FAULT DETECTION SCHEMES FOR DYNAMICAL SYSTEMS

Lamine Mohamadi

PhD

2018

FAULT DETECTION SCHEMES FOR DYNAMICAL SYSTEMS

Lamine Mohamadi

A thesis submitted in partial fulfilment
of the requirement of the
University of Northumbria at Newcastle
for the degree of
Doctor of Philosophy

Research undertaken in the
Faculty of Engineering and Environment
Northumbria University

February 2018

Abstract

In the current digital age where automatic controlled systems are used in many elds such as industrial plants, means of transports and domestic electronics applications, the issue of safety and reliability has become of paramount importance. With that came the need to develop techniques to be implemented in control systems that would allow monitoring those systems and detect if some malfunctions or abnormalities are occurring. Fault detection (FD) emerged as one of the most widely used solutions to this issue and the so-called model based fault detection has received a lot of attention. In this approach, a model of the target system is involved to estimate the expected output of the system under healthy condition and then a fault can be detected by comparing the actual measured output to the estimated healthy output. By making use of the state estimation capability of observers, various observer-based fault detection schemes have been proposed to estimate the system output for the purpose of fault detection.

However, it is worth noting that, state observers have been designed for state estimation and have found use in being adapted for fault detection. In order to try to answer the increasing requirements on systems performances, this thesis focuses on developing observers to be used specifically in FD schemes and for which a systematic way to be designed is proposed.

The first solution based on proportional integral (PI) observer, relies on integrating the systems output to construct an augmented model. This technique has a double effect on achieving better fault detection performances. Indeed, disturbances effect is reduced by using a technique based on replacing part of the model information in the observer design.

Besides, it allows having an additional degree of freedom when optimising the observer design to detect faults.

The second solution to the FD problem, a new type of observers, referred to as output observer is proposed for fault detection in both linear and nonlinear systems, where unnecessary state estimation in observer-based fault detection can be avoided. First, an input/output system representation, upon which the output observer design is based on, is introduced. Then, a new approach of output observer design, in which only the output variables are estimated, is developed. The convergence of the observer with respect to arbitrary initial conditions is proved and the fault detectability capabilities of the scheme are established. Another benefit of the proposed output observer design is the output injection feature, where the measured output is directlly injected in the observer so as to linearise the estimation error dynamics. This feature is fundamental to the solution proposed in this thesis to deal with nonlinearities in the system's model so that the inclusion of those nonlinearities when tuning the observer is avoided. Furthermore, as time delays are ubiquitous in systems, and are one of the most important sources of estimation errors, a solution based on output injection is also proposed to set the condition that ensures the convergence of the observer while maintaining its output estimation performances.

Acknowledgments

This doctoral research was carried out at the Faculty of Engineering and Environment at Northumbria University, Newcastle upon Tyne, UK. I would like to thank my supervisors, Dr. Xuewu Dai and Professor Krishna Busawon for giving me the PhD opportunity as well as their continuous help and guidance. I would also like to express my appreciation to my external based supervisor, Professor Mohamed Djemai for his valuable devices. As a Northumbria University sponsored PhD candidate, I would like to express my gratitude for the scholarship and the support during the PhD.

Finally, special thanks to my family for always believing in me and encouraging me throughout my studies.

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others. Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the University Ethics Committee on 11/08/2014.

Lamine Mohamadi

Contents

List of Figures

List of abbreviations

FD	Fault Detection
FTC	Fault Tolerant Control
FTS	Fault Tolerant Systems
SISO	Single-input Single-output
MIMO	Multi-Input Multi-Output
LTV	Linear Time Varying
LTI	Linear Time Invariant
LMI	Linear Matrix Inequalities
UIO	Unknown Input Observer
PI	Proportional Integral

Chapter 1

Introduction

Nowadays, the issue of safety is the primary concern of many industries. For this reason, fault detection (FD) in the systems involved is one of the most important tasks in many industrial applications. The main concern here is to identify when a fault has occurred, evaluate the type of fault, pinpointing its location as well as taking the necessary maintenance actions to clear the fault. As a result, fault detection is one of the most critical aspects of control systems design.

