrix} < 0. \tag{2.16}$$

is satisfied if there exist $n \times n$ matrix variables X, Y, P, $Q > 0$, V, and $\mathscr{X} \in X(\mathscr{L}, R)$ of appropriate dimensions such that

$$T \begin{bmatrix} \Xi & 0 \\ 0 & \Pi \end{bmatrix} T^{\mathrm{T}} < \mathrm{He} \begin{bmatrix} -\mathscr{X} \\ \mathscr{A}\mathscr{X} + \mathscr{B}YR \end{bmatrix}, \tag{2.17}$$

where $Y = L^{\mathrm{T}}X$ and $T = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix}$ is a permutation matrix defined as follows:

$$\begin{bmatrix} M_1 & M_2 & M_3 & M_4 \end{bmatrix} T = \begin{bmatrix} M_1 & M_3 & M_2 & M_4 \end{bmatrix}, \tag{2.18}$$

and

$$\begin{aligned} X(\mathscr{L}, R) := [\mathscr{L}^{\dagger} XR + (I - \mathscr{L}^{\dagger}\mathscr{L})V | X \in \mathbb{R}^{n\times n}, \\ \det(X) \neq 0, V \in \mathbb{R}^{(n+n_u)\times(2n+n_u+n_r)}]. \end{aligned} \tag{2.19}$$

***Lemma* 6** *(Bounded Real Lemma):* Consider system (2.5) and its corresponding transfer function. Given a $\gamma > 0$, for a symmetric positive definite matrix $P = P^{\mathrm{T}} > 0$, the $\mathscr{H}_\infty$ performance

is equivalent to the following:

$$\begin{bmatrix} A^{\mathrm{T}}P + PA & PB & C^{\mathrm{T}} \\ \star & -\gamma^2 I & D^{\mathrm{T}} \\ \star & \star & -I \end{bmatrix} < 0. \tag{2.20}$$

2.4 Graph Theory

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of vertices $\mathcal{V} = 1, 2, ..., N$ and a set of edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ which present the communication link between each node. The graph $\mathcal{G}$ is called undirected if $(i, j) \in \mathcal{E}$ implies $(j, i) \in \mathcal{E}$. The adjacency matrix $\mathcal{A}$ of graph $\mathcal{G}$ delineates the link between agents i and j. In particular, if there exists a link from agent i to agent j, the element a_{ij} is 1. The neighbor of agent i is defined as $\mathcal{N}_i$ is a set of nodes that have an edge with agent i. The Laplacian matrix of a graph is defined as $\mathcal{L} = \mathcal{D} - \mathcal{A}$ where $\mathcal{D}$ is the in-degree matrix.

2.5 Output Consensus Protocol

In order to design a monitoring system for multi-agent systems in Chapter 4, an output consensus protocol is used as follows. This method is well described in [71] and only the main results are brought in this section. For N homogeneous agents with linear time-invariant state space model given as:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + Bu_i(t), \\ y_i(t) &= Cx_i(t), \end{aligned} \tag{2.21}$$

where it is assumed that (C, A, B) is stabilizable and detectable. The communication topology among the agents is assumed to be directed. It is assumed that agent i collects the output information $y_j, \quad j \in \mathcal{N}_i$ to reach a consensus. The input is generated by the following rule:

$$u_i(t) = \mathcal{K}(t) \sum_{j \in \mathcal{N}_i} a_{ij}(y_i(t) - y_j(t)). \tag{2.22}$$

The consensus problem is solved by considering $\mathcal{K}(t)$ as a stable filter with the transfer function $\mathcal{K}(s) = C_{\mathcal{K}}(SI - A_{\mathcal{K}})^{-1}B_{\mathcal{K}} + D_{\mathcal{K}}$, which represent the following state space model:

$$\begin{aligned} \dot{\chi}(t) =& A_{\mathcal{K}}\chi(t) + B_{\mathcal{K}}z_i(t), \\ u(t) =& C_{\mathcal{K}}\chi(t) + D_{\mathcal{K}}z_i(t), \end{aligned} \tag{2.23}$$

where $A_{\mathcal{K}}$, $B_{\mathcal{K}}$, $C_{\mathcal{K}}$, and $D_{\mathcal{K}}$ are design parameters and $z_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij}(y_i(t) - y_j(t))$. According to Theorem 1 and Theorem 4 in [71], it can be shown that:

$$\begin{aligned} \dot{\chi}(t) =& A\chi(t) + BB^{\mathrm{T}}P\chi(t) - K(y_j(t) - C\chi(t)), \\ u(t) =& B^{\mathrm{T}}P\chi(t), \end{aligned} \tag{2.24}$$

where K is a matrix such that $(A_K C)$ is Hurwitz and P is the unique solution of

$$A^{\mathrm{T}}P + PA - PBB^{\mathrm{T}}P = 0. \tag{2.25}$$

Chapter 3

Robust Fault and Cyber-Attack Detection in Cyber-Physical Systems

The problem of fault and cyber-attack detection in a cyber-physical systems in the presence of disturbance has been investigated in this chapter. A robust Luenberger-based fault detection observer is designed in the plant locally which generates a residual signal which is sensitive to fault signal. The residual signal corresponding to fault detector is transmitted through a wireless network to the command and control station. The cyber-attack is injected on the output signal of the plan into the communication link between the plant and the command and control station. Hence, the fault detector residual is independent of the cyber-attack. Then, a Luenberger-based cyber-attack detector is designed in the command and control station. The generated residual from cyber-attack detector is sensitive to cyber-attack while the effect of disturbance and fault has been attenuated on this signal. However, the cyber-attack detector residual experiences a huge overshoot in the presence of fault. Thus, a new filter is designed for this purpose and their performance has been compared. It is shown by the simulation that the new filter named as normal filter has better performance in cyber-attack detection.

3.1 System Description

Consider the following linear time-invariant system in the presence of sensor and actuator faults, sensor cyber-attacks, and disturbances, namely

$$\begin{aligned} \dot{x}(t) &= A_{\text{op}}x(t) + Bu(t) + B_{\text{f}_{\text{op}}}f(t) + B_{\text{d}_{\text{op}}}d(t), \\ y(t) &= Cx(t) + D_{\text{f}}f(t) + D_{\text{d}}d(t), \\ y_{\text{a}}(t) &= y(t) + D_{\text{a}}a(t), \end{aligned} \tag{3.1}$$

where $x(t)$, $u(t) \in \mathbb{R}^{n_u}$, $y(t) \in \mathbb{R}^{n_y}$, and $y_{\text{a}}(t) \in \mathbb{R}^{n_y}$ denote the state, input, healthy output, and under cyber-attack output of the system, $f(t) \in \mathbb{R}^{n_f}$, $a(t) \in \mathbb{R}^{n_a}$, and $d(t) \in \mathbb{R}^{n_d}$ present $\mathbb{L}_2$-norm bounded fault, cyber-attack, and disturbance signals, respectively. The matrices A_{op}, B, B_{f}, $B_{\text{d}_{\text{op}}}$, C, D_{f}, D_{a}, and D_{d} are known with appropriate dimensions. System (3.1) is represented by considering a local output feedback controller $u(t) = K_1 y(t) + K_2 y_{\text{ref}}(t) = K_1(Cx(t) + D_{\text{f}}f(t) + D_{\text{d}}d(t)) + K_2 y_{\text{ref}}(t)$ as follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_{\text{f}}f(t) + B_{\text{d}}d(t) + B_{\text{c}}y_{\text{ref}}(t), \\ y_{\text{a}}(t) &= Cx(t) + D_{\text{a}}a(t) + D_{\text{f}}f(t) + D_{\text{d}}d(t), \end{aligned} \tag{3.2}$$

where $A = (A_{\text{op}} + BKC)$, $B_{\text{f}} = BK_1D_{\text{f}} + B_{\text{f}_{\text{op}}}$, $B_{\text{d}} = BK_1D_{\text{d}} + B_{\text{d}_{\text{op}}}$, and $B_{\text{c}} = BK_2$. This is now shown in Figure 3.1 which represents a CPS under faults and cyber-attacks. Accordingly, system (3.2) will be considered for cyber-attack detection filter design and system (3.1) is used for fault detection filter in the rest of this thesis.

It is assumed that only sensor measurements are transmitted through the cyber system so the actuators are not prone to cyber-attacks while physical faults may occur on actuators and sensors. The reference signal is sent through a secure network to the local controller in the

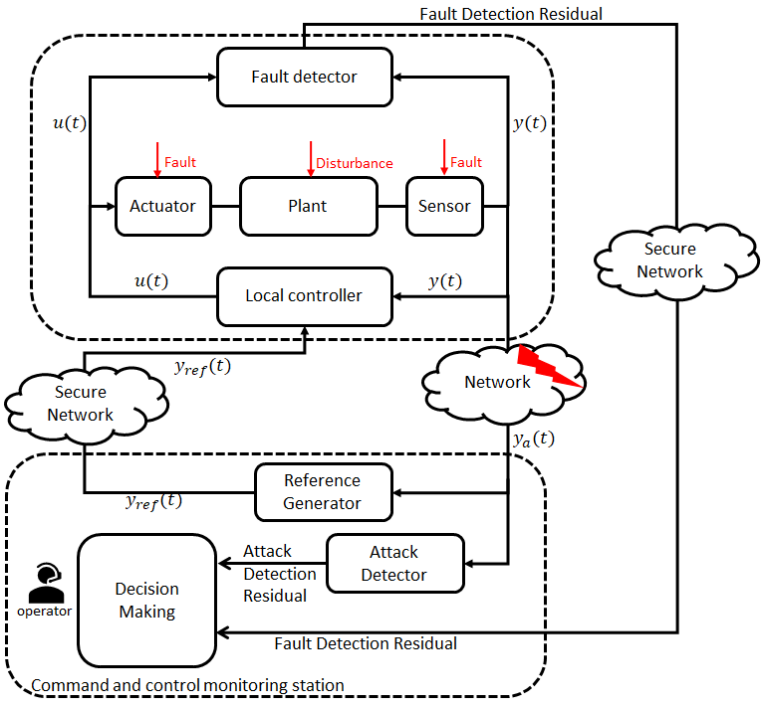


Figure 3.1: CPS in the presence of fault and cyber-attack.

physical system. Moreover, the fault detection residual signal is also sent through a secure network channel and is not prone to cyber-attack. Also, some coding scheme can be used for transmitting the reference signal to ensure its secure transmission. The output from monitoring station is a residual signal which is observed by an operator to take a proper action in the presence of anomalies.

3.2 Problem Formulation

As discussed earlier, the main aim of this thesis is to design fault and cyber-attack detection mechanisms for cyber-physical systems in the presence of disturbances. This problem is solved by using multi-objective framework and $\mathscr{H}_-$ and $\mathscr{H}_\infty$ formulation. Two types of observers are designed as follows. In particular, a cyber-attack detection filter is designed such that the effect of cyber-attack signal on the residual signal is maximized and the effect of disturbance and fault signal on the residual signal is minimized. Also, the fault detection filter is designed such

that the sensitivity of the residual signal to the fault signal is maximized while the effect of disturbance is minimized. It is worth mentioning that the fault detector is designed locally in the plant. The main motivation behind designing the second type of observer which is called Normal filter is to provide more design degrees of freedom. Accordingly, the second type of observer (Normal Filter) is chosen because it has more flexibility in design and more design parameters.

3.2.1 Luenberger Observer

Two Luenberger observers are designed for fault and cyber-attack detection, respectively as follows:

Fault Detection Observer: The dynamics of the observer is given by,

$$\begin{aligned} \dot{\hat{x}}(t) &= A_{\text{op}}\hat{x}(t) + Bu(t) + L_{\text{fd}}(y(t) - \hat{y}(t)), \\ \hat{y}(t) &= C\hat{x}(t), \\ r_{\text{Lfd}}(t) &= y(t) - \hat{y}(t), \end{aligned} \tag{3.3}$$

where $\hat{x}(t) \in \mathbb{R}^n$, $\hat{y}(t) \in \mathbb{R}^{n_y}$, and $r_{\text{Lfd}}(t) \in \mathbb{R}^{n_y}$ are state, output, and residual signal of the Luenberger-based fault detection observer, respectively and the observer gain L_{fd} is a design parameter. As mentioned earlier in this chapter, the fault detector is designed locally by considering the open loop dynamic equation given in (3.1). Let $e_{\text{fd}}(t) = x(t) - \hat{x}(t)$, the state error dynamics can be written as follows:

$$\begin{aligned} \dot{e}_{fd}(t) &= \bar{A}_1 e_{\text{fd}}(t) + \bar{B}_{\text{f1}} f(t) + \bar{B}_{\text{d1}} d(t), \\ r_{\text{Lfd}}(t) &= C e_{\text{fd}}(t) + D_{\text{f}} f(t) + D_{\text{d}} d(t), \end{aligned} \tag{3.4}$$

where $\bar{A}_1 = A_{\text{op}} - L_{\text{fd}}C$, $\bar{B}_{\text{f1}} = B_{\text{f}_{\text{op}}} - L_{\text{fd}}D_{\text{f}}$, and $\bar{B}_{\text{d1}} = B_{\text{d}_{\text{op}}} - L_{\text{fd}}D_{\text{d}}$. Accordingly, L_{fd} should be designed such that the following conditions are met:

i) $A - L_{\text{fd}}C$ is Hurwitz

ii) $||G_{\text{r}_{\text{Lfd}}\text{f}}(j\omega_{\text{f}})||_{\_} > \beta_1, \quad \forall \quad \omega_{\text{f}} \in [0, \bar{\omega}_{\text{f}}]$

iii) $||G_{\text{r}_{\text{Lfd}}\text{d}}(j\omega_{\text{d}})||_{\infty} < \gamma_1, \quad \forall \quad \omega_{\text{d}} \in [0, \bar{\omega}_{\text{d}}]$

where β_1 and γ_1 are design parameters and $\bar{\omega}_{\text{f}}$ and $\bar{\omega}_{\text{d}}$ are the given parameter denoted as upper bound of frequency range corresponding to fault and disturbance, respectively. Transfer matrices $G_{\text{r}_{\text{Lfd}}\text{f}}$ and $G_{\text{r}_{\text{Lfd}}\text{d}}$ that represent the effect of fault and disturbance on the residual signal, respectively are as follows:

$$\begin{aligned} G_{\text{r}_{\text{Lfd}}\text{f}}(j\omega_{\text{f}}) &= C(j\omega_{\text{f}}I - A_{\text{op}} + L_{\text{fd}}C)^{-1}(B_{\text{f}_{\text{op}}} - L_{\text{fd}}D_{\text{f}}) + D_{\text{f}}, \\ G_{\text{r}_{\text{Lfd}}\text{d}}(j\omega_{\text{d}}) &= C(j\omega_{\text{d}}I - A_{\text{op}} + L_{\text{fd}}C)^{-1}(B_{\text{d}_{\text{op}}} - L_{\text{fd}}D_{\text{d}}) + D_{\text{d}}, \end{aligned} \tag{3.5}$$

In order to achieve the aim of fault detection observer, its corresponding residual should be sensitive to fault while attenuating the effect of cyber-attack, disturbance, and reference signal. Hence, according to Definitions 1 and 2 and Lemmas 1 and 2, the best performance will be achieved by maximizing β_1 and minimizing γ_1.

Cyber-Attack Detection Observer: The dynamics of the observer is given by,

$$\begin{aligned} \dot{\tilde{x}}(t) &= A\tilde{x}(t) + L_{\text{ad}}(y_{\text{a}}(t) - \tilde{y}(t)), \\ \tilde{y}(t) &= C\tilde{x}(t), \\ r_{\text{Lad}}(t) &= y_{\text{a}}(t) - \tilde{y}(t), \end{aligned} \tag{3.6}$$

where $\tilde{x}(t) \in \mathbb{R}^n$, $\tilde{y}(t) \in \mathbb{R}^{n_y}$, and $r_{\text{Lad}}(t) \in \mathbb{R}^{n_y}$ are state, output, and residual signal of the Luenberger-based cyber-attack detector observer, respectively and the observer gain L_{ad} is a design parameter. Let $e_{\text{ad}}(t) = x(t) - \tilde{x}(t)$, considering the error dynamics given in (3.2) the

state error dynamics can be written as follows:

$$\begin{aligned} \dot{e}_{ad}(t) &= \bar{A}_2 e_{\mathrm{ad}}(t) + \bar{B}_{\mathrm{f2}} f(t) + \bar{B}_{\mathrm{a2}} a(t) + \bar{B}_{\mathrm{d2}} d(t) + B_c y_{\mathrm{ref}}(t), \\ r_{\mathrm{Lad}}(t) &= C e_{\mathrm{ad}}(t) + D_{\mathrm{a}} a(t) + D_{\mathrm{d}} d(t) + D_{\mathrm{f}} f(t), \end{aligned} \tag{3.7}$$

where $\bar{A}_2 = A - L_{\mathrm{ad}} C$, $\bar{B}_{\mathrm{a2}} = -L_{\mathrm{ad}} D_{\mathrm{a}}$, $\bar{B}_{\mathrm{f2}} = B_{\mathrm{f}} - L_{\mathrm{ad}} D_{\mathrm{f}}$, and $\bar{B}_{\mathrm{d2}} = B_{\mathrm{d}} - L_{\mathrm{ad}} D_{\mathrm{d}}$. The gain of cyber-attack detection Luenberger observer L_{ad} should be designed such that the following conditions hold:

i) $A - L_{\mathrm{ad}} C$ is Hurwitz

ii) $||G_{\mathrm{r_{Lad}a}}(j\omega_{\mathrm{a}})||_{\_} > \beta_2, \quad \forall \quad \omega_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}]$

iii) $||G_{\mathrm{r_{Lad}d}}(j\omega_{\mathrm{d}})||_{\infty} < \gamma_2, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}]$

iv) $||G_{\mathrm{r_{Lad}f}}(j\omega_{\mathrm{f}})||_{\infty} < \gamma_3, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}]$

v) $||G_{\mathrm{r_{Lad}y_{ref}}}(j\omega_{\mathrm{y}})||_{\infty} < \gamma_4, \quad \forall \quad \omega_{\mathrm{y}} \in [0, \bar{\omega}_{\mathrm{y}}]$

where

$$\begin{aligned} G_{\mathrm{r_{Lad}f}}(j\omega_{\mathrm{f}}) &= C(j\omega_{\mathrm{f}} I - A + L_{\mathrm{ad}} C)^{-1}(B_{\mathrm{f}} - L_{\mathrm{ad}} D_{\mathrm{f}}) + D_{\mathrm{f}}, \\ G_{\mathrm{r_{Lad}d}}(j\omega_{\mathrm{d}}) &= C(j\omega_{\mathrm{d}} I - A + L_{\mathrm{ad}} C)^{-1}(B_{\mathrm{d}} - L_{\mathrm{ad}} D_{\mathrm{d}}) + D_{\mathrm{d}}, \\ G_{\mathrm{r_{Lad}a}}(j\omega_{\mathrm{a}}) &= C(j\omega_{\mathrm{a}} I - A + L_{\mathrm{ad}} C)^{-1}(-L_{\mathrm{ad}} D_{\mathrm{a}}) + D_{\mathrm{a}}, \\ G_{\mathrm{r_{Lad}y_{ref}}}(j\omega_{\mathrm{y}}) &= C(j\omega_{\mathrm{y}} I - A + L_{\mathrm{ad}} C)^{-1} B_{\mathrm{c}}. \end{aligned} \tag{3.8}$$

represents the transfer matrices that measure the effect of fault, disturbance, cyber-attack, and reference signal, respectively on the residual signal corresponding to the cyber-attack detector observer.

Remark: It is worth considering that the fault, cyber-attack, and disturbance frequency ranges may overlap and it affects the efficiency of the detection scheme. However, the worst-case

scenario is considered in this chapter in which the frequency range of cyber-attack and fault overlaps. In addition, disturbances usually occur in high frequencies while faults occur in low frequency [23]. Accordingly, the upper bound corresponding to frequency of disturbance is set much higher than that of fault. However, to consider the worst-case scenario, the frequency range for cyber-attack and fault is considered to be identical.

Remark: The fault frequency varies depending on the system and application. The frequency of cyber-attack may vary based on cyber-attacker design. There exists some approaches in the literature to estimate the frequency range for cyber-attack [39], however, it is beyond the scope of this thesis and to prove the efficacy of the proposed method, a low frequency range cyber-attack is considered.

3.2.2 Normal Filter

Similarly, a normal filter is designed for cyber-attack detection as follows. It is worth mentioning that this filter is well designed for detecting anomalies on the output signal in the [26; 36]. Also, in order to generate a residual signal it just requires the output signal. Accordingly, it is designed for cyber-attack detection and its performance is compared with the Luenberger-based cyber-attack detector in Section 3.6. Without loss of generality, cyber-attack detection filter is formulated as follows:

$$\begin{aligned} \dot{x}_{\mathrm{ad}}(t) &= A_{\mathrm{ad}} x_{\mathrm{ad}}(t) + B_{\mathrm{ad}} y_{\mathrm{a}}(t), \\ r_{\mathrm{ad}}(t) &= C_{\mathrm{ad}} x_{\mathrm{ad}}(t) + D_{\mathrm{ad}} y_{\mathrm{a}}(t), \end{aligned} \tag{3.9}$$

where $x_{\mathrm{ad}}(t) \in \mathbb{R}^n$ and $r_{\mathrm{ad}}(t) \in \mathbb{R}^r$ denote the state of the filter and residual signal, respectively and matrices A_{ad}, B_{ad}, C_{ad}, and D_{ad} are design parameters. The main motivation behind

choosing this filter is that it has four parameters to be designed and provides more design flexibility compared to the Luenberger observer which has only one design parameter. Let $X(t) = \begin{bmatrix} x(t) \\ x_{\mathrm{ad}}(t) \end{bmatrix} \in \mathbb{R}^{2n}$, the augmented system can be written as follows:

$$\begin{aligned} \dot{X}(t) &= \bar{A}X(t) + \bar{B}_{\mathrm{f}} f(t) + \bar{B}_{\mathrm{d}} d(t) + \bar{B}_{\mathrm{a}} a(t) + \bar{B}_{\mathrm{c}} y_{\mathrm{ref}}(t), \\ r_{\mathrm{ad}}(t) &= \bar{C}X(t) + \bar{D}_{\mathrm{f}} f(t) + \bar{D}_{d} d(t) + \bar{D}_{\mathrm{a}} a(t), \end{aligned} \tag{3.10}$$

where

$$\begin{aligned} \bar{A} &= \begin{bmatrix} A & 0 \\ B_{\mathrm{ad}}C & A_{\mathrm{ad}} \end{bmatrix}, \quad \bar{B}_{\mathrm{f}} = \begin{bmatrix} B_{\mathrm{f}} \\ B_{\mathrm{ad}}D_{\mathrm{f}} \end{bmatrix}, \\ \bar{B}_{\mathrm{d}} &= \begin{bmatrix} B_{\mathrm{d}} \\ B_{\mathrm{ad}}D_{\mathrm{d}} \end{bmatrix}, \qquad \bar{B}_{\mathrm{a}} = \begin{bmatrix} 0 \\ B_{\mathrm{ad}}D_{\mathrm{a}} \end{bmatrix}, \\ \bar{B}_{\mathrm{c}} &= \begin{bmatrix} B_{\mathrm{c}} \\ 0 \end{bmatrix}, \qquad \hat{C} = \begin{bmatrix} D_{\mathrm{ad}}C & C_{\mathrm{ad}} \end{bmatrix}, \\ \bar{D}_{\mathrm{a}} &= D_{\mathrm{ad}}D_{\mathrm{a}}, \qquad \bar{D}_{d} = D_{\mathrm{ad}}D_{\mathrm{d}}, \qquad \bar{D}_{f} = D_{\mathrm{ad}}D_{\mathrm{f}}. \end{aligned} \tag{3.11}$$

Then, a low frequency cyber-attack detection observer is designed such that the following conditions are met:

$$\begin{aligned} &\text{i)} \ ||\bar{G}_{\mathrm{r_{ad}a}}(j\omega_{\mathrm{a}})||\_ > \beta_3, \quad \forall \quad \omega_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}] \\ &\text{ii)} \ ||\bar{G}_{\mathrm{r_{ad}d}}(j\omega_{\mathrm{d}})||_\infty < \gamma_5, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}] \\ &\text{iii)} \ ||\bar{G}_{\mathrm{r_{ad}f}}(j\omega_{\mathrm{f}})||_\infty < \gamma_6, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}] \\ &\text{iv)} \ ||\bar{G}_{\mathrm{r_{ad}y}}(j\omega_{\mathrm{y}})||_\infty < \gamma_7, \quad \forall \quad \omega_{\mathrm{y}} \in [0, \bar{\omega}_{\mathrm{y}}] \end{aligned} \tag{3.12}$$

where

$$\begin{aligned}
\bar{G}_{\mathrm{r_{ad}f}}(j\omega_{\mathrm{f}}) &= \bar{C}(j\omega_{\mathrm{f}} I - \bar{A})^{-1}\bar{B}_{\mathrm{f}} + \bar{D}_{\mathrm{f}}, \\
\bar{G}_{\mathrm{r_{ad}d}}(j\omega_{\mathrm{d}}) &= \bar{C}(j\omega_{\mathrm{d}} I - \bar{A})^{-1}\bar{B}_{\mathrm{d}} + \bar{D}_{\mathrm{d}}, \\
\bar{G}_{\mathrm{r_{ad}a}}(j\omega_{\mathrm{a}}) &= \bar{C}(j\omega_{\mathrm{a}} I - \bar{A})^{-1}\bar{B}_{\mathrm{a}} + \bar{D}_{\mathrm{a}}, \\
\bar{G}_{\mathrm{r_{ad}y}}(j\omega_{\mathrm{y}}) &= \bar{C}(j\omega_{\mathrm{y}} I - \bar{A})^{-1}\bar{B}_{\mathrm{c}}.
\end{aligned} \tag{3.13}$$

The main advantage of designing the normal filter as compared to the Luenberger filter is as follows. The proposed filters in this thesis are designed based on the definition of transfer matrices presenting the effects of faults, cyber-attacks, disturbances, and reference signal with respect to the residual. In the case of cyber-attack detection Luenberger observer, the transfer matrices given in (3.8) has only one design parameter L_{ad}. However, in the case of normal filter design the transfer matrices provided in (3.13) has A_{ad}, B_{ad}, C_{ad}, and D_{ad} to be designed which bring higher flexibility in design. This is similar to design of fixed gain feedback controller in comparison with the design of a dynamic controller. The dynamic controller has more design parameters which brings more flexibility in changing the place of poles of the system to achieve a desired aim which fixed gain controller has less design flexibility and is more restricted in changing the poles.

3.2.3 Detection and Decision Making

Up to this point, a fault detector is designed on the plant side and it generates a residual signal. A fixed-value threshold is calculated using Monte Carlo simulation in the presence of randomly generated cyber-attacks and disturbances. Thus, a fault is detected if the residual corresponding to the fault detection (r_{fd}) exceeds the threshold (i.e. $r_{\mathrm{fd}} = 1$). On the other hand, similar ap-

proach has been utilized in order to detect cyber-attack in the presence of fault and disturbance. Hence, if the residual signal corresponding to cyber-attack detector (r_{ad}) exceeds a predefined threshold (i.e. $r_{\text{ad}} = 1$), the cyber-attack is detected. It is worth mentioning that if the residuals remain below threshold, they considered as zero flag as shown in table 4.1. In the case of concurrent fault and cyber-attack the classification can be done by simply observing the generated residuals. The decision making process can be summarized in Table 3.1.

Table 3.1: *Decision Making*

r_{fd}	1	0	1
r_{ad}	0	1	1
Decision	Fault	Attack	Fault and Attack

3.3 Luenberger Observer Design

In this section, four main theorems have been developed for cyber-attack detection filter according to the conditions i-v specified in Section 3.2.1. Also, the theorem corresponding to the fault detection filter is provided according to condition i-iii in Section 3.2.1. However, the proof corresponding to fault detection filter design is omitted to avoid duplication. It is worth mentioning that the cyber-attack detection filter is designed for closed-loop system provided in Equation (3.2) while the fault detection filter is designed for system (3.1).

3.3.1 Cyber-Attack Sensitivity Analysis

Let us assume in equation (3.7) $d(t) = 0$ and $f(t) = 0$. Then, the state dynamic error equation can be written as follows:

$$\begin{aligned} \dot{e}_{ad}(t) &= \bar{A}_2 e_{\mathrm{ad}}(t) + \bar{B}_{\mathrm{a}2} a(t), \\ r_{\mathrm{ad}}(t) &= C e_{\mathrm{ad}}(t) + D_{\mathrm{a}} a(t). \end{aligned} \tag{3.14}$$

In order to increase the sensitivity of the residual signal to the cyber-attack signal which is formulated as $||G_{\mathrm{r_{Lad}a}}(j\omega_{\mathrm{a}})||_{\_} > \beta_2, \quad \forall \quad \omega_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}]$, the following theorem should hold.

***Theorem* 1** Considering the dynamic equation (3.14), for $\Pi = \begin{bmatrix} -I & 0 \\ 0 & \beta_2^2 I \end{bmatrix} \in \mathbb{R}^{(n+n_a)\times(n+n_a)}$ and a positive scalar $\bar{\omega}_{\mathrm{a}}$, the inequality

$$\sigma_{min}(G_{\mathrm{r_{Lad}a}}(j\bar{\omega}_{\mathrm{a}})) > \beta_2, \quad \forall \bar{\omega}_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}] \tag{3.15}$$

holds if there exist matrices $P_1 \in \mathbb{R}^{n\times n}$, $Q_1 > 0 \in \mathbb{R}^{n\times n}$, $X \in \mathbb{R}^{n\times n}$, $Y \in \mathbb{R}^{n\times n}$, $V_{a1} \in \mathbb{R}^{n_r \times n}$, $V_{b1} \in \mathbb{R}^{n_r \times n_r}$, $V_{c1} \in \mathbb{R}^{n_r \times n}$, and $V_{d1} \in \mathbb{R}^{n_r \times n_a}$ such that the following LMI holds:

$$\begin{bmatrix} \phi_1 & V_{a1}^{\mathrm{T}} & \phi_2 & \phi_3 \\ \star & \phi_4 & V_{c1} - V_{b1}^{\mathrm{T}} C & V_{d1} - V_{b1}^{\mathrm{T}} D_{\mathrm{a}} \\ \star & \star & \phi_5 & \phi_6 \\ \star & \star & \star & \phi_7 \end{bmatrix} < 0, \tag{3.16}$$

where $Y = L_{\mathrm{ad}}^{\mathrm{T}} X$ and

$$
\begin{aligned}
\phi_1 &= Q_1 + X + X^{\mathrm{T}}, \\
\phi_2 &= P_1 + X - X^{\mathrm{T}} A - V_{a1}^{\mathrm{T}} C + Y^{\mathrm{T}} C, \\
\phi_3 &= -X B_a - X^{\mathrm{T}} B_a - V_{a1}^{\mathrm{T}} D_{\mathrm{a}} + Y^{\mathrm{T}} D_{\mathrm{a}}, \\
\phi_4 &= -I + V_{b1} + V_{b1}^{\mathrm{T}}, \\
\phi_5 &= \bar{\omega}_{\mathrm{a}}^2 Q_1 + \mathrm{He}(-A^{\mathrm{T}} X - C^{\mathrm{T}} V_{c1} + C^{\mathrm{T}} Y), \\
\phi_6 &= A^{\mathrm{T}} X B_a - C^{\mathrm{T}} V_{d1} - C^{\mathrm{T}} Y B_a - X^{\mathrm{T}} B_a - V_{c1}^{\mathrm{T}} D_{\mathrm{a}} + Y^{\mathrm{T}} D_{\mathrm{a}}, \\
\phi_7 &= \beta_2^2 I + \mathrm{He}(B_a^{\mathrm{T}} X B_a - D_{\mathrm{a}}^{\mathrm{T}} V_{d1} - D_{\mathrm{a}}^{\mathrm{T}} Y B_a).
\end{aligned}
\tag{3.17}
$$

Proof: According to Lemma 1, (3.15) is equivalent to

$$
\begin{bmatrix} \bar{A}_2 & \bar{B}_{\mathrm{a}2} \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} \bar{A}_2 & \bar{B}_{\mathrm{a}2} \\ I & 0 \end{bmatrix} + \Phi < 0, \tag{3.18}
$$

where $\Xi = \begin{bmatrix} -Q_1 & P_1 \\ P_1 & \bar{\omega}_{\mathrm{a}}^2 Q_1 \end{bmatrix}$ and $\Phi = \begin{bmatrix} C & D_{\mathrm{a}} \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} C & D_{\mathrm{a}} \\ 0 & I \end{bmatrix}$. According to Lemma 5, it can be concluded that (3.18) holds if

$$
T \begin{bmatrix} \Xi & 0 \\ 0 & \Pi \end{bmatrix} T^{\mathrm{T}} < \mathrm{He} \begin{bmatrix} -\mathscr{X} \\ \mathscr{A}\mathscr{X} + \mathscr{B} Y R \end{bmatrix}, \tag{3.19}
$$

where $\mathscr{A} = \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ B_a^{\mathrm{T}} & D_{\mathrm{a}}^{\mathrm{T}} \end{bmatrix}$, $\mathscr{B} = \begin{bmatrix} -C^{\mathrm{T}} \\ -D_{\mathrm{a}}^{\mathrm{T}} \end{bmatrix}$, and let $R = \begin{bmatrix} I & 0 & I & -B_a \end{bmatrix}$. T is a permutation matrix defined in Lemma 5 and $\mathscr{X} = \begin{bmatrix} I \\ 0 \end{bmatrix} XR + \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} V$ where $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$. Accordingly, condition (3.19) can be written as follows:

$$T \begin{bmatrix} \Xi & 0 \\ 0 & \Pi \end{bmatrix} T^{\mathrm{T}} < \mathrm{He}\Bigg(\begin{bmatrix} -I & 0 & 0 \\ 0 & -I & 0 \\ A^{\mathrm{T}} & C^{\mathrm{T}} & -C^{\mathrm{T}} \\ B_a^{\mathrm{T}} & D_{\mathrm{a}}^{\mathrm{T}} & -D_{\mathrm{a}}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} XR \\ V_2 \\ YR \end{bmatrix} \Bigg), \tag{3.20}$$

where $V_2 = \begin{bmatrix} V_{a1} & V_{b1} & V_{c1} & V_{d1} \end{bmatrix}$. By some algebraic manipulation it can be shown that (3.20) is equivalent to (3.16). □

Remark: The complex-valued LMIs in the case of considering medium frequency are turned in to real-valued LMIs by observing that complex Hermitian matrix $Q < 0$ if and only if $\begin{bmatrix} \mathrm{Re}(Q) & \mathrm{Im}(Q) \\ -\mathrm{Im}(Q) & \mathrm{Re}(Q) \end{bmatrix} < 0$. It is worth considering that this transformation is done by Matlab LMI solver automatically.

3.3.2 Fault, Reference Signal, and Disturbance Attenuation Analysis

Assuming that in equation (3.7) $d(t) = 0$ and $a(t) = 0$, the state error dynamics can be written as follows:

$$\begin{aligned} \dot{e}_{ad}(t) &= \bar{A}_2 e_{\mathrm{ad}}(t) + \bar{B}_{\mathrm{f2}} f(t), \\ r_{\mathrm{ad}}(t) &= C e_{\mathrm{ad}}(t) + D_{\mathrm{f}} f(t). \end{aligned} \tag{3.21}$$

In order to attenuate the effect of the fault on the residual signal which has been formulated as $||G_{\mathrm{r_{Ladf}}}(j\omega_{\mathrm{f}})||_\infty < \gamma_3, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{d}}]$, the following theorem should be satisfied.

***Theorem* 2** Considering the state dynamic equation (3.21), for $\Pi = \begin{bmatrix} I & 0 \\ 0 & -\gamma_3^2 I \end{bmatrix} \in \mathbb{R}^{(n+n_f)\times(n+n_f)}$ and a positive scalar $\bar{\omega}_{\mathrm{f}}$, the inequality

$$\sigma_{\max}((G_{r_{\mathrm{Lad}}f}(j\bar{\omega}_{\mathrm{f}})) < \gamma_3, \quad \forall \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}] \tag{3.22}$$

holds if there exist matrices $P_3,\ Q_3 > 0,\ X,\ Y,\ V_{a2} \in \mathbb{R}^{n_r \times n},\ V_{b2} \in \mathbb{R}^{n_\times n_r},\ V_{c2} \in \mathbb{R}^{n_r \times n}$, and $V_{d2} \in \mathbb{R}^{n_r \times n_f}$ such that the following LMI condition holds:

$$\begin{bmatrix} -Q_3 & V_{a2}^{\mathrm{T}} & \theta_1 & -V_{a2}^{\mathrm{T}} D_{\mathrm{f}} \\ \star & \theta_2 & V_{c2} - V_{b2}^{\mathrm{T}} C & V_{d2} - V_{b2}^{\mathrm{T}} D_{\mathrm{f}} \\ \star & \star & \theta_3 & \theta_4 \\ \star & \star & \star & \theta_5 \end{bmatrix} < 0, \tag{3.23}$$

where

$$\begin{aligned}
\theta_1 &= P_3 + X - V_{a2}^{\mathrm{T}} C, \\
\theta_2 &= I + V_{b2} + V_{b2}^{\mathrm{T}}, \\
\theta_3 &= \bar{\omega}_{\mathrm{f}}^2 Q_3 + \mathrm{He}(-A^{\mathrm{T}} X - C^{\mathrm{T}} V_{c2} + C^{\mathrm{T}} Y), \\
\theta_4 &= -C^{\mathrm{T}} V_{d2} - X^{\mathrm{T}} B_{\mathrm{f}} - V_{c2}^{\mathrm{T}} D_{\mathrm{f}} + Y^{\mathrm{T}} D_{\mathrm{f}}, \\
\theta_5 &= -\gamma_3^2 I - V_{d2}^{\mathrm{T}} D_{\mathrm{f}} - D_{\mathrm{f}}^{\mathrm{T}} V_{d2}.
\end{aligned} \tag{3.24}$$

Proof: Similar to the proof of Theorem 1. According to Lemma 2 for $R = \begin{bmatrix} 0 & 0 & I & 0 \end{bmatrix}$, the derivation is immediate.

Similar to fault attenuation mechanism, an LMI is derived to attenuate the effect of reference signal which is transmitted to the agent from command and control subsystem on residual signal ans formulated as $||G_{\mathrm{r_{Lad}y_{ref}}}(j\omega_{\mathrm{y}})||_\infty < \gamma_4, \quad \forall \quad \omega_{\mathrm{y}} \in [0, \bar{\omega}_{\mathrm{y}}]$. Hence, assuming that in equation (3.7) $d(t) = 0$, $a(t) = 0$, and $f(t) = 0$, the state error dynamics can be written as follows:

$$\begin{aligned}
\dot{e}_{ad}(t) &= \bar{A}_2 e_{\mathrm{ad}}(t) + \bar{B}_{\mathrm{c}} y_{\mathrm{ref}}(t), \\
r_{\mathrm{ad}}(t) &= C e_{\mathrm{ad}}(t).
\end{aligned} \tag{3.25}$$

***Theorem* 3** Considering the state dynamic equation (3.25), for $\Pi = \begin{bmatrix} I & 0 \\ 0 & -\gamma_4^2 I \end{bmatrix} \in \mathbb{R}^{(n+n_y)\times(n+n_y)}$ and a positive scalar $\bar{\omega}_y$, the inequality

$$\sigma_{\max}(G_{\mathrm{r_{Lad}y_{ref}}}(j\bar{\omega}_y)) < \gamma_4, \quad \forall \omega_{\mathrm{y}} \in [0, \bar{\omega}_y] \tag{3.26}$$

holds if there exist matrices $P_4,\ Q_4 > 0,\ X,\ Y,\ V_{a3} \in \mathbb{R}^{n_r \times n},\ V_{b3} \in \mathbb{R}^{n_r \times n_r},\ V_{c3} \in \mathbb{R}^{n_r \times n}$, and $V_{d3} \in \mathbb{R}^{n_r \times n_y}$ such that the following LMI condition holds:

$$\begin{bmatrix} -Q_4 & V_{a3}^{\mathrm{T}} & \rho_1 & 0 \\ \star & \rho_2 & V_{c3} - V_{b3}^{\mathrm{T}} C & V_{d3} \\ \star & \star & \rho_3 & \rho_4 \\ \star & \star & \star & \rho_5 \end{bmatrix} < 0, \tag{3.27}$$

where

$$\begin{aligned} \rho_1 &= P_4 + X - V_{a3}^{\mathrm{T}} C, \\ \rho_2 &= I + V_{b3} + V_{b3}^{\mathrm{T}}, \\ \rho_3 &= \bar{\omega}_{\mathrm{y}}^2 Q_4 + \mathrm{He}(-A^{\mathrm{T}} X - C^{\mathrm{T}} V_{c3} + C^{\mathrm{T}} Y), \\ \rho_4 &= -C^{\mathrm{T}} V_{d3} - X^{\mathrm{T}} B_{\mathrm{c}}, \\ \rho_5 &= -\gamma_4^2 I. \end{aligned} \tag{3.28}$$

The proof is similar to the proof of Theorem 2.

In addition, the observer gain should be designed such that the effect of disturbance is minimized on the residual signal. Let us consider the following error dynamics equation:

$$\begin{aligned} \dot{e}_{ad}(t) &= \bar{A}_2 e_{\mathrm{ad}}(t) + \bar{B}_{\mathrm{d}2} d(t), \\ r_{\mathrm{Lad}}(t) &= C e_{\mathrm{ad}}(t) + D_{\mathrm{d}} d(t). \end{aligned} \tag{3.29}$$

Hence, the following theorem should be satisfied for cyber-attack detection observers to attenuate the effect of disturbance on residual signal which refers to $||G_{\mathrm{r_{Lad}d}}(j\omega_{\mathrm{d}})||_\infty < \gamma_2, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}]$.

***Theorem* 4** Considering the dynamic equation (3.29), for $\Pi = \begin{bmatrix} I & 0 \\ 0 & -\gamma_2^2 I \end{bmatrix} \in \mathbb{R}^{(n+n_d)\times(n+n_d)}$ and a positive scalar $\bar{\omega_{\mathrm{d}}}$, the inequality

$$\sigma_{\max}(G_{\mathrm{r_{Lad}d}}(j\bar{\omega}_{\mathrm{d}})) < \gamma_3, \quad \forall \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}] \tag{3.30}$$

holds if there exist matrices $P_2, Q_2 > 0$, $X, Y, V_{a4} \in \mathbb{R}^{n_r \times n}, V_{b4} \in \mathbb{R}^{n_r \times n_r}, V_{c4} \in \mathbb{R}^{n_r \times n}$, and $V_{d4} \in \mathbb{R}^{n_r \times n_d}$ such that the following LMI condition holds:

$$\begin{bmatrix} -Q_2 & V_{a4}^{\mathrm{T}} & \psi_1 & -V_{a4}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \psi_2 & V_{c4} - V_{b4}^{\mathrm{T}} C & V_{d4} - V_{b4}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \star & \psi_3 & \psi_4 \\ \star & \star & \star & \psi_5 \end{bmatrix} < 0, \tag{3.31}$$

where

$$\begin{aligned} \psi_1 &= P_2 + X - V_{a4}^{\mathrm{T}} C, \\ \psi_2 &= I + V_{b4} + V_{b4}^{\mathrm{T}}, \\ \psi_3 &= \bar{\omega}_{\mathrm{d}}^2 Q_2 + \mathrm{He}(-A^{\mathrm{T}} X - C^{\mathrm{T}} V_{c4} + C^{\mathrm{T}} Y), \\ \psi_4 &= -C^{\mathrm{T}} V_{d4} - X^{\mathrm{T}} B_{\mathrm{d}} - V_{c4}^{\mathrm{T}} D_{\mathrm{d}} + Y^{\mathrm{T}} D_{\mathrm{d}}, \\ \psi_5 &= -\gamma_2^2 I - V_{d4}^{\mathrm{T}} D_{\mathrm{d}} - D_{\mathrm{d}}^{\mathrm{T}} V_{d4}. \end{aligned} \tag{3.32}$$

The proof is similar to proof of Theorem 2 and it is omitted to avoid duplication.

3.3.3 Stability Analysis

Observe gain L_{ad} should be designed such that the stability of the closed-loop system is preserved. Hence, the following Theorem should hold.

***Theorem* 5** Considering error state dynamic (3.14), eigenvalues of $\bar{A}_2 = A - L_{\text{ad}}C$ are in the left half plain if the following condition is satisfied for variable matrices X, Y, and $P_0 > 0 \in \mathbb{R}^{n\times n}$:

$$\begin{bmatrix} 0 & P_0 \\ P_0 & 0 \end{bmatrix} < \text{He}\Big(\begin{bmatrix} -X \\ A^{\text{T}}X - C^{\text{T}}Y \end{bmatrix} \times \begin{bmatrix} -qI & pI \end{bmatrix} \Big), \tag{3.33}$$

where $Y = L_{\text{ad}}^{\text{T}}X$, p and q are arbitrary fixed real numbers and $pq < 0$

Proof: Using Lyapunov theory, $\bar{A}_2$ is stable if and only if there exist a $P_0 > 0$ such that

$$\bar{A}_2^{\text{T}}P_0 + P_0\bar{A}_2 < 0, \tag{3.34}$$

which can be written as follows:

$$\begin{bmatrix} \bar{A}_2^{\text{T}} & I \end{bmatrix} \begin{bmatrix} 0 & P_0 \\ P_0 & 0 \end{bmatrix} \begin{bmatrix} \bar{A}_2 \\ I \end{bmatrix} < 0. \tag{3.35}$$

Also, since $pq + qp < 0$,

$$\begin{bmatrix} pI & qI \end{bmatrix} \begin{bmatrix} 0 & P_0 \\ P_0 & 0 \end{bmatrix} \begin{bmatrix} pI & qI \end{bmatrix}^{\text{T}} = (qp + pq)P_0 < 0. \tag{3.36}$$

It should be noted that $\begin{bmatrix} \bar{A}_2^{\mathrm{T}} & I \end{bmatrix}$ is null space of $\begin{bmatrix} -I & \bar{A}_2^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ and $\begin{bmatrix} pI & qI \end{bmatrix}$ is null space of $\begin{bmatrix} -qI & pI \end{bmatrix}$. Hence, using projection lemma,

$$\begin{bmatrix} 0 & P_0 \\ P_0 & 0 \end{bmatrix} < \mathrm{He}\Big(\begin{bmatrix} -I \\ \bar{A}_2^{\mathrm{T}} \end{bmatrix} X \begin{bmatrix} -qI & pI \end{bmatrix} \Big), \tag{3.37}$$

which is equivalent to condition (3.33). □

***Theorem* 6** Considering system (3.2), error dynamics and residual (3.7), there exists a cyber-attack detection observer (3.6), such that error equation is stable and the following finite frequency performance indices are satisfied:

$||G_{\mathrm{r_{Lad}a}}(j\omega_{\mathrm{a}})||_{\_} > \beta_2, \quad \forall \quad \omega_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}]$

$||G_{\mathrm{r_{Lad}d}}(j\omega_{\mathrm{d}})||_{\infty} < \gamma_2, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}]$

$||G_{\mathrm{r_{Lad}f}}(j\omega_{\mathrm{f}})||_{\infty} < \gamma_3, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}]$

$||G_{\mathrm{r_{Lad}y_{ref}}}(j\omega_{\mathrm{y}})||_{\infty} < \gamma_4, \quad \forall \quad \omega_{\mathrm{y}} \in [0, \bar{\omega}_{\mathrm{y}}]$

if there exist Hermitian matrices P_i, $i = 0, ..., 4$, $Q_i > 0$, $i = 1, ..., 4$ variable matrices X, Y, V_{ai}, V_{bi}, V_{ci}, and V_{di}, $i = 1, 2, 3, 4$ with appropriate dimension such that

$$\begin{bmatrix} 0 & P_0 \\ P_0 & 0 \end{bmatrix} - \mathrm{He}\Big(\begin{bmatrix} -X \\ A^{\mathrm{T}} - C^{\mathrm{T}} Y \end{bmatrix} \begin{bmatrix} -q_1 I & p_1 I \end{bmatrix} \Big) < 0, \tag{3.38}$$

$$\begin{bmatrix} \phi_1 & V_{a1}^{\mathrm{T}} & \phi_2 & \phi_3 \\ \star & \phi_4 & V_{c1} - V_{b1}^{\mathrm{T}} C & V_{d1} - V_{b1}^{\mathrm{T}} D_{\mathrm{a}} \\ \star & \star & \phi_5 & \phi_6 \\ \star & \star & \star & \phi_7 \end{bmatrix} < 0, \tag{3.39}$$

$$\begin{bmatrix} -Q_2 & V_{a4}^{\mathrm{T}} & \psi_1 & -V_{a4}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \psi_2 & V_{c4} - V_{b4}^{\mathrm{T}} C & V_{d4} - V_{b4}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \star & \psi_3 & \psi_4 \\ \star & \star & \star & \psi_5 \end{bmatrix} < 0, \tag{3.40}$$

$$\begin{bmatrix} -Q_3 & V_{a2}^{\mathrm{T}} & \theta_1 & 0 \\ \star & \theta_2 & V_{c2} - V_{b2}^{\mathrm{T}} C & V_{d2} \\ \star & \star & \theta_3 & -C^{\mathrm{T}} V_{d2} - X^{\mathrm{T}} B_{\mathrm{f}} \\ \star & \star & \star & -\gamma_2^2 I \end{bmatrix} < 0, \tag{3.41}$$

$$\begin{bmatrix} -Q_4 & V_{a3}^{\mathrm{T}} & \rho_1 & 0 \\ \star & \rho_2 & V_{c3} - V_{b3}^{\mathrm{T}} C & V_{d3} \\ \star & \star & \rho_3 & \rho_4 \\ \star & \star & \star & \rho_5 \end{bmatrix} < 0. \tag{3.42}$$

Consequently, the observer gain L_{ad} can be determined by solving the following optimization problem:

$$\begin{aligned} &\min -\beta_2 + \gamma_2 + \gamma_3 + \gamma_4 \\ &s.t. \quad (3.38) - (3.41) \quad \text{holds.} \end{aligned} \tag{3.43}$$

Similarly, Luenberger finite frequency fault detection filter is designed as follows:

Theorem 7 considering system (3.1), error dynamics and residual (3.4), there exist a fault detection observer (3.3), such that error equation is stable and the following finite frequency

performance indices are satisfied:

$||G_{\mathrm{r_{Lfd}f}}(j\omega_{\mathrm{f}})||\_ > \beta_1, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}]$

$||G_{\mathrm{r_{Lfd}d}}(j\omega_{\mathrm{d}})||_\infty < \gamma_1, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}]$

if there exist Hermitian matrices P_i, $i = 5,6,7$, and $Q_i > 0$, $i = 5,6$, variable matrices X_1, Y_1, V_{ai}, V_{bi}, V_{ci}, and V_{di}, $i = 5,6$ with appropriate dimension such that

$$\begin{bmatrix} 0 & P_7 \\ P_7 & 0 \end{bmatrix} - \mathrm{He}\left(\begin{bmatrix} -X_1 \\ A_{\mathrm{op}}^{\mathrm{T}} - C^{\mathrm{T}} Y_1 \end{bmatrix} \begin{bmatrix} -qI & pI \end{bmatrix} \right) < 0, \tag{3.44}$$

$$\begin{bmatrix} \phi_1 & V_{a5}^{\mathrm{T}} & \phi_2 & \phi_3 \\ \star & \phi_4 & V_{c5} - V_{b5}^{\mathrm{T}} C & V_{d5} - V_{b5}^{\mathrm{T}} D_{\mathrm{f}} \\ \star & \star & \phi_5 & \phi_6 \\ \star & \star & \star & \phi_7 \end{bmatrix} < 0, \tag{3.45}$$

$$\begin{bmatrix} -Q_6 & V_{a6}^{\mathrm{T}} & \psi_1 & -V_{a6}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \psi_2 & V_{c6} - V_{b6}^{\mathrm{T}} C & V_{d6} - V_{b6}^{\mathrm{T}} D_{\mathrm{d}} \\ \star & \star & \psi_3 & \psi_4 \\ \star & \star & \star & \psi_5 \end{bmatrix} < 0, \tag{3.46}$$

where $Y_1 = L_{\mathrm{fd}}^{\mathrm{T}} X_1$,

$$
\begin{aligned}
\phi_1 &= Q_5 + X + X^{\mathrm{T}},\\
\phi_2 &= P_5 + X - X^{\mathrm{T}} A_{\mathrm{op}} - V_{a5}^{\mathrm{T}} C + Y^{\mathrm{T}} C,\\
\phi_3 &= -X B_{\mathrm{f_{op}}} - X^{\mathrm{T}} B_{\mathrm{f_{op}}} - V_{a5}^{\mathrm{T}} D_{\mathrm{f}} + Y^{\mathrm{T}} D_{\mathrm{f}},\\
\phi_4 &= -I + V_{b5} + V_{b5}^{\mathrm{T}},\\
\phi_5 &= \bar{\omega}_{\mathrm{f}}^2 Q_5 + \mathrm{He}(-A_{\mathrm{op}}^{\mathrm{T}} X - C^{\mathrm{T}} V_{c5} + C^{\mathrm{T}} Y),\\
\phi_6 &= A_{\mathrm{op}}^{\mathrm{T}} X B_{\mathrm{f_{op}}} - C^{\mathrm{T}} V_{d5} - C^{\mathrm{T}} Y B_{\mathrm{f_{op}}} - X^{\mathrm{T}} B_{\mathrm{f_{op}}} - V_{c5}^{\mathrm{T}} D_{\mathrm{f}} + Y^{\mathrm{T}} D_{\mathrm{f}},\\
\phi_7 &= \beta_1^2 I + \mathrm{He}(B_{\mathrm{f_{op}}}^{\mathrm{T}} X B_{\mathrm{f_{op}}} - D_{\mathrm{f}}^{\mathrm{T}} V_{d5} - D_{\mathrm{f}}^{\mathrm{T}} Y B_{\mathrm{f_{op}}}),
\end{aligned}
\tag{3.47}
$$

and

$$
\begin{aligned}
\psi_1 &= P_2 + X - V_{a4}^{\mathrm{T}} C,\\
\psi_2 &= I + V_{b4} + V_{b4}^{\mathrm{T}},\\
\psi_3 &= \bar{\omega}_{\mathrm{d}}^2 Q_2 + \mathrm{He}(-A^{\mathrm{T}} X - C^{\mathrm{T}} V_{c4} + C^{\mathrm{T}} Y),\\
\psi_4 &= -C^{\mathrm{T}} V_{d4} - X^{\mathrm{T}} B_{\mathrm{f_{op}}} - V_{c4}^{\mathrm{T}} D_{\mathrm{d}} + Y^{\mathrm{T}} D_{\mathrm{d}},\\
\psi_5 &= -\gamma_1^2 I - V_{d4}^{\mathrm{T}} D_{\mathrm{d}} - D_{\mathrm{d}}^{\mathrm{T}} V_{d4}.
\end{aligned}
\tag{3.48}
$$

Consequently, the observer gain L_{fd} can be determined by solving the following optimization problem:

$$
\begin{aligned}
&\min -\beta_1 + \gamma_1\\
&s.t. \quad (3.44) - (3.46) \quad \text{holds.}
\end{aligned}
\tag{3.49}
$$

It should be noted that since the fault detection observer is designed locally, it can use the input as shown in the filter model (3.3) and the effect of reference signal and cyber-attack does

not need to be attenuated on the residual signal corresponding to the fault detector.

To sum up, the detailed mathematical derivation for designing two Luenberger observer gain namely L_{ad} and L_{fd} is provided in this section to detect cyber-attack and fault, respectively. In particular, the cyber-attack detection observer gain L_{ad} is designed in the monitoring station to generate a residual signal which is sensitive to the presence of cyber-attack while the effect of fault, disturbance, and reference signal on the residual signal is attenuated. Similarly, the residual signal corresponding to fault detection observer triggers in the presence of fault regardless of presence of cyber-attack, disturbance, and reference signal. Gains L_{ad} and L_{fd} are designed by using $\mathscr{H}_{\_}$ and $\mathscr{H}_{\infty}$ formulation. The efficiency of the proposed method is validated by simulation in Section 3.6.

3.4 Normal Observer Design

Another type of filter is designed for cyber-attack detection in this section to solve the problems defined in Section 3.2.2. In order to design a cyber-attack detection filter such that the conditions i-iv in (3.12) are met, respectively the following theorem is proposed.

***Theorem* 8** The low frequency cyber-attack detection filter satisfies the performance indices i-iv in (3.12) if there exist $n \times n$ symmetric matrices P_a, P_b, P_c, P_d, Q_a, Q_b, Q_c, and Q_d, and matrices $\tilde{A}_{\text{ad}} \in \mathbb{R}^{n\times n}$, $\tilde{B}_{\text{ad}} \in \mathbb{R}^{n\times n_y}$, $\hat{C}_{\text{ad}} \in \mathbb{R}^{n_r\times n}$,$D_{\text{ad}} \in \mathbb{R}^{n_r\times n_y}$, $M \in \mathbb{R}^{n\times n}$, $N \in \mathbb{R}^{n\times n}$, and

$Y \in \mathbb{R}^{n\times n}$, such that the following LMI conditions hold:

$$\begin{bmatrix} -Q_{a_{11}} & -Q_{a_{12}} & P_{a_{11}} - Y^{\mathrm{T}} & P_{a_{12}} - M^{\mathrm{T}} & 0 & a_{16} \\ \star & -Q_{a_{22}} & P_{a_{21}} - N^{\mathrm{T}} & P_{a_{22}} - N^{\mathrm{T}} & 0 & a_{26} \\ \star & \star & a_{33} & a_{34} & a_{35} & a_{36} \\ \star & \star & \star & a_{44} & a_{45} & a_{46} \\ \star & \star & \star & \star & -2I & a_{56} \\ \star & \star & \star & \star & \star & a_{66} \end{bmatrix} < 0, \tag{3.50}$$

$$\begin{bmatrix} -Q_{b_{11}} & -Q_{b_{12}} & P_{b_{11}} - Y^{\mathrm{T}} & P_{b_{12}} - M^{\mathrm{T}} & 0 & 0 \\ \star & -Q_{b_{22}} & P_{b_{21}} - N^{\mathrm{T}} & P_{b_{22}} - N^{\mathrm{T}} & 0 & 0 \\ \star & \star & b_{33} & b_{34} & b_{35} & b_{36} \\ \star & \star & \star & b_{44} & b_{45} & b_{46} \\ \star & \star & \star & \star & -\gamma_5^2 I & b_{56} \\ \star & \star & \star & \star & \star & -I \end{bmatrix} < 0, \tag{3.51}$$

$$\begin{bmatrix} -Q_{c_{11}} & -Q_{c_{12}} & P_{c_{11}} - Y^{\mathrm{T}} & P_{c_{12}} - M^{\mathrm{T}} & 0 & 0 \\ \star & -Q_{c_{22}} & P_{c_{21}} - N^{\mathrm{T}} & P_{c_{22}} - N^{\mathrm{T}} & 0 & 0 \\ \star & \star & c_{33} & c_{34} & c_{35} & c_{36} \\ \star & \star & \star & c_{44} & c_{45} & c_{46} \\ \star & \star & \star & \star & -\gamma_6^2 I & c_{56} \\ \star & \star & \star & \star & \star & -I \end{bmatrix} < 0, \tag{3.52}$$

$$\begin{bmatrix} -Q_{d_{11}} & -Q_{d_{12}} & P_{d_{11}} - Y^{\mathrm{T}} & P_{d_{12}} - M^{\mathrm{T}} & 0 & 0 \\ \star & -Q_{d_{22}} & P_{d_{21}} - N^{\mathrm{T}} & P_{d_{22}} - N^{\mathrm{T}} & 0 & 0 \\ \star & \star & d_{33} & d_{34} & d_{35} & d_{36} \\ \star & \star & \star & d_{44} & d_{45} & d_{46} \\ \star & \star & \star & \star & -\gamma_7^2 I & 0 \\ \star & \star & \star & \star & \star & -I \end{bmatrix} < 0. \tag{3.53}$$

For LMI (3.50),

$$a_{16} = -Y^{\mathrm{T}} \upsilon,$$

$$a_{26} = -N^{\mathrm{T}} \upsilon,$$

$$a_{33} = \bar{\omega}_{\mathrm{f}}^2 Q_{a_{11}} + He(YA + \tilde{B}_{\mathrm{ad}} C + L_{11} D_{\mathrm{ad}} C),$$

$$a_{34} = \bar{\omega}_{\mathrm{a}}^2 Q_{a_{12}} + \tilde{A}_{\mathrm{ad}} + A^{\mathrm{T}} M^{\mathrm{T}} + C^{\mathrm{T}} \tilde{B}_{\mathrm{ad}}^{\mathrm{T}} + L_{11} \hat{C}_{\mathrm{ad}} + C^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}} L_{12}^{\mathrm{T}},$$

$$a_{35} = -L_{11} + 0.5 C^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}},$$

$$a_{36} = (A^{\mathrm{T}} Y^{\mathrm{T}} + C^{\mathrm{T}} \tilde{B}_{\mathrm{ad}}^{\mathrm{T}}) \upsilon + Y B_{\mathrm{a}} + C^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}} L_2^{\mathrm{T}},$$

$$a_{44} = \bar{\omega}_{\mathrm{a}}^2 Q_{a_{22}} + He(\tilde{A}_{\mathrm{ad}} + L_{12} \hat{C}_{\mathrm{ad}}),$$

$$a_{45} = -L_{12} + 0.5 \hat{C}_{\mathrm{fd}}^{\mathrm{T}},$$

$$a_{46} = \tilde{A}_{\mathrm{ad}}^{\mathrm{T}} \upsilon + M B_{\mathrm{a}} + \hat{C}_{\mathrm{ad}}^{\mathrm{T}} L_2^{\mathrm{T}},$$

$$a_{56} = -L_2^{\mathrm{T}},$$

$$a_{66} = \beta_3^2 I + L_2 D_{\mathrm{a}} + B_{\mathrm{a}}^{\mathrm{T}} Y^{\mathrm{T}} \upsilon + \upsilon^{\mathrm{T}} Y B_{\mathrm{a}}.$$

For LMI (3.51),

$$b_{33} = \bar{\omega}_{\mathrm{f}}^2 Q_{b_{11}} + He(YA + \tilde{B}_{\mathrm{ad}} C),$$

$$b_{34} = \bar{\omega}_{\mathrm{f}}^2 Q_{b_{12}} + \tilde{A}_{\mathrm{ad}} + A^{\mathrm{T}} M^{\mathrm{T}} + C^{\mathrm{T}} \tilde{B}_{\mathrm{ad}}^{\mathrm{T}},$$

$$b_{35} = \tilde{B}_{\mathrm{ad}} D_{\mathrm{f}},$$

$$b_{36} = C^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}},$$

$$b_{44} = \bar{\omega}_{\mathrm{f}}^2 Q_{b_{22}} + He(\tilde{A}_{\mathrm{ad}}),$$

$$b_{45} = \tilde{B}_{\mathrm{ad}} D_{\mathrm{f}},$$

$$b_{46} = \hat{C}_{\mathrm{ad}}^{\mathrm{T}},$$

$$b_{56} = D_{\mathrm{f}}^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}}.$$

For LMI (3.52),

$$c_{33} = \bar{\omega}_{\mathrm{d}}^2 Q_{c_{11}} + He(YA + \tilde{B}_{\mathrm{ad}} C),$$

$$c_{34} = \bar{\omega}_{\mathrm{d}}^2 Q_{c_{12}} + \tilde{A}_{\mathrm{ad}} + A^{\mathrm{T}} M^{\mathrm{T}} + C^{\mathrm{T}} \tilde{B}_{\mathrm{ad}}^{\mathrm{T}},$$

$$c_{35} = YB_{\mathrm{d}} + \tilde{B}_{\mathrm{ad}} D_{\mathrm{d}},$$

$$c_{36} = C^{\mathrm{T}} D_{ad}^{\mathrm{T}},$$

$$c_{44} = \bar{\omega}_{\mathrm{d}}^2 Q_{c_{22}} + He(\tilde{A}_{\mathrm{ad}}),$$

$$c_{45} = MB_{\mathrm{d}} + C^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}},$$

$$c_{46} = \hat{C}_{\mathrm{ad}}^{\mathrm{T}},$$

$$c_{56} = D_{\mathrm{d}}^{\mathrm{T}} D_{\mathrm{ad}}^{\mathrm{T}}.$$

For LMI (3.53)

$$d_{33} = \bar{\omega}_{\mathrm{y}}^2 Q_{d_{11}} + He(YA + \tilde{B}_{\mathrm{ad}} C),$$

$$d_{34} = \bar{\omega}_{\mathrm{y}}^2 Q_{d_{12}} + \tilde{A}_{\mathrm{ad}} + A^{\mathrm{T}} M^{\mathrm{T}} + C^{\mathrm{T}} \tilde{B}_{\mathrm{ad}}^{\mathrm{T}},$$

$$d_{35} = Y B_{\mathrm{c}},$$

$$d_{36} = C^{\mathrm{T}} D_{ad}^{\mathrm{T}},$$

$$d_{44} = \bar{\omega}_{\mathrm{y}}^2 Q_{d_{22}} + He(\tilde{A}_{\mathrm{ad}}),$$

$$d_{45} = M B_{\mathrm{c}},$$

$$d_{46} = \hat{C}_{\mathrm{ad}}^{\mathrm{T}}.$$

Proof: In order to proof LMI (3.50) which is driven to meet condition $||\bar{G}_{\mathrm{r_{ad}a}}(j\omega_{\mathrm{a}})||_{\_} > \beta_3, \quad \forall \quad \omega_{\mathrm{a}} \in [0, \bar{\omega}_{\mathrm{a}}]$, let us start from GKYP lemma for variable matrices Q_1 and P_1,

$$\begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{a}} \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{a}} \\ I & 0 \end{bmatrix} + \begin{bmatrix} \bar{C} & \bar{D}_{\mathrm{a}} \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} \bar{C} & \bar{D}_{\mathrm{a}} \\ 0 & I \end{bmatrix} < 0, \tag{3.54}$$

where $\Xi = \begin{bmatrix} -Q_1 & P_1 \\ P_1 & \bar{\omega}_{\mathrm{a}}^2 Q_1 \end{bmatrix}$ and $\Pi = \begin{bmatrix} -I & 0 \\ 0 & \beta_3^2 \end{bmatrix}$. Let us define

$$\gamma = \begin{bmatrix} -Q_1 & P_1 & 0 & 0 \\ P_1 & \bar{\omega}_{\mathrm{a}}^2 Q_1 & 0 & 0 \\ 0 & 0 & -\bar{C}^{\mathrm{T}} \bar{C} & -\bar{C}^{\mathrm{T}} \bar{D}_{\mathrm{a}} \\ 0 & 0 & -\bar{D}_{\mathrm{a}}^{\mathrm{T}} \bar{C} & \beta_3^2 I - \bar{D}_{\mathrm{a}}^{\mathrm{T}} \bar{D}_{\mathrm{a}} \end{bmatrix}, \tag{3.55}$$

and

$$\Theta = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}^{\mathrm{T}} \gamma \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} -Q_1 & P_1 & 0 \\ P_1 & \bar{\omega}_{\mathrm{a}}^2 Q_1 - \bar{C}^{\mathrm{T}}\bar{C} & -\bar{C}^{\mathrm{T}}\bar{D}_{\mathrm{a}} \\ 0 & -\bar{D}_{\mathrm{a}}^{\mathrm{T}}\bar{C} & \beta_3^2 I - \bar{D}_{\mathrm{a}}^{\mathrm{T}}\bar{D}_{\mathrm{a}} \end{bmatrix}. \tag{3.56}$$

Now, let $N_u = \begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{a}} \\ I & 0 \\ 0 & I \end{bmatrix}$, then $N_u^{\mathrm{T}}\Theta N_u < 0$ because it is equivalent to the left hand side of (3.54). Also, let $N_v = \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix}$, then $N_v^{\mathrm{T}}\Theta N_v = -Q_1 < 0$. According to null space bases calculation, $U = \begin{bmatrix} -I & \bar{A} & \bar{B}_{\mathrm{a}} \end{bmatrix}$ and $V = \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$. Thus, according to projection lemma, it follows that

$$\Theta + \begin{bmatrix} -I \\ \bar{A}^{\mathrm{T}} \\ \bar{B}_{\mathrm{a}}^{\mathrm{T}} \end{bmatrix} F \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix} F \begin{bmatrix} -I & \bar{A} & \bar{B}_{\mathrm{a}} \end{bmatrix} < 0, \tag{3.57}$$

where $F = \begin{bmatrix} X & XV \end{bmatrix}$, X is a non-singular matrix, $V = [v \;\; 0]^{\mathrm{T}}$, and v should be chosen such that (3.57)<0. It is simple to show that equation (3.57) is equivalent to the following condition:

$$\begin{bmatrix} -Q_1 & P_1 - X & -XV \\ \star & \zeta_1 & \zeta_2 \\ \star & \star & \zeta_3 \end{bmatrix} < 0, \tag{3.58}$$

where $\zeta_1 = \bar{\omega}_{\mathrm{a}}^2 Q_1 - \bar{C}^{\mathrm{T}}\bar{C} + He(\bar{A}^{\mathrm{T}}X)$, $\zeta_2 = -\bar{C}^{\mathrm{T}}\bar{D}_{\mathrm{a}} + \bar{A}^{\mathrm{T}}XV + X^{\mathrm{T}}\bar{B}_{\mathrm{a}}$, and $\zeta_3 = \beta_3^2 I - \bar{D}_{\mathrm{a}}^{\mathrm{T}}\bar{D}_{\mathrm{a}} + He(\bar{B}_{\mathrm{a}}^{\mathrm{T}}XV)$.

Let us assume that $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ with non-singular X_{22} and X_{21} and $J = \begin{bmatrix} I & 0 \\ 0 & X_{22}^{-1}X_{21} \end{bmatrix}$. Using Congruence transformation $\mathrm{diag}(J, J, I)$ for inequality (3.58), it can be written as follows:

$$\begin{bmatrix} -Q_a & P_a - X_{\mathrm{a}} & -X_{\mathrm{a}}V \\ \star & \zeta_4 & \zeta_5 \\ \star & \star & \zeta_6 \end{bmatrix} < 0, \tag{3.59}$$

where $\zeta_4 = \bar{\omega}_{\mathrm{a}}^2 Q_a - \bar{C}_{\mathrm{a}}^{\mathrm{T}} \bar{C}_{\mathrm{a}} + He(\bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}})$, $\zeta_5 = -\bar{C}_{\mathrm{a}}^{\mathrm{T}} \bar{D}_{\mathrm{aa}} + \bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}} V + X_{\mathrm{a}}^{\mathrm{T}} \bar{B}_{\mathrm{aa}}$, and $\zeta_6 = \beta_3^2 I - \bar{D}_{\mathrm{aa}}^{\mathrm{T}} \bar{D}_{\mathrm{aa}} + He(\bar{B}_{\mathrm{aa}}^{\mathrm{T}} X_{\mathrm{a}} V)$ and

$$X_{\mathrm{a}} = J^{\mathrm{T}} X J = \begin{bmatrix} Y^{\mathrm{T}} & M^{\mathrm{T}} \\ N^{\mathrm{T}} & N^{\mathrm{T}} \end{bmatrix},$$

$$Q_a = J^{\mathrm{T}} Q_1 J, \quad P_a = J^{\mathrm{T}} P_1 J,$$

$$\bar{A}_{\mathrm{a}} = J^{-1} \bar{A} J = \begin{bmatrix} A & 0 \\ \hat{B}_{\mathrm{ad}} C & \hat{A}_{\mathrm{ad}} \end{bmatrix} = \begin{bmatrix} A & 0 \\ X_{21}^{-1} X_{22} B_{\mathrm{ad}} C & X_{21}^{-1} X_{22} A_{\mathrm{ad}} X_{22}^{-1} X_{21} \end{bmatrix},$$

$$\bar{B}_{\mathrm{aa}} = J^{-1} \bar{B}_{\mathrm{a}} = \begin{bmatrix} 0 \\ \hat{B}_{\mathrm{ad}} D_{\mathrm{a}} \end{bmatrix},$$

$$\bar{C}_{\mathrm{a}} = \bar{C} J = \begin{bmatrix} D_{\mathrm{ad}} C & \hat{C}_{\mathrm{ad}} \end{bmatrix} = \begin{bmatrix} D_{\mathrm{ad}} C & C_{\mathrm{ad}} X_{22}^{-1} X_{21} \end{bmatrix}$$

$$\bar{D}_{\mathrm{aa}} = \bar{D}_{\mathrm{a}}.$$

In order to eliminate nonlinearities in inequality (3.59), it is converted to LMI using Finsler's lemma. equation (3.59) can be written as follows:

$$G^T P G < 0, \tag{3.60}$$

where

$$G = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & \bar{C}_{\mathrm{a}} & \bar{D}_{\mathrm{aa}} \\ 0 & 0 & I \end{bmatrix}, \tag{3.61}$$

and

$$P = \begin{bmatrix} -Q_a & P_a - X_{\mathrm{a}} & 0 & -X_{\mathrm{a}} V \\ \star & P_{22} & -\frac{\bar{C}_{\mathrm{a}}^{\mathrm{T}}}{2} & P_{24} \\ \star & \star & 0 & -\frac{\bar{D}_{\mathrm{aa}}^{\mathrm{T}}}{2} \\ \star & \star & \star & \beta_3^2 I \end{bmatrix}, \tag{3.62}$$

where $P_{22} = \bar{\omega}_{\mathrm{a}}^2 Q_a + He(\bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}})$ and $P_{24} = \bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}} V + X_{\mathrm{a}}^{\mathrm{T}} \bar{B}_{\mathrm{aa}}$.

According to definition of negative definite matrix,

$$G^{\mathrm{T}} P G < 0 \equiv \zeta^{\mathrm{T}} P \zeta < 0; \;\; \zeta = G\mu, \tag{3.63}$$

where μ is a non-zero vector. It is obvious that G is a full rank matrix and $HG = 0$ where

$$H = \begin{bmatrix} 0 & \bar{C}_{\mathrm{a}} & -I & \bar{D}_{\mathrm{aa}} \end{bmatrix}, \tag{3.64}$$

Hence, according to Finsler's lemma, (3.63) holds if and only if there exist a matrix L such that:

$$P + LH + H^{\mathrm{T}} L^{\mathrm{T}} < 0. \tag{3.65}$$

Matrix $L = \begin{bmatrix} 0 & L_1^{\mathrm{T}} & I & L_2^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, L_1 and L_2 should be chosen such that (3.65) holds.