In order to meet actual performance and safety requirement of critical systems, various fault-tolerant control (FTC) algorithms together with the corresponding mathematical tools are developed [1]. Obviously, FD is fundamental to solving the FTC problem. In this context, model-based FD has received much more attention since it was originally introduced in the early of 1970s and these methods have been already remarkably developed in the last four decades [2]. Thanks to the advent of efficient computers and the advancement in control theory, observer-based fault detection and fault diagnosis methods have been applied in various fields such as mechanical, electrical and chemical engineering [3]. Plenty of work has already been done in this area and some of the most known contributions on model-based FD to the literature have been made by Paul M. Frank, Ron Patton and Rolf Isermann [4]. Nevertheless, there is still a lot of work to be done regarding some of the fundamentals of model-based FD. Indeed, the high-performance requirements of industrial processes has suggested the need in finding more appropriate methods for FD.

Because of the requirement for more efficient machines, high quality products, better profitability, high complexity and continuous increasing the degree of automation of the industrial processes, the safety and reliability of systems have become increasingly important. Today, one of the major issues surrounding the design of automatic systems is reliability and dependability.

A traditional way to improve reliability and operation safety of systems is to improve the quality, reliability and robustness of each component such as sensors, actuators or controllers but despite this improvement, fault in systems remain a risk that must be considered. Process monitoring and fault diagnosis are therefore increasingly integrated in a modern automatic control system. In general, these are not all the faults that lead directly to the system failure. Normally, the system is robust enough to withstand some faults and can still provide services in degraded mode. When a fault occurs in a fault tolerant system, a sequence of actions occurs to bring the system in a stable condition such as:

- **Fault detection:** To identify whether a fault happens in the system.

- **Fault isolation and identification:** To determine the location of the fault and to decide the nature of the fault.

- **Fault accommodation:** To reconfigure control laws.

As a result, a disaster or a failure can be avoided.

One of the main issues of current model-based fault detection schemes is that the observers employed are full-state observers. However, for the purpose of fault detection only, these types of observers are not suitable since only a comparison between the real output of the system and the estimated output is carried out.

So, in order to enhance fault detection capabilities of the fault detection scheme, the identified approach is to design observers specifically for fault detection in systems. In those observers, the main target will be that the observer structure is optimised toward estimating the system's output rather than estimating unnecessary states for fault detection, while reducing the complexity of the observer design. Thus, observer tuning for fault detection is simplified and better performances can be achieved.

Furthermore, although successive integrals cannot be avoided in observer design, we aim to reduce the system complexity by avoiding to estimate states that are combinations of those integrals.

1.1 Scope of the thesis

There are several methods for fault detection such as signal based, knowledge based and model based. The scope of this thesis fall in the area of model-based fault detection

Model-based fault detection employs state estimators/observers to generate a residual signal which carries the information of the fault. The main focus of this thesis is to redesign the observers employed for fault detection with the aim of improving efficiency, the computation time for generating the residual and detecting the faults.

The problem of fault detection under disturbances for multi-input and multi output linear and nonlinear systems is investigated. The considered fault detection approach relies on the developed proportional integral and output observers. A new algorithm is proposed to calculate the gain of the observers with the purpose of optimising the fault detection performances. The proposed approach based on the output observer is also applied to systems with time delay.

Application is made to a direct current (DC) motor as well as a three wheeled robot model in order to show the performance of the proposed fault detection schemes.

1.2 Contributions

As a result, the main contributions of this thesis are:

- A PI observer theory is developed for the first time for MIMO systems. Then, an optimised PI based fault detection scheme for MIMO linear systems under disturbances is designed.

- The output observer design methodology is developed. The basic steps towards designing the observer for linear and nonlinear, SISO and MIMO systems are detailed and the observer estimation error asymptotic convergence is proved. Then, the proposed observer is applied to design a fault detection scheme for systems under disturbances. Finally, the proposed observer is also applied as a solution to deal with time delay in interconnected systems.