Finally, (3.65) can be written as follows:

$$\Phi_a = \begin{bmatrix} -Q_a & P_a - X_{\mathrm{a}} & 0 & -X_{\mathrm{a}}V \\ \star & \Phi_{a_{22}} & \Phi_{a_{23}} & \Phi_{a_{24}} \\ \star & \star & -2I & \Phi_{a_{34}} \\ \star & \star & \star & \Phi_{a_{44}} \end{bmatrix} < 0, \tag{3.66}$$

where

$$\begin{aligned} \Phi_{a_{22}} &= \bar{\omega}_{\mathrm{a}}^2 Q_a + He(\bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}} + L_1 \bar{C}_{\mathrm{a}}), \\ \Phi_{a_{23}} &= -L_a + \frac{\bar{C}_{\mathrm{a}}^{\mathrm{T}}}{2}, \\ \Phi_{a_{24}} &= \bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}} V + X_{\mathrm{a}}^{\mathrm{T}} \bar{B}_{\mathrm{aa}} + L_1 \bar{D}_{\mathrm{aa}} + \bar{C}_{\mathrm{a}}^{\mathrm{T}} L_2^{\mathrm{T}}, \\ \Phi_{a_{34}} &= -L_2^{\mathrm{T}} + \frac{\bar{D}_{\mathrm{aa}}^{\mathrm{T}}}{2}, \\ \Phi_{a_{44}} &= \beta_3^2 I + He(\bar{B}_{\mathrm{aa}}^{\mathrm{T}} X_{\mathrm{a}} V + L_2 \bar{D}_{\mathrm{aa}}). \end{aligned}$$

It should be noted that to eliminate nonlinearity, $\tilde{A}_{\mathrm{ad}} = N\bar{A}_{\mathrm{fad}}$ and $\tilde{B}_{\mathrm{ad}} = N\bar{B}_{\mathrm{ad}}$. It is worth mentioning that the LMI (3.50) is the expanded version of LMI (3.66).

To proof inequality (3.51) which is corresponding to condition $||\bar{G}_{\mathrm{r_{ad}f}}(j\omega_{\mathrm{f}})||_\infty < \gamma_6, \quad \forall \quad \omega_{\mathrm{f}} \in [0, \bar{\omega}_{\mathrm{f}}]$, according to GKYP lemma for variable matrices Q_2 and P_2, the following condition holds;

$$\begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{f}} \\ I & 0 \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{f}} \\ I & 0 \end{bmatrix} + \begin{bmatrix} \bar{C} & \bar{D}_{\mathrm{f}} \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Pi \begin{bmatrix} \bar{C} & \bar{D}_{\mathrm{f}} \\ 0 & I \end{bmatrix} < 0, \tag{3.67}$$

where $\Xi_1 = \begin{bmatrix} -Q_2 & P_2 \\ P_2 & \bar{\omega}_{\mathrm{f}}^2 Q_2 \end{bmatrix}$ and $\Pi_1 = \begin{bmatrix} I & 0 \\ 0 & -\gamma_5^2 \end{bmatrix}$. Let us define

$$\bar{\gamma} = \begin{bmatrix} -Q_2 & P_2 & 0 & 0 \\ P_2 & \bar{\omega}_{\mathrm{f}}^2 Q_2 & 0 & 0 \\ 0 & 0 & \bar{C}^{\mathrm{T}}\bar{C} & \bar{C}^{\mathrm{T}}\bar{D}_{\mathrm{f}} \\ 0 & 0 & \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{C} & -\gamma_5^2 I + \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{D}_{\mathrm{f}} \end{bmatrix}, \tag{3.68}$$

and

$$\Theta_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}^{\mathrm{T}} \bar{\gamma} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} -Q_2 & P_2 & 0 \\ P_2 & \bar{\omega}_{\mathrm{f}}^2 Q_2 + \bar{C}^{\mathrm{T}}\bar{C} & \bar{C}^{\mathrm{T}}\bar{D}_{\mathrm{f}} \\ 0 & \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{C} & -\gamma_5^2 I + \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{D}_{\mathrm{f}} \end{bmatrix}, \tag{3.69}$$

Similar to previous case, let $N_u = \begin{bmatrix} \bar{A} & \bar{B}_{\mathrm{f}} \\ I & 0 \\ 0 & I \end{bmatrix}$, then $N_u^{\mathrm{T}}\Theta N_u < 0$ because it is equivalent to the left hand side of (3.67). Also, let $N_v = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$, then

$$N_v^{\mathrm{T}}\Theta N_v = \begin{bmatrix} -Q_2 & 0 \\ 0 & -\gamma_5^2 I + \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{D}_{\mathrm{f}} \end{bmatrix} < 0.$$

According to null space bases calculation, $U = \begin{bmatrix} -I & \bar{A} & \bar{B}_{\mathrm{f}} \end{bmatrix}$ and $V = \begin{bmatrix} 0 & I & 0 \end{bmatrix}$. Thus, accord-

ing to projection lemma,

$$\Theta + \begin{bmatrix} -I \\ \bar{A}^{\mathrm{T}} \\ \bar{B}_{\mathrm{f}}^{\mathrm{T}} \end{bmatrix} F \begin{bmatrix} 0 & I & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} F \begin{bmatrix} -I & \bar{A} & \bar{B}_{\mathrm{f}} \end{bmatrix} < 0, \tag{3.70}$$

where F is a variable matrix $F = X$. Hence, it can be shown that,

$$\begin{bmatrix} -Q_2 & P_2 - X & 0 \\ \star & \zeta_7 & \zeta_8 \\ \star & \star & \zeta_9 \end{bmatrix} < 0, \tag{3.71}$$

where $\zeta_7 = \bar{\omega}_{\mathrm{f}}^2 Q_2 - \bar{C}^{\mathrm{T}}\bar{C} + He(\bar{A}^{\mathrm{T}}X)$, $\zeta_8 = -\bar{C}^{\mathrm{T}}\bar{D}_{\mathrm{f}} + X^{\mathrm{T}}\bar{B}_{\mathrm{f}}$, and $\zeta_9 = \gamma_5^2 I - \bar{D}_{\mathrm{f}}^{\mathrm{T}}\bar{D}_{\mathrm{f}}$.

Let us assume that $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ with non-singular X_{22} and X_{21} and $J = \begin{bmatrix} I & 0 \\ 0 & X_{22}^{-1}X_{21} \end{bmatrix}$.
Using Congruence transformation diag(J, J, I) for inequality (3.71), it can be written as follows:

$$\begin{bmatrix} -Q_b & P_b - X_{\mathrm{a}} & 0 \\ \star & \zeta_{10} & \zeta_{11} \\ \star & \star & \zeta_{12} \end{bmatrix} < 0, \tag{3.72}$$

where $\zeta_{10} = \bar{\omega}_{\mathrm{f}}^2 Q_b - \bar{C}_{\mathrm{a}}^{\mathrm{T}} \bar{C}_{\mathrm{a}} + He(\bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}})$, $\zeta_{11} = -\bar{C}_{\mathrm{a}}^{\mathrm{T}} \bar{D}_{\mathrm{fa}} + X_{\mathrm{a}}^{\mathrm{T}} \bar{B}_{\mathrm{fa}}$, $\zeta_{12} = \gamma_5^2 I - \bar{D}_{\mathrm{fa}}^{\mathrm{T}} \bar{D}_{\mathrm{fa}}$, and

$$X_{\mathrm{a}} = J^{\mathrm{T}} X J = \begin{bmatrix} Y^{\mathrm{T}} & M^{\mathrm{T}} \\ N^{\mathrm{T}} & N^{\mathrm{T}} \end{bmatrix},$$

$$Q_b = J^{\mathrm{T}} Q_2 J, \quad P_b = J^{\mathrm{T}} P_2 J,$$

$$\bar{A}_{\mathrm{a}} = J^{-1} \bar{A} J = \begin{bmatrix} A & 0 \\ \hat{B}_{\mathrm{ad}} C & \hat{A}_{\mathrm{ad}} \end{bmatrix},$$

$$\bar{B}_{\mathrm{fa}} = J^{-1} \bar{B}_{\mathrm{f}} = \begin{bmatrix} B_{\mathrm{f}} \\ \hat{B}_{\mathrm{ad}} D_{\mathrm{f}} \end{bmatrix},$$

$$\bar{C}_{\mathrm{a}} = \bar{C} J = \begin{bmatrix} D_{\mathrm{ad}} C & \hat{C}_{\mathrm{ad}} \end{bmatrix} = \begin{bmatrix} D_{\mathrm{ad}} C & C_{\mathrm{ad}} X_{22}^{-1} X_{21} \end{bmatrix},$$

$$\bar{D}_{\mathrm{fa}} = \bar{D}_{\mathrm{f}}.$$

Finally, by applying Schur complement lemma, inequality (3.72) can be written as follows:

$$\begin{bmatrix} -Q_b & P_b - X_{\mathrm{a}} & 0 & 0 \\ \star & \omega_{\mathrm{f}}^2 Q_b + He(\bar{A}_{\mathrm{a}}^{\mathrm{T}} X_{\mathrm{a}}) & X_{\mathrm{a}}^{\mathrm{T}} \bar{B}_{\mathrm{fa}} & \bar{C}_{\mathrm{a}}^{\mathrm{T}} \\ \star & \star & \gamma_5^2 I & \bar{D}_{\mathrm{fa}}^{\mathrm{T}} \\ \star & \star & \star & -I \end{bmatrix} < 0, \tag{3.73}$$

where $\tilde{A}_{\mathrm{ad}} = N \hat{A}_{\mathrm{ad}}$ and $\tilde{B}_{\mathrm{fd}} = N \hat{B}_{\mathrm{fd}}$. The LMI (3.51) is the expanded version of (3.73). Accordingly, the designed parameters corresponding to low frequency fault detection filter is as

follows:

$$
\begin{aligned}
A_{\mathrm{ad}} &= DU^{\mathrm{T}}N^{-1}\tilde{A}_{\mathrm{ad}}U^{-\mathrm{T}}D^{-1},\\
B_{\mathrm{ad}} &= DU^{\mathrm{T}}N^{-1}\tilde{B}_{\mathrm{ad}},\\
C_{\mathrm{ad}} &= \hat{C}_{\mathrm{ad}}U^{-\mathrm{T}}D^{-1}.
\end{aligned}
\tag{3.74}
$$

where $N = UDU^{\mathrm{T}}$ represents the singular value decomposition of symmetric matrix N. □

The following optimization problem has been solved in order to find fault detection filter parameters:

$$
\begin{aligned}
&\min -\beta_3 + \gamma_5 + \gamma_6 + \gamma_7\\
&s.t. \quad (3.50) - (3.53) \quad \text{holds.}
\end{aligned}
\tag{3.75}
$$

It should be noted that the proofs of inequalities (3.52) and (3.53) which derived to meet the conditions $||\bar{G}_{\mathrm{r_{ad}d}}(j\omega_{\mathrm{d}})||_\infty < \gamma_5, \quad \forall \quad \omega_{\mathrm{d}} \in [0, \bar{\omega}_{\mathrm{d}}]$ and $||\bar{G}_{\mathrm{r_{ad}y}}(j\omega_{\mathrm{y}})||_\infty < \gamma_7, \quad \forall \quad \omega_{\mathrm{y}} \in [0, \bar{\omega}_{\mathrm{y}}]$, respectively are omitted due to similarity.

3.5 Detection Mechanism

The following evaluation function $J(t)$ is used to evaluate the residual signals.

$$
J(t) = \sqrt{\frac{1}{\Delta t}\int_0^{\Delta t} r^{\mathrm{T}}(t)r(t)dt},
\tag{3.76}
$$

where Δt is a time window and r refers to any residual signal defined in this chapter. A Monte Carlo Simulation is used to design a threshold value such that

$$J_{th} = \sup J(t). \tag{3.77}$$

In particular, the threshold value corresponding to cyber-attack detection observer is calculated by running a Monte Carlo simulation for different values of fault. Similarly, the threshold value corresponding to fault detection observer is estimated using a Monte Carlo simulation for a specific range of cyber-attack values. The range of cyber-attack values are chosen such that the performance of the system is degraded while the fault detection observer is not triggered. This consideration is important because if the fault detection filter trigger in a false alarm in the presence of cyber-attack, the system operator will notice the existence of anomaly and can investigate and compensate in early stage. Hence, it is important for cyber-attacker to remain stealthy to all monitoring systems. In the other words, if the cyber-attacker is aware of fault detection filter, he/she can design the cyber-attack signal in a way that the fault detection filter does not trigger.

In this section, two new filter are designed for cyber-attack detection using $\mathscr{H}_-$ and $\mathscr{H}_\infty$ formulation. The main advantage of these filters compared to Luenberger observers is that they have more design parameters and it is expected that they show more satisfactory results in detection of fault and cyber-attack from each other. In the case of cyber-attack detection filter the parameters A_{ad}, B_{ad}, C_{ad}, and D_{ad} are designed such that the residual signal is sensitive to cyber-attack and the effect of fault and disturbance is attenuated.

3.6 Simulation results

3.6.1 Simulation results- Luenberger observer

The efficiency of the proposed method in Section 3.3 is verified in this section. The continuous-time state space model of VTOL aircraft is used for simulation as follows [94]:

$$\begin{aligned} \dot{x}(t) &= A_{\text{op}}x(t) + Bu(t) + B_{\text{f}}f(t) + B_{\text{d}}d(t), \\ y(t) &= Cx(t) + D_{\text{a}}a(t) + D_{\text{f}}f(t) + D_{\text{d}}d(t) + v(t) \end{aligned} \tag{3.78}$$

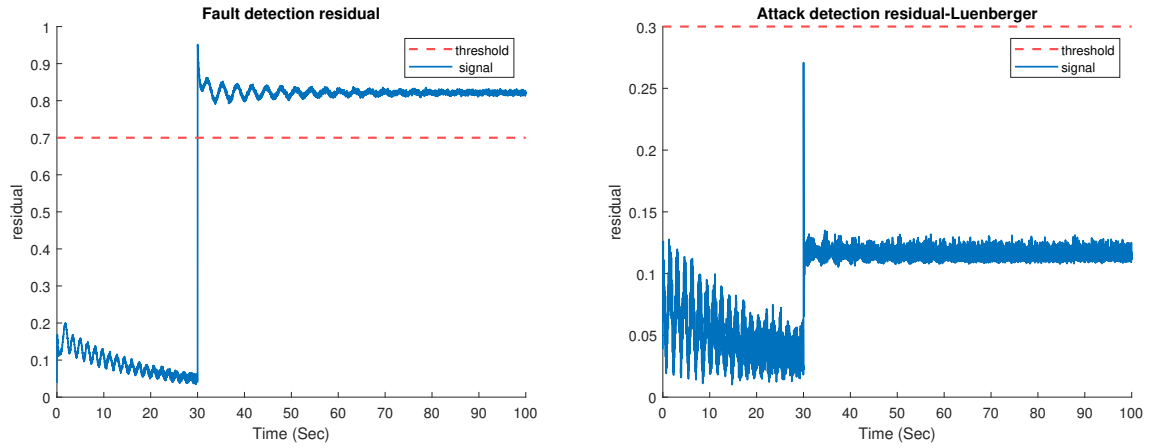
where $x_1(t)$ is the horizontal velocity (knot), $x_2(t)$ is the vertical velocity (knot), $x_3(t)$ is the pitch rate (degree/sec), $x_4(t)$ is the pitch angle (degree), $u_1(t)$ is the collective pitch control, and $u_2(t)$ is the longitudinal pitch control. A random white Gaussian noise $v(t)$ is also added to the sensor measurement. Although the noise was not considered in the mathematical derivations, it is added in the simulation to verify the effectiveness of the proposed method. The closed-loop dynamic is derived using a fixed gain state feedback controller designed in [94; 95]. This dynamics hold for a typical loading and flight condition of the VTOL at the air speed of 135 knot. As mentioned earlier in this chapter, a fault detection Luenberger-based observer is designed locally for the open loop system A_{op} while the cyber-attack detectors are designed for

the closed loop system A.

$$
\begin{aligned}
A_{\text{op}} &= \begin{bmatrix} -0.0366 & 0.0271 & 0.0188 & -0.4555 \\ 0.0482 & -1.01 & 0.0024 & -4.0208 \\ 0.1002 & 0.3681 & -0.707 & 1.420 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\
A &= \begin{bmatrix} -9.9477 & -0.7476 & 0.2632 & 5.0337 \\ 52.1659 & 2.7452 & 5.5532 & -24.4221 \\ 26.0922 & 2.6361 & -4.1975 & -19.2274 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\
B_{\text{f}} &= \begin{bmatrix} 0.4422 & 0.1761 \\ 3.5446 & -7.5922 \\ -5.52 & 4.49 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\
D_{\text{d}} &= \begin{bmatrix} 0 & 0.2 \\ 0 & 0.1 \\ 0.3 & 0 \end{bmatrix}, \quad D_{\text{a}} = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}.
\end{aligned}
\tag{3.79}
$$

Different scenarios are considered to guarantee the efficiency of the proposed method.

The same frequency range for fault and cyber-attack is considered as (0,1) ($\bar{\omega}_{\text{a}} = \bar{\omega}_{\text{f}} = 1$). The disturbance frequency is set to (0,10) ($\bar{\omega}_{\text{d}} = 10$). For $q = -1$ and $p = 1$, the optimization problem (3.43) and (3.49) are solved to calculate observers gain L_{ad} and L_{fd}, respectively. The reference signal y_{ref} is considered to be zero. To verify the effectiveness of the proposed Luenberger observer the following three scenarios are proposed. In all scenarios disturbance is



(a) Fault detection residual (b) Attack detection residual

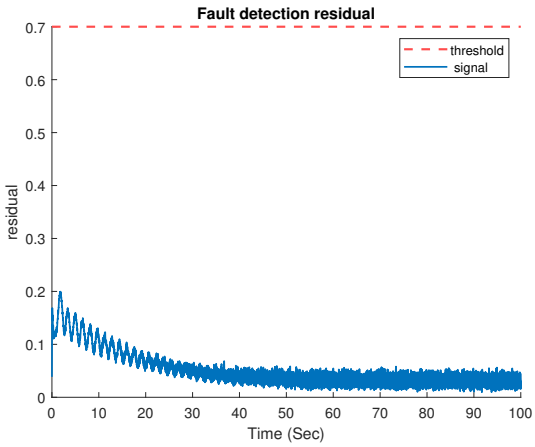
Figure 3.2: Residual signal corresponding to fault (a) and cyber-attack (b) detection filter in the presence of fault.

$d(t) = [\sin(4t)e^{-0.05t} \quad \cos(4t)e^{-0.05t}]^T$.

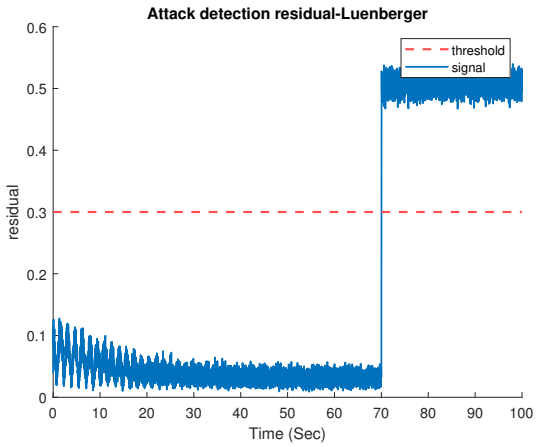
Scenario 1: In the first scenario, a fault signal is a vector of two constants as $f(t) = [1.2, 0.8]$. The fault is occurred at $t \geq 30$ seconds. As shown in Figure 3.2a the Luenberger-based fault detector detects the fault signal. The cyber-attack detection residual signal remains below threshold as shown in Figure 3.2b. However, the residual signal had a big jump at 30 seconds in which the fault has been injected.

Scenario 2: In this scenario, a cyber-attack signal $a(t) = 0.5$ is injected at $t = 70$ seconds. The effectiveness of the proposed cyber-attack detection filter is verified in the presence of noise as shown in Figures 3.3a and 3.3b. In particular, the cyber-attack detection residual signal triggered as soon as the cyber-attack is injected while the fault detection residual signal remains below the threshold. The output of the VTOL aircraft in the presence of cyber attack, noise, and disturbances is shown in Figure 3.4 to show the severity of cyber-attack on the system.

Scenario 3: The effectiveness of the proposed method is verified in the presence of concurrent fault and cyber-attack as follows. A fault signal $f(t) = [1.2, 0.8]$ is injected at 30 seconds and a cyber-attack signal $a(t) = 0.5$ is injected at 70 seconds of the simulation time. According to



(a) Fault detection residual



(b) Attack detection residual

Figure 3.3: Residual signal corresponding to fault (a) and cyber-attack (b) detection filter in the presence of cyber-attack.

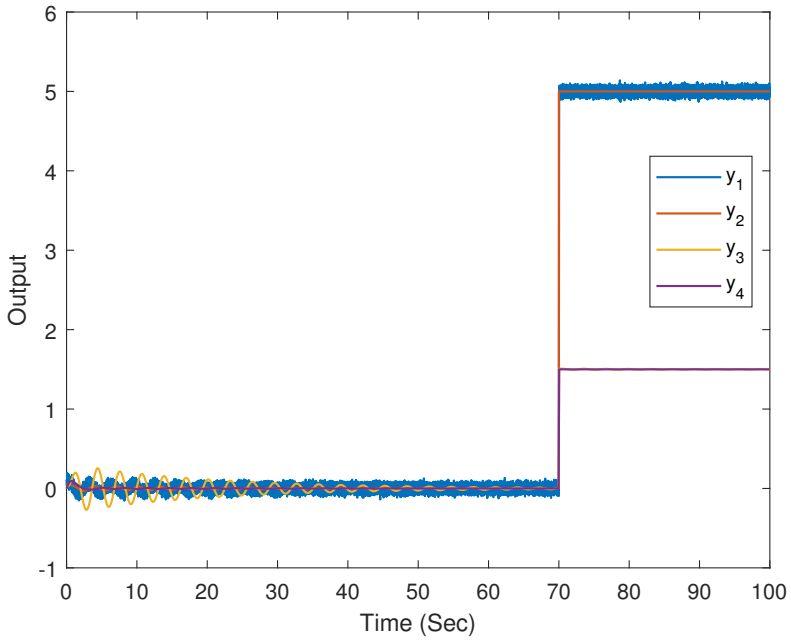
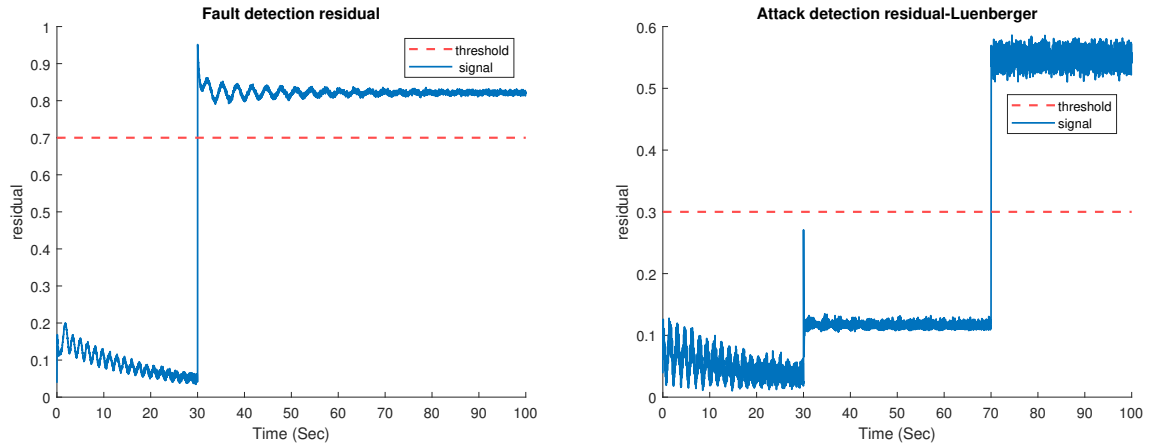


Figure 3.4: Output of VTOL aircraft in the presence of cyber attack, disturbances,and noise.



(a) Fault detection residual

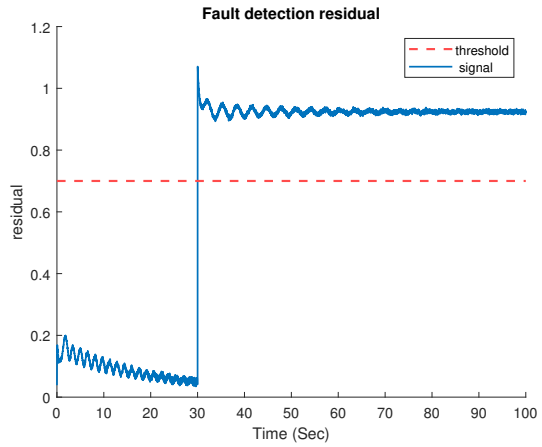
(b) Attack detection residual

Figure 3.5: Residual signal corresponding to fault (a) and cyber-attack (b) detection filter in the presence of fault and cyber-attack.

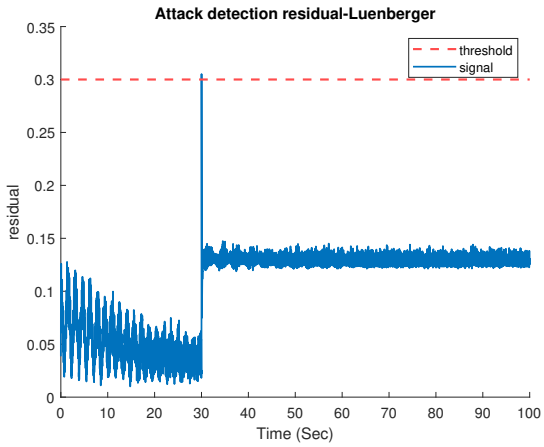
Figure 3.5a, the fault signal is triggered as soon as the fault is injected. Moreover, Figure 3.5b delineates that the residual signal corresponding to Luenberger-based cyber-attack detector remain below threshold until the cyber-attack signal is injected to the system.

It is worth considering that the Luenberger-based cyber-attack detector is prone to false alarms in the presence of fault. In particular, due to the big overshoot that occurs in the residual signal as soon as the fault is injected to the system, the cyber-attack detector may generate a false alarm for few sampling time. For instance, Figure 3.6b shows that in the presence of a fault signal $f(t) = [1.5, 0.8]$ at $t \geq 30$ seconds, the residual signal corresponding to cyber-attack detector showed a false alarm at 30 seconds and the it remains below threshold. It is worth mentioning that fault is detected successfully as shown in Figure 3.6a.

According to the results shown above, it is concluded that fault detection filter is working perfectly. However, in some scenarios depending on the fault signal, the residual signal corresponding to the cyber-attack detector shows a false alarm. Accordingly, new type of observer is designed in which more parameters are required to be designed and hence, it is expected to have more flexibility in detecting cyber-attacks.



(a) Attack detection residual



(b) Attack detection residual

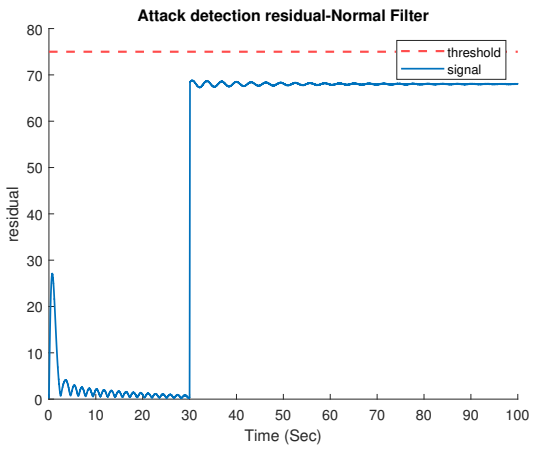
Figure 3.6: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of fault.

3.6.2 Simulation Results- Normal filter

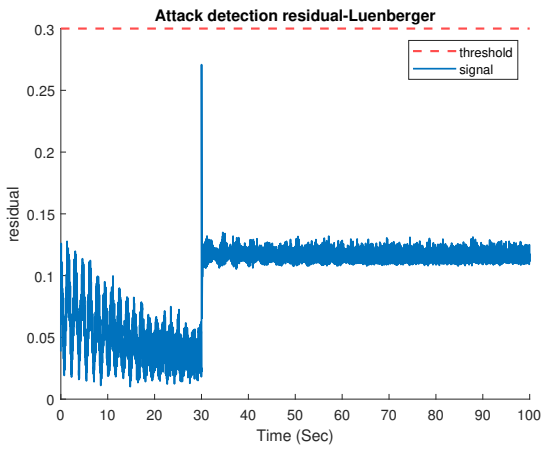
This section presents the efficiency of the designed filter proposed in Section 3.4. The same frequency ranges as in the case of Luenberger observer is considered for fault, cyber-attack, and disturbance. The parameters corresponding to cyber-attack detection filter are given as follows, respectively:

$$v = \begin{bmatrix} 1 & -1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \end{bmatrix}, \quad L_1 = I_{4\times 4}, \quad L_2 = -\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \tag{3.80}$$

The effectiveness of the proposed detector namely normal filter is compared with the Luenberger-based cyber-attack detection observer. Different scenarios are considered to show the efficacy of the proposed filters in the presence of disturbance $d(t) = [\sin(4t)e^{-0.05t} \quad \cos(4t)e^{-0.05t}]^T$. The Figures corresponding to the fault detection observer is omitted in this section because it



(a) Attack detection residual



(b) Attack detection residual

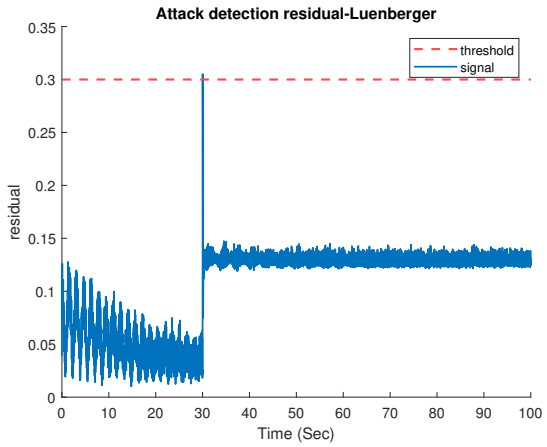
Figure 3.7: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of fault.

has not been changed.

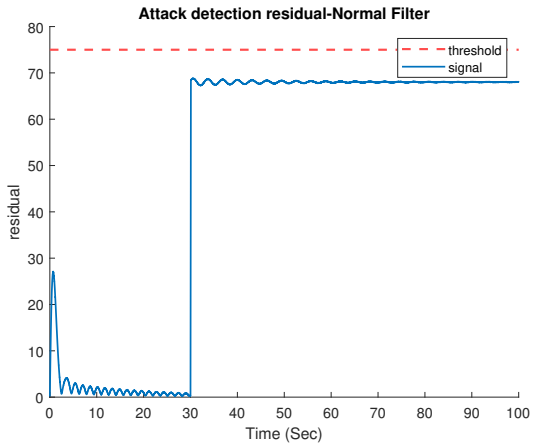
Scenario 1: In the first case, a fault signal is considered as $f(t) = [1.2, 0.8]^{\mathrm{T}}$ in the presence of disturbance and a Gaussian white noise. The fault is occurred in $t \geq 30$ seconds. According to Figure 3.2a, the fault detection filter is working perfectly and is sensitive to the existence of fault. The residual signal corresponding to Luenberger-based cyber-attack detection filter is shown in Figure 3.7b with a huge overshoot at 30 seconds. On the other hand, The residual signal corresponding to the normal filter-based cyber-attack detector is delineated in Figure 3.7a which remains below threshold and performs smooth detection with no overshoot.

Scenario 2: In this scenario a fault signal $f(t) = [1.5, 0.8]^{\mathrm{T}}$ is injected at $t \geq 30$ seconds. It is obvious from Figure 3.8b that the residual signal corresponding to normal filter cyber-attack detector remains below threshold while the Luenberger-based cyber-attack detector shows a false alarm (See Figure 3.8a).

Scenario 3: The effect of cyber-attack signal $a(t) = 0.5$ injected at $t = 70$ seconds is delineated in Figures 3.9a and 3.9b in the presence of disturbance and noise, corresponding to normal filter and Luenberger-based observer, respectively. It is worth mentioning that in the absence



(a) Attack detection residual



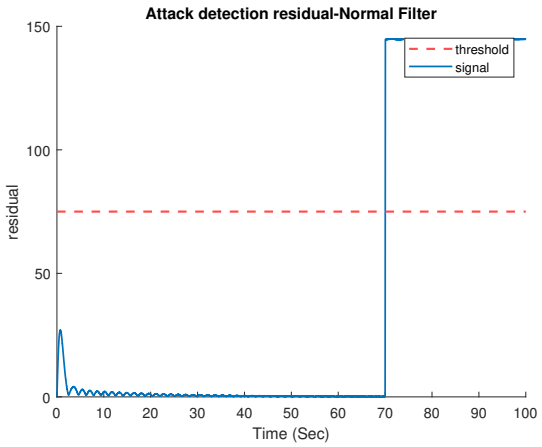
(b) Attack detection residual

Figure 3.8: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of fault.

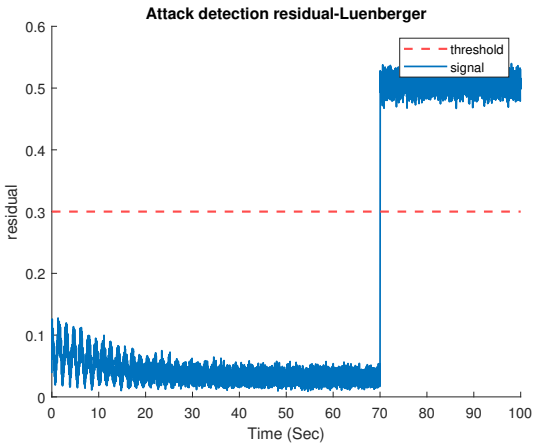
of fault the performance of the both cyber-attacks are similar to each other.

Scenario 4: In this scenario, a fault signal $f(t) = [1.1, 0.9]$ is injected at $t \geq 30$ seconds and -attack signal $a(t) = 0.7$ injected at $t \geq 70$ seconds in the presence of disturbance and noise. Figure 3.10 shows that the fault is detected successfully at $t = 30.015$ seconds. The residual signal corresponding to the normal filter remains below threshold until a cyber-attack is injected at 70 seconds as shown in Figure 3.11a. The Luenberger-based observer is also detecting the cyber-attack with a false alarm (Figure 3.11b).

Scenario 5: In this scenario, a cyber-attack signal $a(t) = 1.2$ is injected at $t \geq 40$ seconds and the fault signal $f(t) = [1.3, 0.5]^T$ is injected at $t \geq 80$ seconds. As shown in Figure 3.12 the fault detection filter detects the fault at $t = 80.01$ seconds. Both Figures 3.13a and 3.13b delineate that the cyber-attack is detected as soon as the cyber-attack signal is injected. Hence the cyber-attack and fault can be detected easily by observing the residuals. By comparing the simulation results of the proposed filters, it is observed that the cyber-attack detection Normal filter-based observer has better performance compared to the Luenberger-based cyber-attack detection filter. Moreover, the performance of the fault detection filter is identical in the pres-



(a) Attack detection residual



(b) Attack detection residual

Figure 3.9: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of cyber-attack.

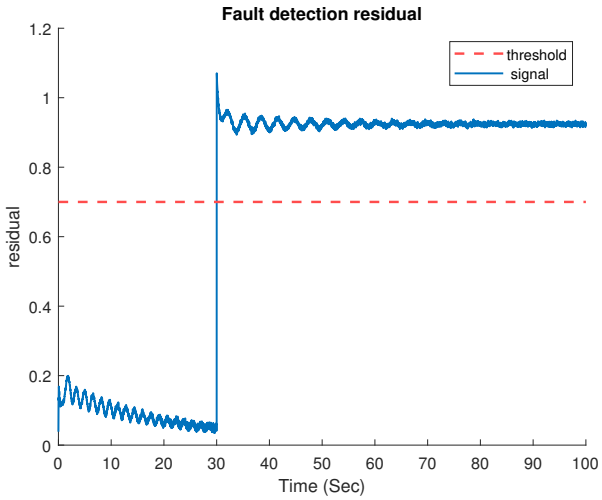
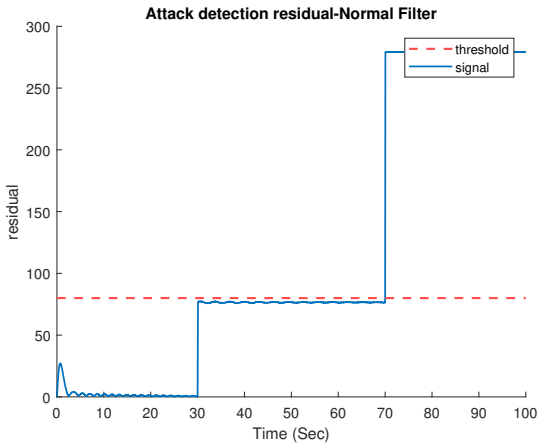
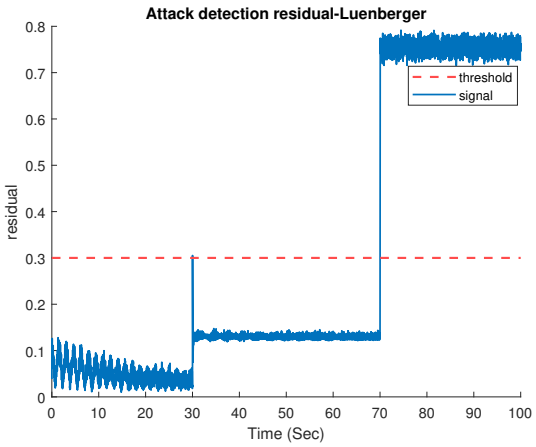


Figure 3.10: Residual signal corresponding to the fault detection filter in the presence of fault and cyber-attack.



(a) Attack detection residual



(b) Attack detection residual

Figure 3.11: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of fault and cyber-attack.

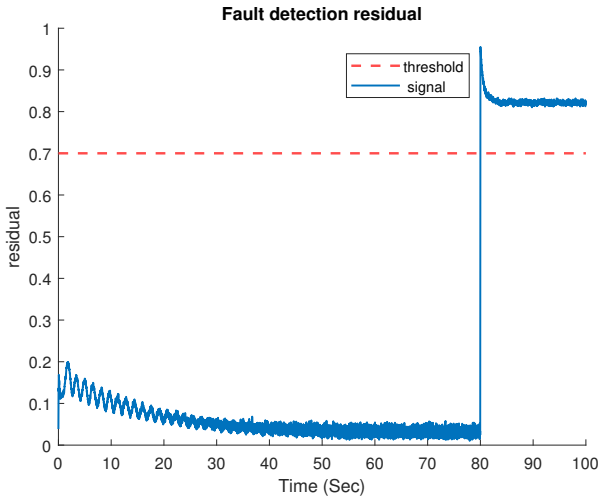
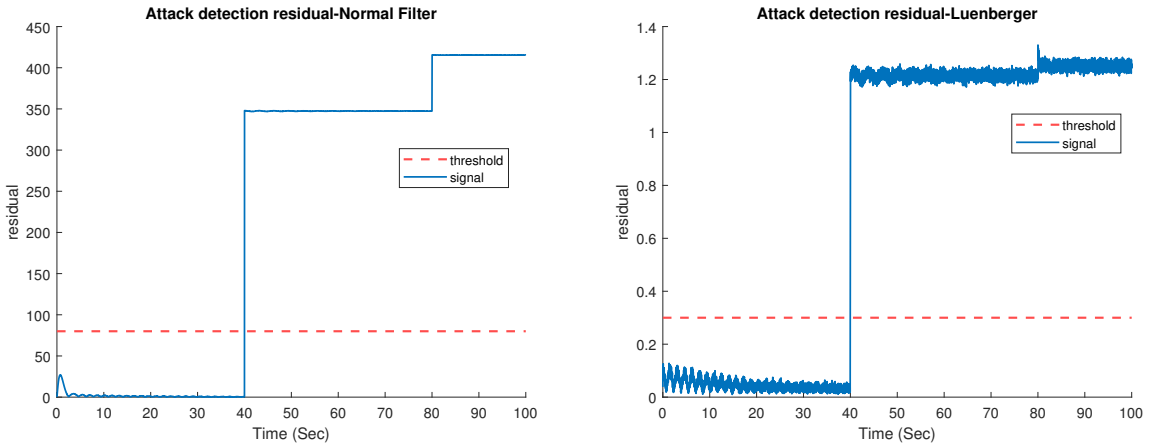


Figure 3.12: Residual signal corresponding to the fault detection filter in the presence of fault and cyber-attack.



(a) Attack detection residual (b) Attack detection residual

Figure 3.13: Residual signal corresponding to normal (a) and Luenberger-based (b) detection filter in the presence of fault and cyber-attack.

ence or absence of cyber-attack as it is designed locally. Accordingly, it is more reliable to use normal filter for cyber-attack detection rather than the Luenberger observer due to its false alarms. In particular, in order to detect faults and cyber-attacks in a cyber-physical system simultaneously, two filters are designed. One filter is designed to detect fault locally and the second is a normal filter-based observer which is designed to detect cyber-attacks.

Scenario 6: Moreover, a fault signal $f(t) = [1.3, 0.8]^{\mathrm{T}}$ is injected at $30 \leq t \leq 70$ seconds and a cyber-attack signal $a(t) = 0.7\sin(0.5t)$ is injected at $t \geq 40$ seconds. The fault detection residual which is designed in the plant side, triggers as soon as the fault is injected as shown in Figure 3.14 . Thus, according to Table 4.1, the fault is detected. The cyber-attack detection residual signal corresponding to Luenberger observer and Normal filter are shown in Figures 3.15a and 3.15b, respectively. It can be seen that the Luenberger observer shows a false alarm in the presence of fault at $t = 30$ seconds. However, the normal filter can detect cyber-attack without any false alarm in the presence of fault.

It is worth mentioning that limiting the frequency ranges in the design procedure does not affect the detection performance. To verify this statement, a cyber-attack signal is injected

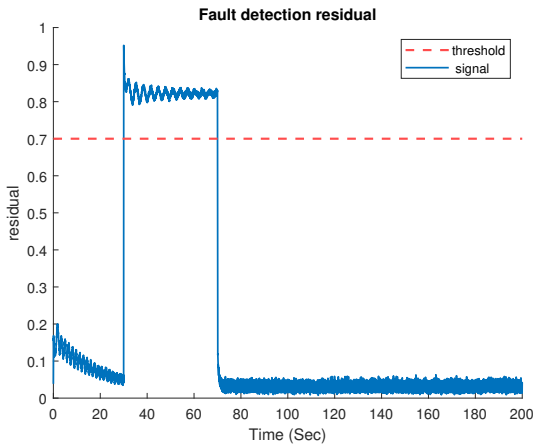
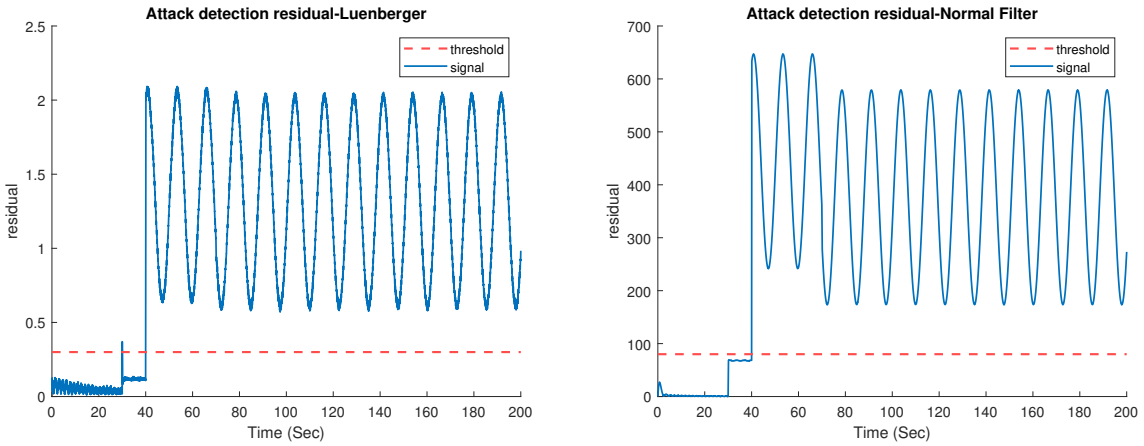
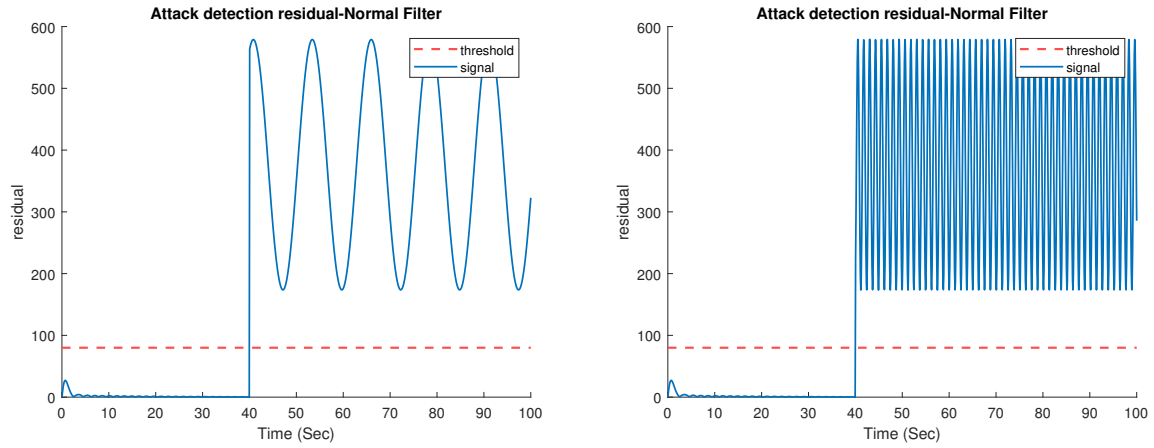


Figure 3.14: Residual signal corresponding to the fault detection filter in the presence of fault, cyber-attack, disturbance, and noise.



(a) Attack detection residual (b) Attack detection residual

Figure 3.15: Residual signal corresponding to cyber-attack detection Luenberger-based observer (a) and normal filter cyber-attack detector (b) in the presence of fault, cyber-attack, disturbance, and noise.



(a) Attack detection residual (b) Attack detection residual

Figure 3.16: Residual signal corresponding to cyber-attack detection filter in the presence of Low frequency cyber-attack (a) and High frequency cyber-attack (b).

using two different frequencies and the corresponding residual signals are shown as follows. Figure 3.16a delineates the residual signal corresponding to cyber-attack detection using normal filter in the presence of cyber-attack signal $a(t) = 0.7\sin(0.5t)$. The frequency of this signal is within the range which is considered in the design procedure ($\bar{\omega}_a = 1$). On the other hand, Figure 3.16b presents the residual signal generated by cyber-attack detection filter (Normal filter) in the presence of cyber-attack signal $a(t) = 0.7\sin(20t)$ which is beyond the assumed range in design procedure. As mentioned earlier, in both scenarios the cyber-attack is detected successfully. This is an advantage of the proposed method in this thesis from defender point of view as the frequency of cyber-attack neither degrade nor enhance the detection performance and detection performance is only depends on the magnitude of the cyber-attack signal. The proposed method can detect any cyber-attack with the magnitude in the range of $[0.1, 1.2]$. Attacks with greater magnitudes result in triggering the residual of both fault and cyber-attack detectors.

To further verify the effectiveness of the proposed detectors in this chapter, a Monte Carlo simulation is conducted. In particular, to show the effectiveness of the fault detector, 100 simu-

lations have been performed with different values of fault and cyber-attack generated randomly. The faults magnitudes are in the range of $([0.1, 1.8]; [0.1, 0.9])$ and cyber-attacks magnitudes are in the range of $(0.1, 1.2)$. Table 3.2 delineates the confusion matrix corresponding to the fault detector. The confusion matrix corresponding to fault detector consists of the following main components:

TP = when fault is detected correctly

TN = when no fault is detected correctly.

FP = when fault is detected wrongly (there is no fault but residual is triggred.)

FN = when fault is not detected correctly (there is fault but not detected.)

Table 3.2: *Confusion Matrix Corresponding to the Fault Detector*

		Expected	
	N=100	Positive	Negative
Detected	Positive	TP = 65	FP = 0
	Negative	FN = 2	TN = 33

According to this table, the detection rate of the Luenberger-based fault detector can be calculated as follows:

$$Dr = \frac{TP}{TP + FN} * 100 = 97.01\%. \tag{3.81}$$

Similarly, two Monte-Carlo simulations are conducted for Luenberger-based cyber-attack detector and the cyber-attack detector based on normal filter and their corresponding Confusion matrices are shown in Tables 3.3 and 3.4, respectively. It is worth considering that the simulations has been done for the same range of cyber-attack and fault. The confusion matrix corresponding to cyber-attack detector consists of the following main components:

TP = when cyber-attack is detected correctly

TN = when no cyber-attack is detected correctly.

FP = when cyber-attack is detected wrongly (there is no cyber-attack but residual is triggered.)

FN = when cyber-attack is not detected correctly (there is cyber-attack but not detected.)

Table 3.3: *Confusion Matrix Corresponding to the Luenberger-Based Cyber-Attack Detector*

		Expected	
	N=100	Positive	Negative
Detected	Positive	TP = 45	FP = 10
	Negative	FN = 4	TN = 31

Table 3.4: *Confusion Matrix Corresponding to the Cyber-Attack Detector Based on Normal Filter*

		Expected	
	N=100	Positive	Negative
Detected	Positive	TP = 68	FP = 2
	Negative	FN = 3	TN = 27

According to Table 3.3, the detection rate of the Luenberger-based cyber-attack detector is 91.84% while the detection rate of normal-based cyber-attack detector is 95.77% according to Table 3.4. Hence, it verifies the better performance of the normal filter as mentioned earlier in this chapter.

Table 3.5 presents extra measures on the performance of each filter in this chapter as follows:

$$precision = PR = \frac{TP}{TP + FP} \tag{3.82}$$

$$Accuracy = ACC = \frac{TP + TN}{P + N} \tag{3.83}$$

where $P = TP + FN$ denote the number of real positive and $N = FP + TN$ denotes the number of real negative cases. Moreover, Table 3.5 shows that the accuracy and precision of the normal

Table 3.5: *Extra Measures on Confusion Matrices*

Filter Name	Table	PR	ACC
Luenberger-based Fault Detector	3.2	100%	98%
Luenberger-based Attack Detector	3.3	81.8%	84.4%
Normal Filter-based Attack Detector	3.4	97.14%	95%

filter in detecting cyber-attack is higher than that of Luenberger-based cyber-attack detector.

3.7 Conclusion

In conclusion, two different types of filters are designed in this thesis to simultaneously detect concurrent fault and cyber-attack in the presence of disturbance and noise in a cyber-physical system. This problem is defined as a multi-objective framework and $\mathscr{H}_\infty/\mathscr{H}_-$ formulation. A set of two Luenberger-based observers and a normal filters is designed and their performances are compared. It is observed that the best performance is obtained by combining two types of filters. In particular, the fault detection is achieved by Luenberger-based observer and normal filter showed the best performance in detecting and isolating cyber-attacks in this thesis however, the performance of designed detection filter depends on the application. The efficiency of the proposed method in simultaneous detection of fault and cyber-attack is verified by simulating a VTOL aircraft in the presence of measurement noise and disturbance.