1.3 Thesis structure

This thesis is divided into 7 chapters:

- Chapter 1 mainly introduces the scope of the thesis. Also, the aim, objectives, key contributions and the structure of the thesis are presented
- Chapter 2 gives an overview of dynamical systems and model-based fault detection schemes and presents a literature review on model-based fault detection applied to general dynamical systems
- Chapter 3 presents PI observer development for fault detection in MIMO systems
- Chapter 4 is devoted to develop output observer design for linear and nonlinear systems
- Chapter 5 proposes a fault detection method based on output observers for SISO and MIMO systems
- Chapter 6 is dedicated to a solution using output observer for time delays in systems
- Finally, conclusions are drawn and future works are discussed in Chapter 7.

Chapter 2

Overview on model-based fault detection

2.1 Introduction

Fault detection (FD) is the process of monitoring a system in order to identify a significant change that is indicative of a developing fault. The advantage of fault detection is its ability of reducing system's unplanned down time by detecting a fault early before the system suffers severe damage and has to be shutdown. Then preventive maintenance can be carried out and system time being out of work is considerably reduced. As a result, FD has been playing an important role in the industrial processes that have restricted safety requirements. In practice, the most frequently used FD method is the model based fault detection.

The most important element model based FD is the model that is built using dynamical systems that describe the monitored plant. Dynamical systems are systems that describe the evolution in time of state variables of a physical system. As such, they are used to model various systems in diverse engineering and scientific fields ranging from biology, to economics and chemical processes, to mention a few.

This chapter will mostly concentrate on describing the mathematical proprieties of linear and nonlinear systems that allow to study and classify dynamical systems. It will also, introduce fault detection, and will investigate the current state of the art on observers and fault detectors designs.

2.2 Classification of dynamical systems

The purpose of this section is to classify different types of dynamical systems based on their structure and their distinct proprieties and features as in [5], [6], [7], [8]. Dynamical systems can be classified into :

- Continuous-time dynamical systems;
- Discrete-time dynamical systems.

Several academic examples are given to illustrate various proprieties of continuous and discrete dynamical systems .

2.2.1 Classification of continuous-time dynamical systems

A continuous-time dynamical system (or simply a continuous system), as the name implies, is a system in which the time variable, $t \in R$ (set of real numbers), is continuous and its state variables values change over time. As a result, these systems are described by differential equations of various types.

General nonlinear systems

A general nonlinear system can be described as follows:

$$\begin{cases} \dot{x}(t) = f(x, u) \\ y(t) = h((x, u) \end{cases} \tag{2.1}$$

where $x \in R^n$ denotes the n-dimensional state vector, $y \in R^p$ denotes the p-dimensional output vector and $u \in R^m$ denotes the m-dimensional input vector. The functions $f : R^n \times R^m \to R^n$ and $h : R^n \times R^m \to R^p$ are supposed to be smooth functions (i.e. f and h are of class C^∞).

Example 1 *Take an example of general systems:*

$$\begin{cases} \dot{x}_1(t) = x_1^2(t)u(t) + \sin(u)x_2(t) \\ \dot{x}_2(t) = \ln(x_3) + x_2(t)u^2(t) \\ \dot{x}_3(t) = u(t)\cos(x_2) \\ y(t) = u(t)\exp(x) \end{cases} \tag{2.2}$$

Here $x \in R^3$, $y \in R$ *and* $u \in R$ *with*

$$f(x,u) = \begin{pmatrix} x_1^2(t)u(t) + \sin(u)x_2(t) \\ \\ \ln(x_3) + x_2(t)u^2(t) \\ \\ u(t)\cos(x_2) \end{pmatrix} \tag{2.3}$$

$$h(x,u) = u(t)\exp(x)$$

Control Affine Systems

Roughly speaking, affine means "linear". Therefore, a nonlinear system in which the control, u, appears linearly is called control affine nonlinear system or control affine system. So control affine systems are a particular case of the general nonlinear system as described in Equation (2.1) and are written as follows:

$$\begin{cases} f\left(x,u\right) = f\left(x\right) + \sum_{i=1}^{m} g_i\left(x\right) u_i(t) \\ \\ h\left(x,u\right) = h\left(x\right) \end{cases} \tag{2.4}$$

with f and g_i being smooth functions. The most commonly used form of control affine systems is as follows:

$$\begin{cases} \dot{x}(t) = f\left(x\right) + g\left(x\right) u(t) \\ \\ y(t) = h\left(x\right) \end{cases} \tag{2.5}$$