Chapter 4

Robust Fault and Cyber-Attack Detection in Multi-Agent Systems

The problem of fault and cyber-attack detection has been investigated in presence of disturbance in multi-agent systems. To achieve this goal, a Luenberger-based observer is designed for each agent locally to generate a residual signal that detects the presence of fault on the actuator and sensor in presence of disturbance. The corresponding threshold to the fault detector transmits over a secure channel to the command and control station. It is not a limiting assumption to consider a secure channel for transmitting the residual signal to the command and control because it is a binary signal and can be secured by coding strategies. In addition, a bank of UIOs has been designed for the neighbors of each agent to detect cyber-attack in a simultaneous way in presence of both fault and disturbance. Each UIO generates a residual signal corresponding to a neighboring agent using only the output information of that agent. Similar to the fault detection residual signal, the residual signal corresponding to the cyber-attack detector transmits to command and control over a secure channel. One of the most important challenges in this field is to distinguish the presence of fault and cyber-attack from each other

for the operator in the command and control station. Hence, a decision making algorithm based on the received residuals of each agent is designed in command and control station.

4.1 System Description

Let us consider the following multi-agent linear time-invariant system consisting of N homogeneous agents in presence of actuator and sensor fault and disturbance,

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + Bu_i(t) + B_{\mathrm{f}} f_i(t) + B_{\mathrm{d}} d_i(t) \\ y_i(t) &= Cx_i(t) + D_{\mathrm{f}} f_i(t) + D_{\mathrm{d}} d_i(t) \quad i = 1, 2, ..., N \end{aligned} \tag{4.1}$$

where $x_i(t) \in \mathbb{R}^n$, $u_i(t) \in \mathbb{R}^u$, and $y_i(t) \in \mathbb{R}^y$ denote the state, input and output of the agent i. $f_i(t)$ is fault vector and disturbance is denoted as $d_i(t)$. The matrices A, B, B_{f}, B_{d}, C, D_{d}, and D_{f} are known with appropriate dimensions. The communication topology among the agents is represented as a directed graph and it is assumed that the communication topology has a spanning tree. The exact definition and example of a directed graph is provided in Section 2.4. Agents are interacting with their neighbors using a wireless communication network. It is assumed that only the output measurement of each agent is transmitted to its neighbors. $\mathcal{N}_j$ represents the neighboring set of agent j which consists of agents which send their output information to agent j. Thus, let $y_i^j(t)$ be the output of agent i received by agent j which is prone to cyber-attack,

$$y_i^j(t) = \quad y_i(t) + D_{\mathrm{a}} a_i(t), \;\; i \in \mathcal{N}_j \tag{4.2}$$

where $a_i(t)$ is the cyber-attack vector representing data corruption caused by false data injection on the communication link and D_{a} is known with appropriate dimension. The attack is aimed

to fool the monitoring system and not the instability of the system. Also, it is injected to remain stealthy to the fault detection system. Hence, the main aim of this chapter is to classify anomalies as fault and cyber-attack. To consider a more general case, it is assumed that all the communication links from agent i will be under cyber-attack and not only one of them. For instance, if agent 3 communicates its output to agents 2, 4, and command and control subsystem the cyber-attack will comprise the communication links between all neighbors of agent 3 and not only one of them.

The block diagram of a MAS in presence of fault, cyber-attack and disturbance is given in Figure 4.1. It is worth mentioning that Figure 4.1 presents the components of the system and not the communication topology among the agents. It is realistic to assume a secure channel for transmitting the reference signal as it is not changing every time instant. The reference signal can be a desired value of outputs to be reached by the agents. However, for a consensus control which is considered in this chapter the reference signal is set to be zero. It is worth mentioning that the existence of the reference signal does not affect the filter design in this chapter. Each agent is equipped with a fault detection filter and its corresponding residual signal is transmitted to command and control subsystem over a secure channel. On the other hand, each agent is able to detect the injected cyber-attack on the communication link between itself and its neighbors by generating a residual. The output of the decision making station is observed by the operator to take a proper action in presence of anomalies. The fault detector is a Luenberger observer which is developed in Section 4.3.1 and the UIO is derived in Section 4.3.2.

4.2 Problem Formulation

As discussed earlier, the main aim of this chapter is to detect cyber-attack and fault in presence of disturbances in a multi-agent system which is equipped with an output consensus controller.

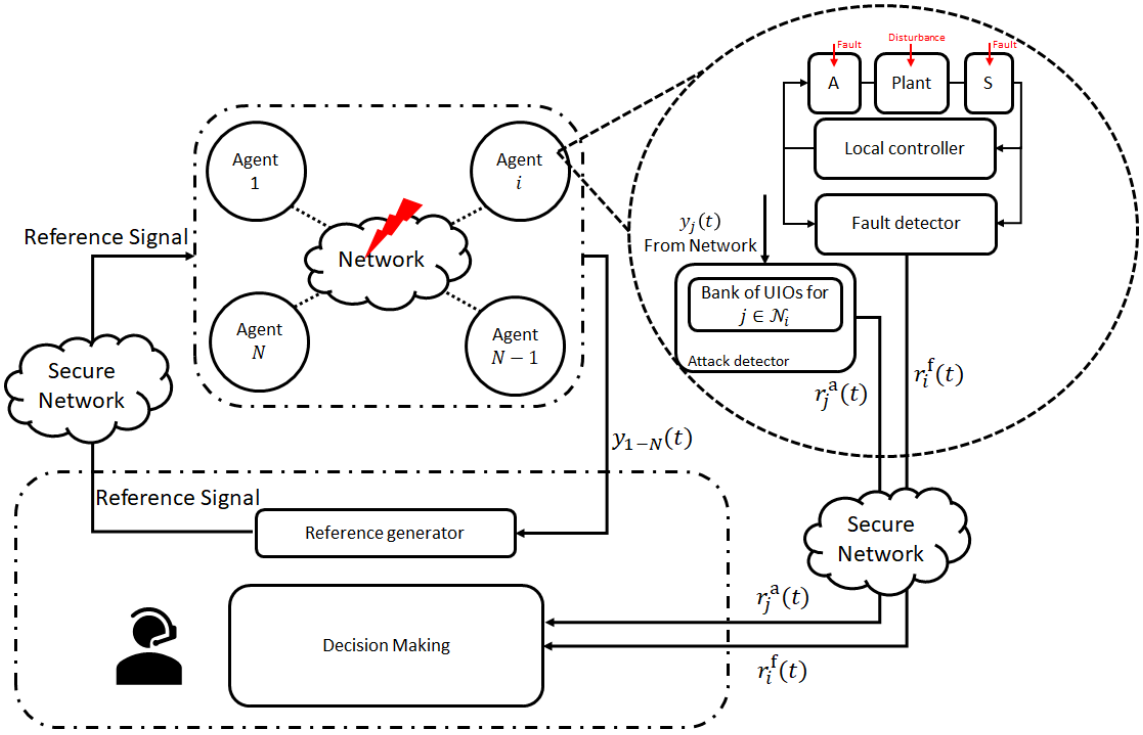


Figure 4.1: MASs Block Diagram with N agents.

The fault detection filter is designed for each agent utilizing multi-objective framework and $\mathscr{H}_-$ and $\mathscr{H}_\infty$ formulation. On the other hand, a bank of unknown input observers is designed for each agent to detect the cyber-attack injected on the communication link between its neighbors. It is worth considering that the Luenberger observer is not able to detect cyber-attack because in the MASs with output consensus controller, the input of each agent depends on the output of its neighbors which is not transmitted over communication network. It is assumed that the output information of each agent is transmitted through a directed graph among the neighboring agents as shown in Figure 4.1. Also, it is assumed that the directed graph has a spanning tree. It should be noted that in the case of undirected graph, the graph should be connected. The problem formulation of Luenberger and unknown input observers are provided in Sections 4.2.1 and 4.2.2, respectively. The details of decision making block is provided in Section 4.2.3.

4.2.1 Luenberger Observer - Fault Detection

A Luenberger observer is designed for fault detection in agent i as follows:

$$\begin{aligned} \dot{\hat{x}}_i(t) &= A\hat{x}_i(t) + Bu_i(t) + L(y_i(t) - \hat{y}_i(t)), \\ \hat{y}_i(t) &= C\hat{x}_i(t), \\ r_i(t) &= y_i(t) - \hat{y}_i(t), \end{aligned} \tag{4.3}$$

where $\hat{x}_i(t) \in \mathbb{R}^n$, $\hat{y}_i(t) \in \mathbb{R}^{n_y}$, and $r_i(t) \in \mathbb{R}^{n_y}$ are state, output, and residual signal of the observer in agent i, respectively and the observer gain L is a design parameter. Let $e_i(t) = x_i(t) - \hat{x}_i(t)$, the state error dynamic can be written as follows:

$$\begin{aligned} \dot{e}_i(t) &= \bar{A}e_i(t) + \bar{B}_\mathrm{f} f_i(t) + \bar{B}_\mathrm{d} d_i(t), \\ r_i(t) &= y_i(t) - \hat{y}_i(t) = Ce_i(t) + D_\mathrm{f} f_i(t) + D_\mathrm{d} d_i(t), \end{aligned} \tag{4.4}$$

where $\bar{A} = A - LC$, $\bar{B}_\mathrm{f} = B_\mathrm{f} - LD_\mathrm{f}$, and $\bar{B}_\mathrm{d} = B_\mathrm{d} - LD_\mathrm{d}$. Accordingly, L should be designed such that the following conditions are met:

i) $A - LC$ is Hurwitz

ii) $||G_\mathrm{rf}(s)||_- > \beta$,

iii) $||G_\mathrm{rd}(s)||_\infty < \gamma$,

where β and γ are design parameters. The transfer matrices G_rf and G_rd that present the effect of the fault and disturbance on the residual signal, respectively, are as follows,

$$\begin{aligned} G_\mathrm{rf}(s) &= C(sI - A + LC)^{-1}(B_\mathrm{f} - LD_\mathrm{f}) + D_\mathrm{f}, \\ G_\mathrm{rd}(s) &= C(sI - A + LC)^{-1}(B_\mathrm{d} - LD_\mathrm{d}) + D_\mathrm{d}. \end{aligned} \tag{4.5}$$

Despite the previous chapter in which the filters have been designed in finite frequency domain, the filter design will be in full frequency domain in this chapter. However, as it is mentioned in Section 2.2, the $\mathscr{H}_-$ is always zero for strictly proper and proper system. Hence, an arbitrary large upper bound is considered for the sake of mathematical derivation of LMI. The LMI corresponding to $\mathscr{H}_\infty$ is derived in full frequency domain i.e. $[0, \infty)$.

4.2.2 Unknown Input Observer- Attack Detection

In order to achieve the goal of cyber-attack detection, a bank of UIOs are designed in this section for each agent. In particular, agent i has N_i number of UIOs which is refered to bank of UIOs to detect the injected FDI cyber-attack on the communication link between itself and its neighbors. The UIOs have identical design parameters as the agents are homogeneous. Hence, bank of UIOs consists of N_i number of identical UIOs in agent i. In order to design an UIO in agent j which is in the neighborhood of agent i, let us rewrite (4.1) as follows:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + E_1\rho_i(t), \\ y_i(t) &= Cx_i(t) + E_2\rho_i(t), \\ y_i^j(t) &= y_i(t) + D_a a_i(t) \quad i \in N_j \end{aligned} \tag{4.6}$$

where $E_1 = [B, B_\mathrm{f}, B_\mathrm{d}]$, $E_2 = [0, D_\mathrm{f}, D_\mathrm{d}]$, and $\rho_i(t) = [u_i^\mathrm{T}(t), f_i^\mathrm{T}(t), d_i^\mathrm{T}(t)]^\mathrm{T}$. The matrix $E \in \mathbb{R}^{n_i \times n_q}$, $n_q \leq n_i$ links the unknown inputs to the dynamics of the system. A transformation matrix T is chosen to nullify the term $E_2\rho_i(t)$ from the output equation (i.e. $TE_2 = 0$) as follows:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + E_1\rho_i(t), \\ y_{E_i}(t) &= C_E x_i(t), \quad i \in N_j \end{aligned} \tag{4.7}$$

where $y_{E_i}(t) = Ty_i(t)$ and $C_E = TC$. The structure of full-order unknown input observer for system (4.7) can be defined as follows [96; 97]:

$$\begin{aligned}
\dot{z}_i^j(t) &= Fz_i^j(t) + Ky_i^j(t), \\
\tilde{x}_i^j(t) &= z_i^j(t) + Hy_i^j(t), \\
\tilde{y}_{E_i}^j(t) &= C_E\tilde{x}_i(t), \\
r_i^j(t) &= y_i^j(t) - T^{-1}\tilde{y}_{E_i}^j(t)
\end{aligned} \tag{4.8}$$

where $z_i^j(t)$ is state of the full order observer corresponding to agent i in agent j, $\tilde{x}_i^j(t)$ and $\tilde{y}_{E_i}^j(t)$ are the estimated state and the estimated output of agent i in agent j, respectively. The matrices F, K, and H are design parameters.

Remark: The proposed UIO in this chapter is designed independent of the topology. Simply, each UIO receives information from only one agent in the neighboring set and estimates the state of that agent specifically. Accordingly, each agent is equipped with bank of UIOs each responsible for a single agent in the neighboring set. This type of UIO is well studied in [63] and [69]. In particular, the design procedure and stability analysis of each UIO is independent of the topology.

4.2.3 Detection and Decision Making

Up to this point, in order to solve the problem of fault and cyber-attack detection in multi-agent systems, two types of observers are presented namely Luenberger observer for fault detection and unknown input observer for cyber-attack detection. The fault detection module is designed in the plant side and produces a residual signal. By utilizing the Monte Carlo approach, a fixed threshold will be calculated and the fault is detected if the residual triggers the calculated

threshold. Similarly, each cyber-attack detector generates a residual signal corresponding to its corresponding agent. A threshold is designed using Monte-Carlo approach. The cyber-attack is detected if the residual triggers the threshold. The decision making is done by the operator according to the generated residual. In particular, the decision making can be done according to Table 4.1 in which r_{fd} is considered as flag 1 if the residual corresponding to the fault detector exceeds the threshold and if the residua remains below threshold it is considered as flag 0. Similar logic applies to the attack detection residual r_{ad}. It is worth mentioning that in the case of concurrent fault and cyber-attack, the operator can still classify these anomalies by observing their corresponding residuals.

Table 4.1: *Decision Making*

r_{fd}	1	0	1
r_{ad}	0	1	1
Decision	Fault	Attack	Fault and Attack

4.3 Main Results

This section presents the detailed mathematical derivation of the proposed detection mechanisms. A Luenberger-based fault detection observer is designed for each agent locally to detect sensor and actuator fault in presence of disturbances. In addition, a bank of UIOs is also designed in each agent to detect the cyber-attack in its neighboring agent.

4.3.1 Luenberger Observer Design - Fault Detection

A Luenberger observer is designed for each agent locally to detect fault on the sensor and actuator in presence of disturbance such that the conditions (i)-(iii) in Section 4.2 are met. Thus, in order to detect the fault in the sensor and actuator of an agent locally, the following

theorem should hold.

***Theorem* 9** Considering system (4.1), error dynamics and residual (4.4), there exists a fault detection observer (4.3), such that the error equation is BIBO stable and the following performance indices are satisfied:

$||G_{rf}(s)||_{-} > \beta$,

$||G_{rd}(s)||_{\infty} < \gamma$,

if there exist Hermitian matrices $X > 0$ and a variable matrix Y with appropriate dimension such that

$$\begin{bmatrix} AX + A^{T}X - Y^{T}C - C^{T}Y + C^{T}C & XB_{f} - Y^{T}D_{f} + C^{T}D_{f} \\ \star & D_{f}^{T}D_{f} - \beta^{2}I \end{bmatrix} < 0 \tag{4.9}$$

$$\begin{bmatrix} A^{T}X + XA - C^{T}Y - Y^{T}C & XB_{d} - Y^{T}D_{d} & C^{T} \\ \star & -\gamma^{2}I & D_{d}^{T} \\ \star & \star & -I \end{bmatrix} < 0 \tag{4.10}$$

where $Y = L^{T}X$.

Proof: According to Remark 2.2 and considering $d(t) = 0$, the $\mathscr{H}_{-}$ performance corresponding to the system (4.4) is as follows:

$$\begin{bmatrix} \bar{A}^{T}X + X\bar{A} + C^{T}C & X\bar{B}_{f} + C^{T}D_{f} \\ \star & D_{f}^{T}D_{f} - \beta^{2}I \end{bmatrix} > 0. \tag{4.11}$$

In order to convert this nonlinear matrix inequality into an LMI, let $Y = L^{T}X$ then (4.11) can be written as (4.9).

According to Lemma 6 and considering $f(t) = 0$, the $\mathscr{H}_\infty$ performance for system (4.4) is as follows:

$$\begin{bmatrix} \bar{A}^{\mathrm{T}}X + X\bar{A} & X\bar{B}_{\mathrm{d}} & C^{\mathrm{T}} \\ \star & -\gamma^2 I & D_{\mathrm{d}}^{\mathrm{T}} \\ \star & \star & -I \end{bmatrix} < 0. \tag{4.12}$$

In order to convert this nonlinear matrix inequality into an LMI, let $Y = L^{\mathrm{T}}X$ then (4.12) can be written as (4.10).

□

The observer gain L for each agent is determined by solving the following optimization problem:

$$\begin{aligned} &\min -\beta + \gamma \\ &\text{s.t.(4.9)} - \text{(4.10) hold.} \end{aligned} \tag{4.13}$$

4.3.2 Unknown Input Observer Design - Attack Detection

In order to design the cyber-attack detection mechanism for MASs, a bank of UIOs needs to be designed for each agent. In particular, a bank of UIOs are designed in each agent i which each UIO is corresponding to neighbor j where $j \in \mathcal{N}_i$. To do so, the following theorem should hold.

***Theorem* 10** Considering system (4.7), if and only if

$$\begin{aligned} &\mathrm{rank}(C_E E_1) = \mathrm{rank}(E_1), \\ &(C_E, PA) \text{ is detectable.} \end{aligned} \tag{4.14}$$

there exists an unknown input cyber-attack detection observer (4.8) which decouples the effect of input, fault, and disturbances from the effect of cyber-attack on the residual and its parameters can be calculated as follows:

$$\begin{aligned} PE_1 &= 0 \\ F &= PA - K_1 C_E \\ K_2 &= FH \\ K &= K_1 + K_2 \end{aligned} \tag{4.15}$$

Proof: In order to design the UIO design parameters F, K, and H, let $e_i^j(t) = x_i(t) - \tilde{x}_i^j(t)$, using Equations (4.7) and (4.8) the state error dynamics can be written as follows:

$$\begin{aligned} e_i^j(t) &= x_i(t) - \tilde{x}_i^j(t) = x_i(t) - z_i^j(t) - Hy_i^j(t), \\ &= x_i(t) - z_i^j(t) - HC_E x_i(t) - HD_a a_i(t), \end{aligned} \tag{4.16}$$

Let us assume a cyber-attack-free system, then the error dynamics can be written as follows:

$$e_i^j(t) = x_i(t) - z_i^j(t) - HC_E x_i(t) = (I - HC_E)x_i(t) - z_i^j(t), \tag{4.17}$$

Thus,

$$\dot{e}_i^j(t) = (I - HC_E)\dot{x}_i(t) - \dot{z}_i^j(t), \tag{4.18}$$

Let $P = (I - HC_E)$, then,

$$\dot{e}_i^j(t) = PAx_i(t) + PE_1\rho_i(t) - Fz_i^j(t) - Ky_i^j(t). \tag{4.19}$$

It should be noted that it is assumed that the system is cyber-attack-free. Hence,

$$\begin{aligned}\dot{e}_i^j(t) =& PAx_i(t) + PE_1\rho_i(t) - Fz_i^j(t) - KC_Ex_i(t),\\ =& (PA - KC_E)x_i(t) + PE_1\rho_i(t) - F(\tilde{x}_i^j(t) - HC_Ex_i(t)),\\ =& (PA - KC_E + FHC_E)x_i(t) + PE_1\rho_i(t) - F\tilde{x}_i^j(t).\end{aligned} \tag{4.20}$$

Now, let $K_2 = FH$, and $K = K_1 + K_2$,

$$\begin{aligned}\dot{e}_i^j(t) =& (PA - K_1C_E)x_i(t) + PE_1\rho_i(t) - F\tilde{x}_i^j(t)\\ =& (PA - K_1C_E - F)x_i(t) + PE_1\rho_i(t) + Fe_i^j(t).\end{aligned} \tag{4.21}$$

To ensure the asymptotic stability of the error dynamic inin (4.21) the following equalities should hold,

$$\begin{aligned}PE_1 =& (I - HC_E)E_1 = 0\\ F =& PA - K_1C_E\\ K_2 =& FH\\ K =& K_1 + K_2\end{aligned} \tag{4.22}$$

The first statement in (4.22) decouples the effect of unknown inputs from the residual signal through design of H, while the matrix K_1 is chosen such that F is Hurwitz. According to [96], the following necessary and sufficient conditions have to be satisfied for the system (4.7) to employ the proposed detection mechanism.

$$\begin{aligned}&\text{rank}(C_EE_1) = \text{rank}(E_1),\\ &(C_E, PA) \text{ is detectable.}\end{aligned} \tag{4.23}$$

It is worth mentioning that when the system is under cyber-attack, the cyber-attack term will appear in the error dynamic. However, since the attacker wants to remain stealthy while he/she is degrading the performance of the system, it is logical to consider that cyber-attack signal does not lead to instability in the system and the error dynamic. Thus, without loss of generality the UIO (4.8) can be used to detect cyber-attack and its error dynamic remain BIBO stable. □

Remark: The proposed UIO filter in this chapter is designed in a decentralized manner as in [63] and [69]. In particular, the design of the proposed UIO does not depend on the Laplacian and topology among the agents because each UIO is working independently from other neighbors. In the other word, each UIO receives the output information from a single neighboring agent and estimate the states of that agent only and not all neighbors in the neighboring set. Accordingly, each agent is equipped with a bank of UIOs and each UIO is responsible for estimating states of one agent only. Hence, its design procedure is studied independent of the topology. It is worth mentioning that in [63] and [69], the effect of fault is not considered and the observers are designed for interconnected systems and smart grids systems while in this chapter, the effect of fault is considered and is applied to a group of UAVs.

The threshold value for cyber-attack detection can then be defined based on the upper bound of the rate of change of cyber-attack signal. However, in this thesis, the author did not assume that this bound is known and the threshold is designed using a Monte Carlo simulation and the effectiveness of the proposed method is verified by the simulation in the next Section.

4.4 Detection Mechanism

The following evaluation function $J(t)$ is used to evaluate the residual signals.

$$J(t) = \sqrt{\frac{1}{\Delta t}\int_0^{\Delta t} r^{\mathrm{T}}(t)r(t)dt}, \tag{4.24}$$

where Δt is a time window and r refers to any residual signal defined in this chapter. A Monte Carlo Simulation is used to design a predefined threshold value J_{th} such that

$$J_{th} = \sup J(t), \tag{4.25}$$

and the detection logic is based on Table 4.1 as discussed in Section 4.2.3.

In particular, the threshold value corresponding to cyber-attack detection observer is calculated by running a Monte Carlo simulation for different values of fault and disturbances. It is worth considering that the magnitude of fault can affect the performance of the cyber-attack detector and it is crucial for the cyber-attack detector to perform an acceptable performance in presence of fault. Similarly, the threshold value corresponding to fault detection observer is estimated using a Monte-Carlo simulation in presence of disturbances. It is worth mentioning that since the Luenberger-based cyber-attack detector is designed locally, the effect of the cyber-attack does not appear in the residual signal. Thus, the fault detector response is regardless of existence of cyber-attack.

The range of cyber-attack values are chosen such that the performance of the system is degraded while the system remains stable. This consideration is not restrictive and it is actually the worst-case scenario that can be considered by defender while designing the monitoring system. The cyber-attacker wants to remain stealthy and fool the monitoring system. Thus, he/she

injects a cyber-attack which does not lead the system to instability. Also, the effect of the cyber-attack should not trigger the fault detection system in the neighboring agents in ideal situation from cyber-attacker point of view which is not the case in the proposed scenario. However, even if the cyber-attacker injects a cyber-attack that triggers the fault detector residual, without the cyber-attack detection residual, the operator is not able to perform appropriate reaction in order to compensate the effect of anomaly. The above explanation clarifies the importance of designing the fault and cyber-attack detection scheme in multi-agent system.

4.5 Simulation Results

The efficacy of the proposed method is verified by simulating the lateral dynamic model of a small UAV proposed in [98]. The continues-time state space model of this UAV is given as follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + B_{\mathrm{f}} f(t) + B_{\mathrm{d}} d(t), \\ y(t) &= Cx(t) + D_{\mathrm{a}} a(t) + D_{\mathrm{f}} f(t) + D_{\mathrm{d}} d(t), \end{aligned} \tag{4.26}$$

where $x_1(t)$ is the body axis y direction velocity (knot), $x_2(t)$ and $x_3(t)$ are the roll and yaw angular rate, respectively. $x_4(t)$ and $x_5(t)$ denote the roll and yaw angle, respectively. $u_1(t)$ is the aileron and $u_2(t)$ is the rudder control. The output variables for the lateral model are sideslip angel, roll and yaw angular rate, and roll and yaw angle, respectively. An output-consensus controller is designed as described in Chapter 2. The simulation has been done for a multi-agent system consisting of four agents based on the topology shown in Figure 4.2.

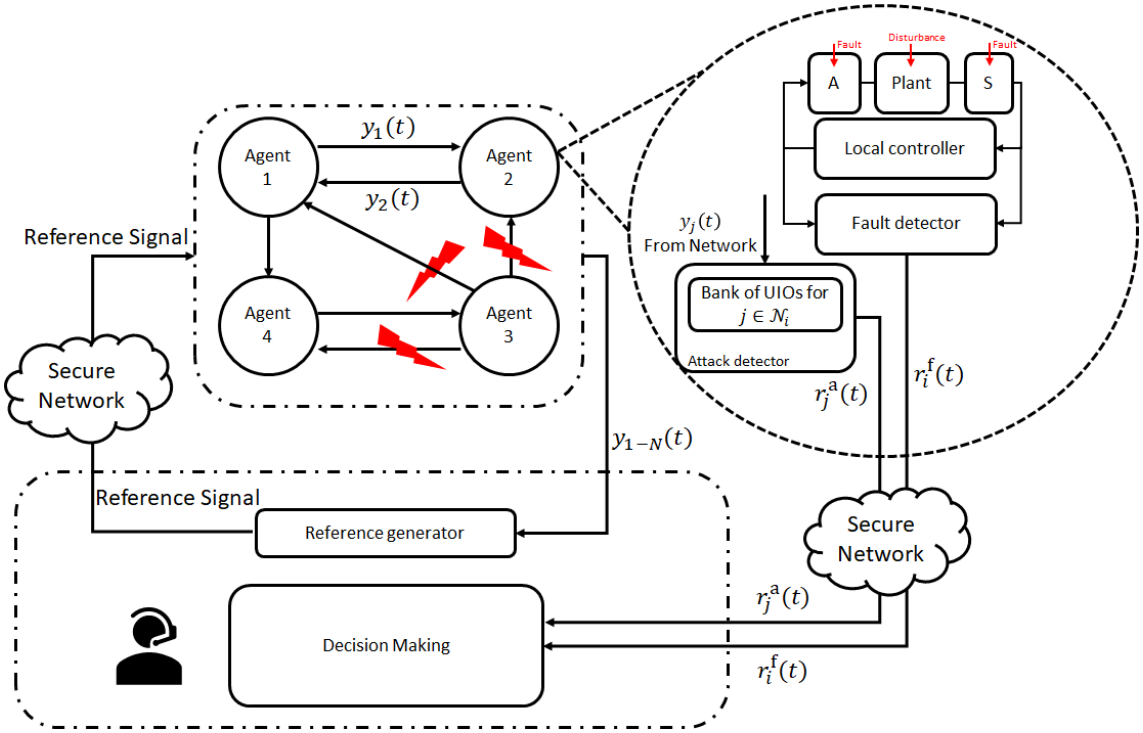


Figure 4.2: MASs Block Diagram.

$$A = \begin{bmatrix} -1.476 & 0.368 & -16.997 & 9.794 & 0.000 \\ -5.703 & -21.709 & -6.809 & 0.000 & 0.000 \\ 1.329 & -0.100 & -2.698 & 0.000 & 0.000 \\ 0.000 & 1.000 & -0.049 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.001 & 0.000 & 0.000 \end{bmatrix},$$

$$B_f = B = \begin{bmatrix} 2.147 & 5.480 \\ -105.001 & -2.792 \\ -6.131 & -12.398 \\ 0.000 & 0.000 \\ 0.000 & 0.000 \end{bmatrix}, \quad D_f = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \\ 0 & 0.5 \\ 0.4 & 0 \\ 0 & 0 \end{bmatrix} \tag{4.27}$$

$$C = \begin{bmatrix} 0.059 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.000 \end{bmatrix},$$

$$B_d = \begin{bmatrix} 0.147 & 0.21 & 0.6131 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$$

$$D_{\mathrm{d}} = \begin{bmatrix} 0.2 & 0 & 0.3 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \quad D_{\mathrm{a}} = \begin{bmatrix} 1 & 0.2 & 0.5 & 0.1 & 0.45 \end{bmatrix}^{\mathrm{T}},$$

The optimization problem (4.13) is solved to calculate observers gain L. In all scenarios disturbance is $d(t) = \sin(4t)e^{-0.05t}$. In order to design the UIO-based cyber-attack detector, the matrix H is chosen such that $(I - HC_E)E_1 = 0$. The matrix K_1 is selected to assign the eigenvalues of F to $\{-10, -15, -5, -5, -5\}$. The rest of UIO matrices are calculated as in (4.22). To verify the effectiveness of the proposed method, three different scenarios have been taken into account as follows:

Scenario 1 - Fault only: In this scenario, a fault signal $f(t) = [0.2, 0.2]^{\mathrm{T}}$ is injected on the actuator of agent 3 at $t \geq 30$ seconds in presence of disturbance. Figure 4.3 delineates the residual signal corresponding to fault detection module in each agent in which the residual signal corresponding to agent 3 is triggered and fault is successfully detected. As mentioned earlier in this chapter, each agent is equipped with a bank of UIOs to detect the cyber-attack on the communication link between it neighbors. In particular, as shown in Figure 4.2 , the neighbors of agent 1 are agent 2 and 3. Accordingly, if the cyber-attack is injected on the communication link between agent 1 and 2 or between the agents 1 and 3, their corresponding residual will be triggered. In this scenario, no cyber-attack is injected on the communication links. Thus,

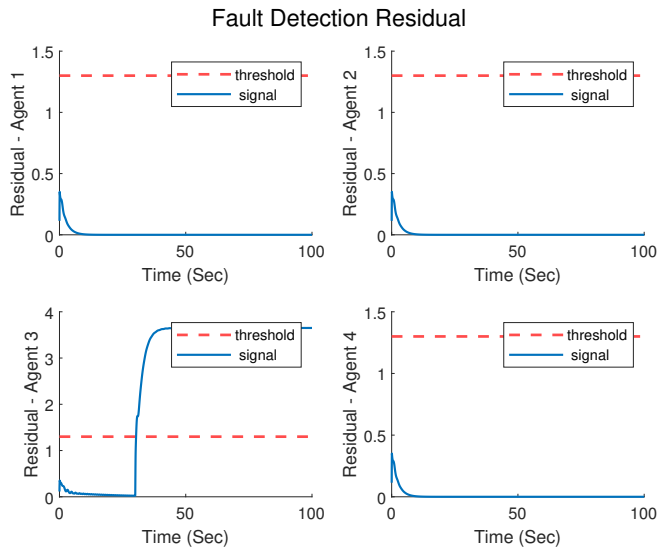
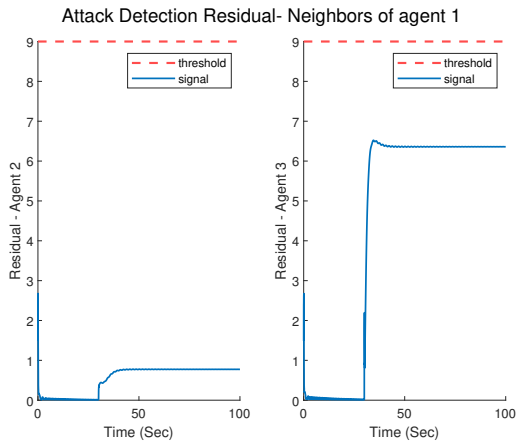


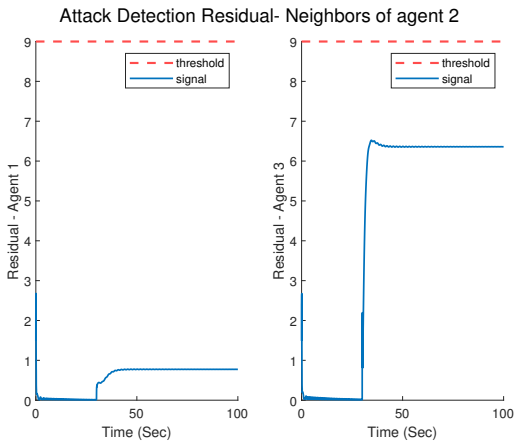
Figure 4.3: Fault detection residual corresponding to agent 1,2,3, and 4.

none of the residuals are triggered. Figure 4.4a presents the cyber-attack detection residuals corresponding to neighbors of agent 1 (agents 2 and 3). Figure 4.4b delineates the cyber-attack detection residuals corresponding to neighbors of agent 2 (agents 1 and 3). The cyber-attack detection residual signal corresponding to the neighbors of agents 3 and 4 are shown in Figures 4.5a and 4.5b, respectively. Note that the neighbor of agent 3 is agent 4 and the neighbors of agent 4 are agents 1 and 3.

Scenario 2 - Cyber-Attack only: In the second scenario, the effectiveness of the proposed UIO-based observer is verified and a cyber-attack signal $a(t) = 0.5$ is injected on the communication link that transmit the output information of agent 1 to all its neighbors at $t \geq 70$ seconds. It is worth mentioning that a directed communication topology is considered in this chapter. The fault detection residuals did not triggered as shown in Figure 4.6 in presence of disturbances and cyber-attack. However, the cyber-attack detection residual signal corresponding to agent 1 in other agents have been triggered as follows. Since agent 1 is in the neighboring set of agent 2, the cyber-attack detection residual corresponding to agent 1 inside agent 2 is triggered once the cyber-attack is injected as is shown in Figure 4.7b. Similarly, the corresponding

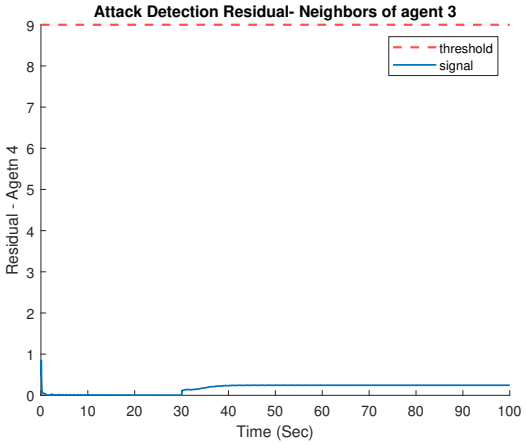


(a) Attack detection residual

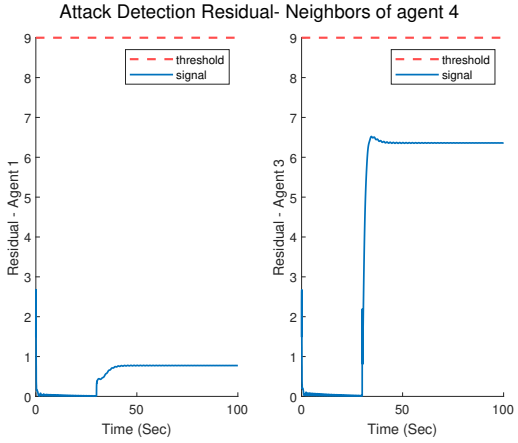


(b) Attack detection residual

Figure 4.4: Attack detection residual corresponding to a) neighbors of agent 1 and b) neighbors of agent 2.



(a) Attack detection residual



(b) Attack detection residual

Figure 4.5: Attack detection residual corresponding to a) neighbors of agent 3 and b) neighbors of agent 4.

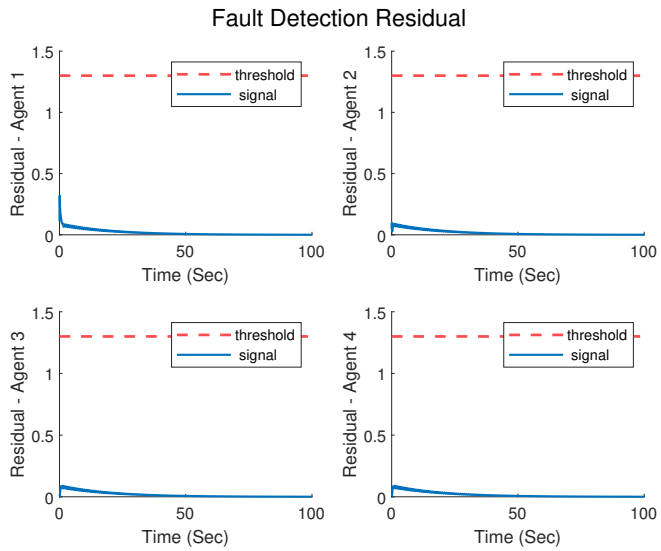
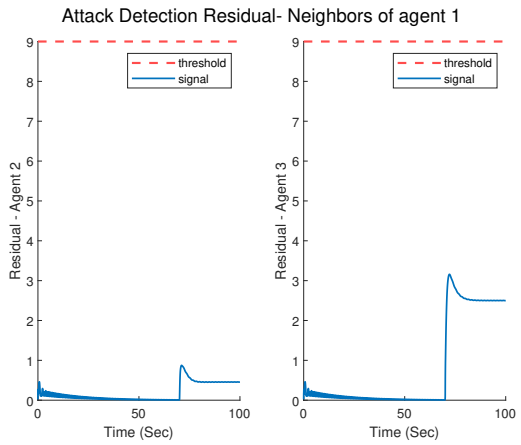


Figure 4.6: Fault detection residual corresponding to agent 1,2,3, and 4.

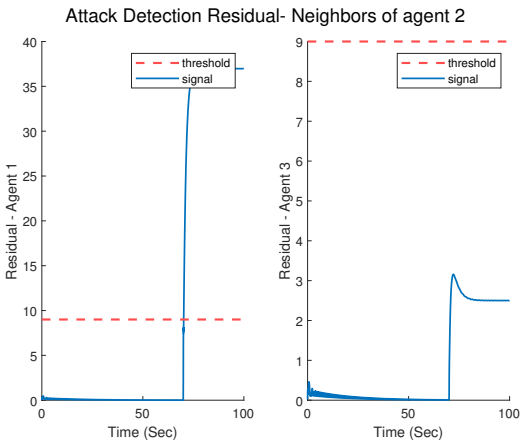
cyber-attack detection residual to agent 1 inside agent 4 is also triggered (see Figure 4.8b). Note that since the cyber-attack is injected on the communication link, it is not detectable by agent 1 itself. Figure 4.9 delineates the effect of cyber-attack on the output of agents 1 to 4. It is obvious that agents deviates from consensus as soon as cyber-attack is injected.

Scenario 3 - Concurrent Fault and Cyber-Attack: Finally, in this scenario, a fault signal $f(t) = [0.4, 0.1]^{\mathrm{T}}$ is injected at $t \geq 30$ seconds on the agent 3. An cyber-attack signal $a(t) = 0.6$ is also injected on the communication links that transmit the output information of agent 3 to other neighbors at $t \geq 70$ seconds. It is obvious from Figure 4.10 that fault is injected at 30 seconds on the agent 3. According to Figures 4.11a, 4.11b, and 4.12b, the communication link from agent 3 is under cyber-attack from $t = 70$ seconds onward.

Accordingly, the efficiency of the proposed detection method is verified by simulation of multiple UAVs in presence of disturbances. A set of agents under directed communication topology is considered which are equipped with a output-consensus controller. In the other word, the agents are required to transmit their output information among their neighbors to reach a consensus. It is first shown that the fault is detectable locally, using a Luenberger-based

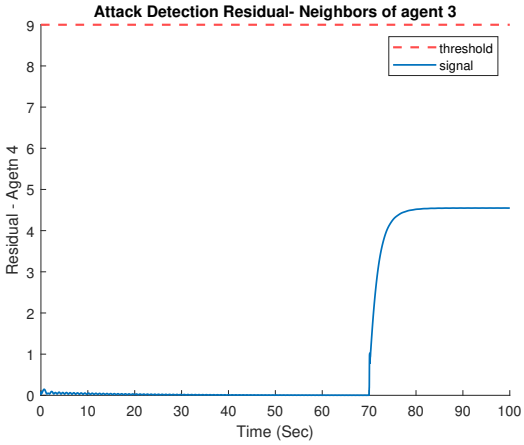


(a) Attack detection residual

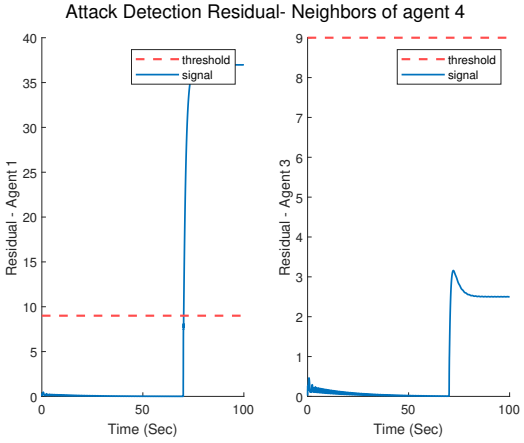


(b) Attack detection residual

Figure 4.7: Attack detection residual corresponding to a) neighbors of agent 1 and b) neighbors of agent 2.



(a) Attack detection residual



(b) Attack detection residual

Figure 4.8: Attack detection residual corresponding to a) neighbors of agent 3 and b) neighbors of agent 4.

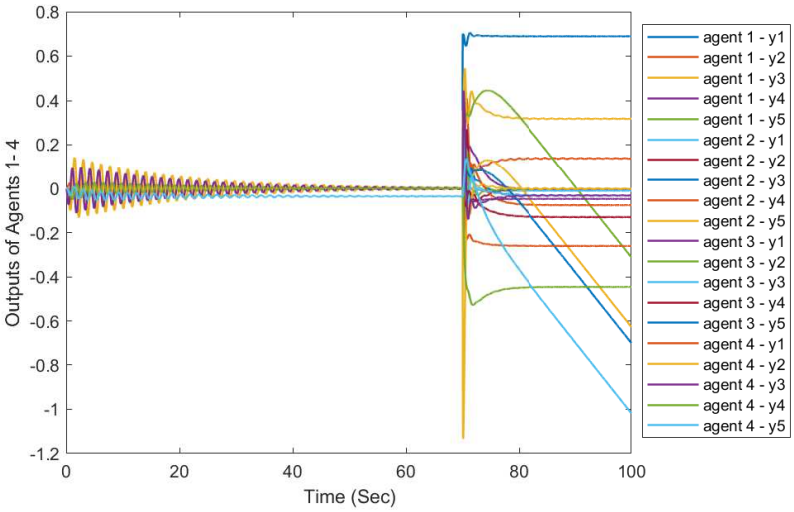


Figure 4.9: Output of agents 1 - 4 in the presence of cyber-attack, disturbances, and noise.

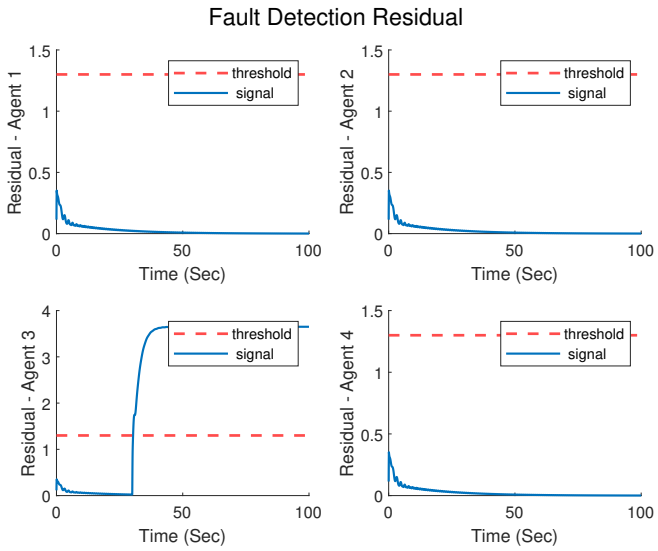
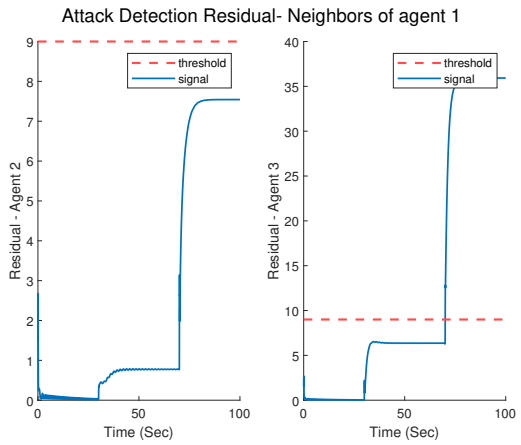
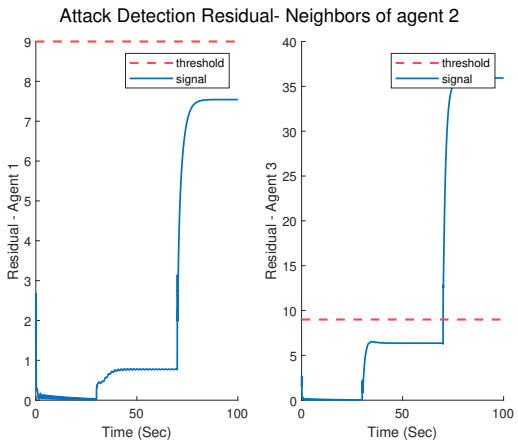


Figure 4.10: Fault detection residual corresponding to agent 1,2,3, and 4.

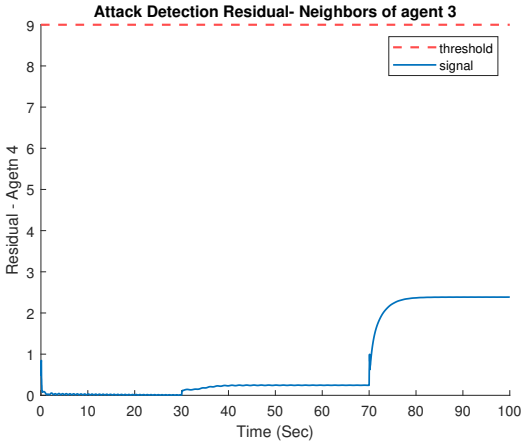


(a) Attack detection residual

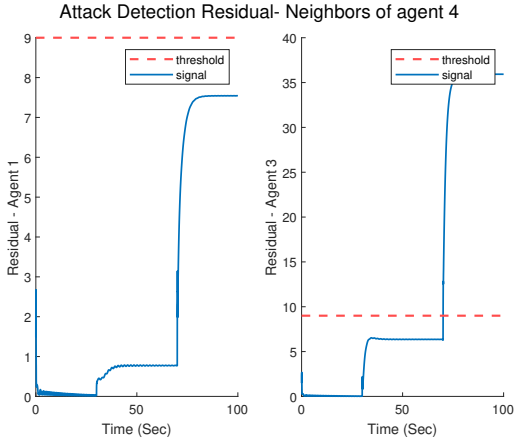


(b) Attack detection residual

Figure 4.11: Attack detection residual corresponding to a) neighbors of agent 1 and b) neighbors of agent 2.



(a) Attack detection residual



(b) Attack detection residual

Figure 4.12: Attack detection residual corresponding to a) neighbors of agent 3 and b) neighbors of agent 4.

observer. Then, the efficacy of the UIO-based cyber-attack detection module is verified in presence of disturbances. In particular, it is assumed that cyber-attack is injected on the output communication links of an agent. Each agent is equipped with a bank of UIOs corresponding to each of its neighbors. Accordingly, it is shown that each agent is able to detect the cyber-attack on the communication links within its neighboring set. Finally, for the case of concurrent fault and cyber-attack, it is observed that the fault and cyber-attack has been easily detected and classified from each other by observing the residual signals in the command and control subsystem.

The problem solved in this chapter is rarely studied in the literature. For instance, in [62] the problem of robust fault and cyber-attack detection and identification has been tackled using multi-objective framework. Two different filter has been designed and in order to achieve the fault and cyber-attack detection and identification in a MAS, a bank of each filter is required to be implemented in each agent. Hence, compared to the work in this chapter it is more complex from both computation and implementation point of view. Also, in [62], the authors consider cyber-attack on communication and they assume that the control input of neighboring agent is communicated over the communication channel but it remains healthy. In contrast, this assumption is relaxed in this chapter and the cyber-attack detector is designed in the absence of input of neighboring agents. In addition, relaxing this assumption reduces the communication costs in this chapter compared to [62].

In order to show the effectiveness of the proposed detectors, two Monte-Carlo simulations performed for a set of 100 simulations. In particular, to evaluate the performance of the fault detector, 100 sets of simulations are conducted for fault range of $([0.1, 1.2]; [0.1, 0.9])$ and cyber-attack range of $[0.1, 1.2]$. The values of cyber-attack and fault are randomly generated for each simulation. Table 4.2 shows the confusion matrix corresponding to fault detector. The main

components of this confusion matrix are as follows:

TP = when fault is detected correctly

TN = when no fault is detected correctly.

FP = when fault is detected wrongly (there is no fault but residual is triggered.)

FN = when fault is not detected correctly (there is fault but not detected.)