Example 2 *To be specific, take this simple example to represent control affine systems:*

$$\begin{cases} \dot{x}_1(t) = x_1^2(t)u_1(t) + x_2(t) + 3u_2(t) \\ \dot{x}_2(t) = \sin(x_1(t) + x_2(t)) + u_2(t)\cos(x_1(t)) \\ y_1(t) = x_1(t) \\ y_2(t) = x_1^2(t) + x_2(t) \end{cases} \tag{2.6}$$

Here $x \in R^2$, $y \in R^2$ *and* $u \in R^2$ *with*

$$\begin{aligned} f(x) &= \begin{pmatrix} x_2(t) \\ \sin(x_1(t) + x_2(t)) \end{pmatrix} \\ g(x) &= \begin{pmatrix} x_1^2(t) & 3 \\ 0 & \cos(x_1(t)) \end{pmatrix} \\ h(x) &= \begin{pmatrix} x_1(t) \\ x_1^2(t) + x_2(t) \end{pmatrix} \end{aligned} \tag{2.7}$$

State affine systems

State affine systems as the name implies are systems in which the state, x, appears linearly. They are a particular case of the general nonlinear systems (2.1) whereby:

$$\begin{cases} f\,(x, u) = F\,(u, y)\,x(t) + g\,(u, y) \\ h\,(x, u) = H\,(u, y)\,x(t) \end{cases} \tag{2.8}$$

The system described above can be re-written in a commonly used form as follows:

$$\begin{cases} \dot{x}(t) = A(u, y)x(t) + g\,(u, y) \\ y(t) = C\,(u)\,x(t) \end{cases} \tag{2.9}$$

where A and C are matrices of appropriate dimensions, and g is a smooth function in u and y.

Example 3 *Consider the following example of state affine system:*

$$\begin{cases} \dot{x}_1(t) = x_1(t)\cos(u) + u(t)x_2(t)x_3(t) + 3u^2(t)x_2(t) \\ \dot{x}_2(t) = 3x_1(t) + 4u(t)x_2^2(t) + 5\sin(u) \\ \dot{x}_3(t) = 2x_2(t) + x_3(t) \\ y(t) = u(t)x_2(t) \end{cases} \tag{2.10}$$

which can be rewritten as:

$$\begin{cases} \dot{x}_1(t) = x_1(t)\cos(u) + y(t)x_3(t) + 3u(t)y(t) \\ \dot{x}_2(t) = 3x_1(t) + 4y(t)x_2(t) + 5\sin\,(u) \\ \dot{x}_3(t) = 2x_2(t) + x_3(t) \\ y(t) = u(t)x_2(t) \end{cases} \tag{2.11}$$

where

$$A(u,y)=\begin{pmatrix}\cos(u) & 0 & y(t)\\ 3 & 4y(t) & 0\\ 0 & 2 & 1\end{pmatrix}$$

$$g(u,y)=\begin{pmatrix}3u(t)y(t)\\ 5\sin(u)\\ 0\end{pmatrix} \tag{2.12}$$

$$C(u)=\begin{pmatrix}0 & u(t) & 0\end{pmatrix}$$

Bilinear Systems

Bilinear systems are special case of state affine system (2.9) where:

$$\begin{cases} A(u,y)=A+u(t)B \\ g(u,y)=D \\ C(u)=C \end{cases} \tag{2.13}$$

More precisely,

$$\begin{cases} \dot{x}(t)=Ax(t)+u(t)Bx(t)+Du(t) \\ y(t)=Cx(t) \end{cases} \tag{2.14}$$

where A, B, D and C are the matrices of the system's parameters with appropriate dimensions.