Table 4.2: *Confusion Matrix Corresponding to Fault Detector*

		Expected	
	N=100	Yes	No
Detected	Yes	TP = 73	FP = 0
	No	FN = 3	TN = 24

The percentage of the detection rate (DR) corresponding to the fault detector is calculated as follows:

$$Dr = \frac{TP}{TP + FN} * 100 = 96.05\%. \tag{4.28}$$

Similarly, for 100 iterations and with the same range of cyber-attack and fault, the confusion matrix shown in Table 4.3 is produced. The main elements of the confusion matrix corresponding to cyber-attack detector are as follows:

TP = when cyber-attack is detected correctly

TN = when no cyber-attack is detected correctly.

FP = when cyber-attack is detected wrongly (there is no cyber-attack but residual is triggered.)

FN = when cyber-attack is not detected correctly (there is cyber-attack but not detected.)

According to this table, the detection rate for the cyber-attack detector is 95.31%. Table 4.4 presents the precision and accuracy of the proposed fault and cyber-attack detector in this chap-

Table 4.3: *Confusion Matrix Corresponding to Cyber-Attack Detector*

		Expected	
	N=100	Positive	Negative
Detected	Positive	TP = 61	FP = 2
	Negative	FN = 3	TN = 44

ter as follows:

$$precision = PR = \frac{TP}{TP + FP} \tag{4.29}$$

$$Accuracy = ACC = \frac{TP + TN}{P + N} \tag{4.30}$$

where $P = TP + FN$ denotes the number of real positive and $N = FP + TN$ denotes the number of real negative cases.

Table 4.4: *Extra Measures on Confusion Matrices*

Filter Name	Table	PR	ACC
Luenberger-based Fault Detector	4.2	100%	97%
UIO-based Attack Detector	4.3	96.83%	95.45%

4.6 Conclusion

To sum up, this chapter presents a detection mechanism to detect fault and cyber-attack from each other in presence of disturbance in multi-agent systems. The cyber-attack is injected on the communication link between two or multiple agents. It can have the same effect as fault on the performance of the system while it cannot be detected by the fault detector mechanism. It is crucial for the monitoring system to distinguish between fault and cyber-attack in order to compensate the effect of anomaly.

Each agent is equipped with a Luenberger-based observer to detect fault on actuator and sen-

sor locally. The fault is detected by evaluating a residual signal generated by the observer. On the other hand, each agent is equipped with a bank of UIO-based observers to detect the cyber-attack on the link between the agent and its neighbors. The UIO-based observer is also generates residual signal corresponding to each of the neighbors. The residuals are then transmitted to the monitoring system to detect fault and cyber-attack in presence of disturbance. The efficiency of the proposed method is validated by simulations using a linear lateral model of a UAV.

Chapter 5

SUMMARY AND FUTURE WORK

The usage of cyber-physical systems in different industries is exponentially increasing in the current era of technology in which there is a huge tendency toward automation of devices and reduce the human operator in the field. Also, development of mutli-agent systems to perform different tasks such as accessing remote areas is of high interest. Accordingly, robust detection of anomalies on cyber-physical systems and multi-agent systems is of growing interest in the recent studies. There exists variety of methods for fault diagnosis in the cyber-physical systems and multi-agent systems in the presence of fault and disturbances. In addition, several authors in the recent studies are focusing on cyber-attack detection and designing robust controller in the presence of attack. However, designing a monitoring system to be able to detect fault and cyber-attack in the cyber-physical systems and multi-agent systems requires more attention. The fault signal and cyber-attack may have an identical effect on the performance of the system however, it is crucial to distinguish them from each other to be able to reach the desired goal of any system. The main focus of this thesis is to design a robust detection method for cyber-physical and multi-agent systems in the presence of concurrent fault and attack and disturbance.

The third chapter describes a robust detection mechanism for a cyber-physical system using

$\mathscr{H}_-$ and $\mathscr{H}_\infty$ performance indeces. The problem is solved in finite frequency using multi-objective framework. Two types of filters have been designed namely Luenberger observer and Normal filter. Based on the simulation results, it is concluded that Luenberger-based observer has false alarms in the case of attack detection in the presence of fault and it is highly dependent on the magnitude of the fault. Thus, normal filter is designed for the attack detection and a combination of the proposed filters have been used to detect faults and attacks in presence of disturbances and noise in the command and control subsystem. The proposed method is verified by simulating a VTOL aircraft in Matlab/Simulink.

In Chapter 4, a robust fault and cyber-attack detection mechanism is proposed for multi-agent systems. The multi-agent system is considered under a directed communication topology with an output-consensus controller under healthy condition. The Luenberger-based observer is designed locally to detect fault in the presence of disturbance in the full frequency domain. In addition, a bank of UIO-based attack detectors are designed in each agent to detect the injected attack on the communication link between its neighbors. The residual signals have been sent to the command and control for decision making process. The proposed method is verified by simulations of a multiple UAVs in Matlab/Simulink.

In order to compare Chapter 3 with Chapter 4, it should be noted that a single agent in the framework of CPS is considered in Chapter 3 while in the later chapter a multi-agent framework is considered. Design of a monitoring scheme is presented in both of these chapters to detect and classified fault and cyber-attack in presence of disturbances. However, Chapter 3 covers an LMI approach in finite frequency domain in order to design detection modules while in Chapter 4, the analysis have been done in the full frequency domain. The computational complexity is higher for the case of finite frequency analysis compare to the full frequency domain. However, it is worth mentioning that the full frequency analysis is not convex for strictly proper system.

In Chapter 3 the combination of Luenberger observer and Normal filter is designed and utilized while in Chapter 4 the combination of Luenberger and UIO is used to detect fault and cyber-attack. The main reason to choose UIO for attack detection is that an agent is unaware of the fault and input of its neighbors, hence it cannot reduce the effects of these unknown inputs on the residual signals. The attack detection module in the case of single agent system is implemented in the command and control center while in the case of multi-agent system, it is implemented in each agent locally. It is worth considering that due to communication burden it is more efficient to design the detection modules locally in the case of multi-agent systems.

The main future directions for this thesis are as follows:

• Development of fault and attack detection scheme for different ranges of frequencies in Chapter 3.

As shown in Chapter 3, the proposed fault and attack detection mechanism is designed for attacks in low frequency. However, the attacker can inject the attack in other frequency ranges as it is assumed. Hence, it is important to design attack detector for different frequency ranges.

• Development of isolation scheme for simultaneous fault and cyber-attack in cyber-physical system in presence of disturbance is one of the important future works to be studied.

• Development of a controller which is able to compensate the effect of detected fault and attack.

In the domain of control theory, to ensure safety and security of an autonomous system, the next step after detection, isolation, and classification of an anomaly is to compensate it so the desired goal can be achieved. Thus, designing a controller that can tolerate the effect of fault and compensate the effect of cyber-attack in the cyber-physical systems simultaneously, is in a future direction of this thesis.

• Development of fault and attack detection methods for other types of cyber-attacks such as

DoS, Covert, and zero dynamics attacks.

In this thesis, the proposed detection methods is designed for the false data injection attack on the output measurement of the system. Hence, considering other types of cyber-attack and designing a diagnosis mechanism in the presence of concurrent fault is considered to be a future work.

• Performing a vulnerability analysis for the designed monitoring system in this thesis.

On of the future works in this thesis is to evaluate the effectiveness of designed monitoring systems in the presence of other types of attacks and recognize which attacks may remain stealthy to the proposed method. This will help the defender to recognize and predict the possible cyber-attacks which are undetectable to enhance the safety and security measures of the CPS.

REFERENCES

[1] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[2] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*. IEEE, 2011, pp. 1–6.

[3] S. K. Khaitan and J. D. McCalley, "Design techniques and applications of cyberphysical systems: A survey," *IEEE Systems Journal*, vol. 9, no. 2, pp. 350–365, 2014.

[4] G. B. K. L. R. R. Rajkumar, "An end-to-end integration framework for automotive cyber-physical systems using sysweaver," *AVICPS 2010*, p. 23, 2010.

[5] H. Chen, "Applications of cyber-physical system: a literature review," *Journal of Industrial Integration and Management*, vol. 2, no. 03, p. 1750012, 2017.

[6] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, 2011.

[7] S. H. Ahmed, G. Kim, and D. Kim, "Cyber physical system: Architecture, applications and research challenges," in *2013 IFIP Wireless Days (WD)*. IEEE, 2013, pp. 1–5.

[8] S. A. Haque, S. M. Aziz, and M. Rahman, "Review of cyber-physical system in healthcare," *International Journal of Distributed Sensor Networks*, vol. 10, no. 4, p. 217415, 2014.

[9] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[10] Y. Li, P. Zhang, L. Zhang, and B. Wang, "Active synchronous detection of deception attacks in microgrid control systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 373–375, 2016.

[11] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *2009 47th Annual Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2009, pp. 911–918.

[12] F. Miao, M. Pajic, and G. J. Pappas, "Stochastic game approach for replay attack detection," in *52nd IEEE Conference on Decision and Control*. IEEE, 2013, pp. 1854–1859.

[13] H. S. Sánchez, D. Rotondo, T. Escobet, V. Puig, and J. Quevedo, "Bibliographical review on cyber attacks from a control oriented perspective," *Annual Reviews in Control*, vol. 48, pp. 103–128, 2019.

[14] H. Habibzadeh, B. H. Nussbaum, F. Anjomshoa, B. Kantarci, and T. Soyata, "A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities," *Sustainable Cities and Society*, vol. 50, p. 101660, 2019.

[15] V. R. Palleti, Y. C. Tan, and L. Samavedham, "A mechanistic fault detection and isolation approach using Kalman filter to improve the security of cyber physical systems," *Journal of Process Control*, vol. 68, pp. 160–170, 2018.

[16] D. Xu, F. Zhu, Z. Zhou, and X. Yan, "Distributed fault detection and estimation in cyber–physical systems subject to actuator faults," *ISA Transactions*, vol. 104, pp. 162–174, 2020.

[17] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Distributed sensor fault diagnosis for a network of interconnected cyberphysical systems," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 11–23, 2015.

[18] J. CHEN, R. J. PATTON, and H.-Y. ZHANG, "Design of unknown input observers and robust fault detection filters," *International Journal of Control*, vol. 63, no. 1, pp. 85–105, 1996.

[19] N. Tudoroiu and K. Khorasani, "Fault detection and diagnosis for satellite's attitude control system (acs) using an interactive multiple model (IMM) approach," in *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005.* IEEE, 2005, pp. 1287–1292.

[20] N. Meskin, E. Naderi, and K. Khorasani, "A multiple model-based approach for fault diagnosis of jet engines," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 254–262, 2011.

[21] B. Pourbabaee, N. Meskin, and K. Khorasani, "Sensor fault detection, isolation, and identification using multiple-model-based hybrid Kalman filter for gas turbine engines," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 4, pp. 1184–1200, 2015.

[22] M. Davoodi, N. Meskin, and K. Khorasani, "A single dynamic observer-based module for design of simultaneous fault detection, isolation and tracking control scheme," *International Journal of Control*, vol. 91, no. 3, pp. 508–523, 2018.

[23] H. Wang and G.-H. Yang, "A finite frequency domain approach to fault detection observer design for linear continuous-time systems," *Asian Journal of Control*, vol. 10, no. 5, pp. 559–568, 2008.

[24] J. Chen and Y.-Y. Cao, "A stable fault detection observer design in finite frequency domain," *International Journal of Control*, vol. 86, no. 2, pp. 290–298, 2013.

[25] 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*, vol. 49, no. 7, pp. 1572–1585, 2018.

[26] Y. Long and G.-H. Yang, "Fault detection in finite frequency domain for networked control systems with missing measurements," *Journal of the Franklin Institute*, vol. 350, no. 9, pp. 2605–2626, 2013.

[27] X. Li and G. Yang, "Fault detection in finite frequency domain for linear systems under feedback control," in *2009 IEEE Control Applications,(CCA) & Intelligent Control,(ISIC)*. IEEE, 2009, pp. 1332–1337.

[28] M. Zhou, Z. Wang, Y. Shen, and M. Shen, "$H\_/H_\infty$ fault detection observer design in finite-frequency domain for lipschitz non-linear systems," *IET Control Theory & Applications*, vol. 11, no. 14, pp. 2361–2369, 2017.

[29] A. Chibani, M. Chadli, and N. B. Braiek, "A finite frequency approach to H_∞ filtering for t–s fuzzy systems with unknown inputs," *Asian Journal of Control*, vol. 18, no. 5, pp. 1608–1618, 2016.

[30] S. Tan, J. M. Guerrero, P. Xie, R. Han, and J. C. Vasquez, "Brief survey on attack detection methods for cyber-physical systems," *IEEE Systems Journal*, vol. 14, no. 4, pp. 5329–5339, 2020.

[31] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[32] A. Sargolzaei, C. D. Crane, A. Abbaspour, and S. Noei, "A machine learning approach for fault detection in vehicular cyber-physical systems," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2016, pp. 636–640.

[33] K. Paridari, N. O'Mahony, A. E.-D. Mady, R. Chabukswar, M. Boubekeur, and H. Sandberg, "A framework for attack-resilient industrial control systems: Attack detection and controller reconfiguration," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 113–128, 2017.

[34] C. De Persis and P. Tesi, "Input-to-state stabilizing control under denial-of-service," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2930–2944, 2015.

[35] D. Ding, Q.-L. Han, Y. Xiang, X. Ge, and X.-M. Zhang, "A survey on security control and attack detection for industrial cyber-physical systems," *Neurocomputing*, vol. 275, pp. 1674–1683, 2018.

[36] C.-Y. Gu, J.-W. Zhu, W.-A. Zhang, and L. Yu, "Sensor attack detection for cyber-physical systems based on frequency domain partition," *IET Control Theory & Applications*, vol. 14, no. 11, pp. 1452–1466, 2020.

[37] Y. Z. Lun, A. D'Innocenzo, F. Smarra, I. Malavolta, and M. D. Di Benedetto, "State of the art of cyber-physical systems security: An automatic control perspective," *Journal of Systems and Software*, vol. 149, pp. 174–216, 2019.

[38] M. Taheri, K. Khorasani, I. Shames, and N. Meskin, "Cyber attack and machine induced fault detection and isolation methodologies for cyber-physical systems," *arXiv preprint arXiv:2009.06196*, 2020.

[39] Y. Wu and J. Dong, "Cyber-physical attacks against state estimators based on a finite frequency approach," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.

[40] Q. Su, Z. Fan, Y. Long, and J. Li, "Attack detection and secure state estimation for cyber-physical systems with finite-frequency observers," *Journal of the Franklin Institute*, vol. 357, no. 17, pp. 12 724–12 741, 2020.

[41] A. Anwar, A. N. Mahmood, and Z. Shah, "A data-driven approach to distinguish cyber-attacks from physical faults in a smart grid," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 2015, pp. 1811–1814.

[42] B. R. Amin, A. Anwar, and M. Hossain, "Distinguishing between cyber injection and faults using machine learning algorithms," in *2018 IEEE Region Ten Symposium (Tensymp)*. IEEE, 2018, pp. 19–24.

[43] A. Eslami, F. Abdollahi, and K. Khorasani, "Stochastic fault and cyber-attack detection and consensus control in multi-agent systems," *International Journal of Control*, pp. 1–19, 2021.

[44] M. Davoodi, M. Chadli, N. Meskin, and J. M. Velni, "Finite frequency fault diagnosis for heterogeneous multi-agent lpv systems," in *2019 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2019, pp. 1018–1023.

[45] D. Ding, Q.-L. Han, Z. Wang, and X. Ge, "A survey on model-based distributed control and filtering for industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2483–2499, 2019.

[46] Z. Zuo, Q.-L. Han, B. Ning, X. Ge, and X.-M. Zhang, "An overview of recent advances in fixed-time cooperative control of multiagent systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2322–2334, 2018.

[47] A. Dorri, S. S. Kanhere, and R. Jurdak, "Multi-agent systems: A survey," *IEEE Access*, vol. 6, pp. 28 573–28 593, 2018.

[48] H. Yang, Q.-L. Han, X. Ge, L. Ding, Y. Xu, B. Jiang, and D. Zhou, "Fault-tolerant cooperative control of multiagent systems: A survey of trends and methodologies," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 4–17, 2019.

[49] M. Davoodi, N. Meskin, and K. Khorasani, "Simultaneous fault detection and consensus control design for a network of multi-agent systems," *Automatica*, vol. 66, pp. 185–194, 2016.

[50] ——, "Event-triggered multiobjective control and fault diagnosis: A unified framework," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 298–311, 2016.

[51] M. R. Davoodi, K. Khorasani, H. A. Talebi, and H. R. Momeni, "Distributed fault detection and isolation filter design for a network of heterogeneous multiagent systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 1061–1069, 2013.

[52] Y. Bai and J. Wang, "Fault detection and isolation using relative information for multi-agent systems," *ISA Transactions*, vol. 116, pp. 182–190, 2021.

[53] I. Shames, A. M. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, no. 12, pp. 2757–2764, 2011.

[54] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Distributed fault detection and isolation resilient to network model uncertainties," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2024–2037, 2014.

[55] X. Li, C. K. Ahn, D. Lu, and S. Guo, "Robust simultaneous fault estimation and nonfragile output feedback fault-tolerant control for markovian jump systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 9, pp. 1769–1776, 2018.

[56] S. Hajshirmohamadi, F. Sheikholeslam, M. Davoodi, and N. Meskin, "Event-triggered simultaneous fault detection and tracking control for multi-agent systems," *International Journal of Control*, vol. 92, no. 8, pp. 1928–1944, 2019.

[57] X. Liu, X. Gao, and J. Han, "Robust unknown input observer based fault detection for high-order multi-agent systems with disturbances," *ISA Transactions*, vol. 61, pp. 15–28, 2016.

[58] X. Gao, X. Liu, and J. Han, "Reduced order unknown input observer based distributed fault detection for multi-agent systems," *Journal of the Franklin Institute*, vol. 354, no. 3, pp. 1464–1483, 2017.

[59] A. Khazraei, H. Kebriaei, and F. R. Salmasi, "Replay attack detection in a multi agent system using stability analysis and loss effective watermarking," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 4778–4783.

[60] L. N. Lemma, S.-H. Kim, and H.-L. Choi, "An unknown-input-observer based approach for cyber attack detection in formation flying uavs," in *AIAA Infotech@ Aerospace*, 2016, p. 0916.

[61] E. Mousavinejad, X. Ge, Q.-L. Han, F. Yang, and L. Vlacic, "Detection of cyber attacks on leader-following multi-agent systems," in *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1. IEEE, 2019, pp. 6243–6248.

[62] Y. Li, H. Fang, and J. Chen, "Anomaly detection and identification for multiagent systems subjected to physical faults and cyber attacks," *IEEE Transactions on Industrial Electronics*, 2019.

[63] A. Barboni, H. Rezaee, F. Boem, and T. Parisini, "Detection of covert cyber-attacks in interconnected systems: A distributed model-based approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3728–3741, 2020.

[64] F. Pasqualetti, A. Bicchi, and F. Bullo, "Distributed intrusion detection for secure consensus computations," in *2007 46th IEEE Conference on Decision and Control*. IEEE, 2007, pp. 5594–5599.

[65] ——, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, 2011.

[66] P. Wang and M. Govindarasu, "Multi-agent based attack-resilient system integrity protection for smart grid," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3447–3456, 2020.

[67] R. Merco, Z. A. Biron, and P. Pisu, "Replay attack detection in a platoon of connected vehicles with cooperative adaptive cruise control," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 5582–5587.

[68] Z. Abdollahi Biron, S. Dey, and P. Pisu, "Real-time detection and estimation of denial of service attack in connected vehicle systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3893–3902, 2018.

[69] A. J. Gallo, M. S. Turan, P. Nahata, F. Boem, T. Parisini, and G. Ferrari-Trecate, "Distributed cyber-attack detection in the secondary control of DC microgrids," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 344–349.

[70] X. Luo, X. Wang, X. Pan, and X. Guan, "Detection and isolation of false data injection attack for smart grids via unknown input observers," *IET Generation, Transmission & Distribution*, vol. 13, no. 8, pp. 1277–1286, 2019.

[71] J. H. Seo, H. Shim, and J. Back, "Consensus of high-order linear systems using dynamic output feedback compensator: Low gain approach," *Automatica*, vol. 45, no. 11, pp. 2659–2664, 2009.

[72] Z. Feng, G. Wen, and G. Hu, "Distributed secure coordinated control for multiagent systems under strategic attacks," *IEEE Transactions on Cybernetics*, vol. 47, no. 5, pp. 1273–1284, 2016.

[73] A. Mustafa, H. Modares, and R. Moghadam, "Resilient synchronization of distributed multi-agent systems under attacks," *Automatica*, vol. 115, p. 108869, 2020.

[74] H. Rezaee, T. Parisini, and M. M. Polycarpou, "Resiliency in dynamic leader–follower multiagent systems," *Automatica*, vol. 125, p. 109384, 2021.

[75] E. M. Amullen, S. Shetty, and L. H. Keel, "Model-based resilient control for a multi-agent system against denial of service attacks," in *2016 World Automation Congress (WAC)*. IEEE, 2016, pp. 1–6.

[76] T.-Y. Zhang and D. Ye, "Distributed event-triggered control for multi-agent systems under intermittently random denial-of-service attacks," *Information Sciences*, vol. 542, pp. 380–390.

[77] R. Moghadam and H. Modares, "Resilient autonomous control of distributed multiagent systems in contested environments," *IEEE Transactions on Cybernetics*, vol. 49, no. 11, pp. 3957–3967, 2019.

[78] A.-Y. Lu and G.-H. Yang, "Distributed consensus control for multi-agent systems under denial-of-service," *Information Sciences*, vol. 439-440, pp. 95–107, 2018.

[79] Z. Zuo, X. Cao, and Y. Wang, "Security control of multi-agent systems under false data injection attacks," *Neurocomputing*, vol. 404, pp. 240–246, 2020.

[80] Y. Yang, H. Xu, and D. Yue, "Observer-based distributed secure consensus control of a class of linear multi-agent systems subject to random attacks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 8, pp. 3089–3099, 2019.

[81] W. He, Z. Mo, Q.-L. Han, and F. Qian, "Secure impulsive synchronization in lipschitz-type multi-agent systems subject to deception attacks," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1326–1334, 2020.

[82] "Secure impulsive synchronization control of multi-agent systems under deception attacks," *Information Sciences*, vol. 459, pp. 354–368, 2018.

[83] Y. Xu, M. Fang, P. Shi, and Z.-G. Wu, "Event-based secure consensus of mutiagent systems against DoS attacks," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3468–3476, 2020.

[84] M. Mola, N. Meskin, K. Khorasani, and A. Massoud, "Distributed event-triggered consensus-based control of DC microgrids in presence of DoS cyber attacks," *IEEE Access*, vol. 9, pp. 54 009–54 021, 2021.

[85] Z. Feng, G. Hu, and G. Wen, "Distributed consensus tracking for multi-agent systems under two types of attacks," *International Journal of Robust and Nonlinear Control*, vol. 26, no. 5, pp. 896–918, 2016.

[86] Z. Cheng, D. Yue, S. Hu, H. Ge, and L. Chen, "Distributed event-triggered consensus of multi-agent systems under periodic DoS jamming attacks," *Neurocomputing*, vol. 400, pp. 458–466, 2020.

[87] Y. Xu, M. Fang, Z.-G. Wu, Y.-J. Pan, M. Chadli, and T. Huang, "Input-based event-triggering consensus of multiagent systems under denial-of-service attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 4, pp. 1455–1464, 2018.

[88] Z. Feng and G. Hu, "Secure cooperative event-triggered control of linear multiagent systems under DoS attacks," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 3, pp. 741–752, 2019.

[89] D. Ding, Z. Wang, D. W. Ho, and G. Wei, "Observer-based event-triggering consensus control for multiagent systems with lossy sensors and cyber-attacks," *IEEE Transactions on Cybernetics*, vol. 47, no. 8, pp. 1936–1947, 2016.

[90] M. Taheri, K. Khorasani, I. Shames, and N. Meskin, "Mitigation and resiliency of multi-agent systems subject to malicious cyber attacks on communication links," in *2020 IEEE Conference on Control Technology and Applications (CCTA)*, 2020, pp. 857–862.

[91] A. Polyakov, D. Efimov, W. Perruquetti, and J.-P. Richard, "Implicit lyapunov-krasovski functionals for stability analysis and control design of time-delay systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, pp. 3344–3349, 2015.

[92] T. Iwasaki and S. Hara, "Generalized KYP lemma: unified frequency domain inequalities with design applications," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 41–59, 2005.

[93] J. Liu, J. L. Wang, and G.-H. Yang, "An lmi approach to minimum sensitivity analysis with application to fault detection," *Automatica*, vol. 41, no. 11, pp. 1995–2004, 2005.

[94] K. S. Narendra and S. S. Tripathi, "Identification and optimization of aircraft dynamics." *Journal of Aircraft*, vol. 10, no. 4, pp. 193–199, 1973.

[95] M. Saif and Y. Guan, "A new approach to robust fault detection and identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 3, pp. 685–695, 1993.

[96] J. Chen, R. J. Patton, and H.-Y. Zhang, "Design of unknown input observers and robust fault detection filters," *International Journal of control*, vol. 63, no. 1, pp. 85–105, 1996.

[97] A. J. Gallo, M. S. Turan, F. Boem, T. Parisini, and G. Ferrari-Trecate, "A distributed cyber-attack detection scheme with application to DC microgrids," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3800–3815, 2020.

[98] Y. C. Paw, "Synthesis and validation of flight control for UAV," *University of Minnesota*, 2009.