Example 4 *An example of bilinear system is given by*

$$\begin{cases} \dot{x}_1(t) = -2x_1(t) - x_2(t) + 2x_1(t)u(t) + u(t) \\ \dot{x}_2(t) = -3x_2(t) + 5x_2(t)u(t) + 3u(t) \\ y_1(t) = x_1(t) - 6x_2(t) \\ y_2(t) = x_1(t) + 2x_2(t) \end{cases} \tag{2.15}$$

where $x \in R^2$, $y \in R^2$, $u \in R$,

$$\begin{array}{cc} A = \begin{pmatrix} -2 & -1 \\ 0 & -3 \end{pmatrix} & B = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \\ D = \begin{pmatrix} 1 \\ 3 \end{pmatrix} & C = \begin{pmatrix} 1 & -6 \\ 1 & 2 \end{pmatrix} \end{array} \tag{2.16}$$

Linear time invariant (LTI) systems

Linear time invariant systems are a particular case of bilinear systems expressed as follows:

$$\begin{cases} \dot{x}\,(t) = Ax\,(t) + Bu\,(t) \\ y\,(t) = Cx\,(t) \end{cases} \tag{2.17}$$

where A, B and C are the matrices of the system parameters of appropriate dimensions.

Example 5 *The following is an example of a 2nd order LTI system:*

$$\begin{cases} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 1 \\ -3 \end{pmatrix} u(t) \\ y(t) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \end{cases} \tag{2.18}$$

Linear time varying (LTV) systems

When the matrices in (2.17) varies with time, the following linear time varying system is obtained:

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ y(t) = C(t)x(t) \end{cases} \tag{2.19}$$

Example 6 *The following gives and example of a time varying system:*

$$\begin{cases} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} e^{-t} & 2t \\ \sin(t) & -t^2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} t \\ -3t \end{pmatrix} u(t) \\ y(t) = \begin{pmatrix} \cos(t) & t \end{pmatrix} x(t) \end{cases} \tag{2.20}$$

Remark

In all the above equations, the issue of delays was not considered. In practice, there may be delay in every part of the system, either in the state or in the input or output of the system. For example, the following is an example of a linear delayed system in the state:

$$\begin{cases} \dot{x}(t) = A_1 x(t) + A_2 x(t-\tau) + Bu(t) \\ y(t) = Cx(t) \end{cases} \tag{2.21}$$

2.2.2 Classification of discrete-time dynamical systems

Discrete-time dynamical system (or simply a discrete system) as its name implies are systems in which the time variable is discrete and is generally denoted by $k \in Z$ (set of integers). As a result, the evolution of the state of a discrete time system is described by a difference equation. In what follows, a brief classification of discrete-time dynamical systems is given. Basically, all the above continuous-time dynamical systems can be written in a discrete setting.

General discrete nonlinear systems

General discrete-time systems described as follow:

$$\begin{cases} x[k+1] = f(x[k], u[k]) \\ y[k] = h(x[k], u[k]) \end{cases} \tag{2.22}$$

where $k \in Z$, $x \in R^n, y \in R^p$and $u \in R^m$ with f and h being smooth functions.

Control affine discrete systems

Likewise control affine continuous systems, a control affine discrete system is described as follows:

$$\begin{cases} x[k+1] = f(x[k]) + g(x[k])\, u[k] \\ y[k] = h(x[k]) \end{cases} \tag{2.23}$$

State affine discrete systems

In the discrete case, state affine systems are generally expressed as follows:

$$\begin{cases} x[k+1] = A(u[k], y[k])x[k] + g\left(u[k], y[k]\right) \\ y[k] = C\left(u[k], y[k]\right)x[k] \end{cases} \tag{2.24}$$

where A and C are matrices of system's parameters with appropriate dimensions, and g is a smooth function in u and y.

Bilinear Systems

Discrete bilinear systems expressed as follows:

$$\begin{cases} x[k+1] = Ax[k] + Bu[k]x[k] + Du[k] \\ y[k] = Cx[k] \end{cases} \tag{2.25}$$

where A, B, C and D are the matrices of system's parameters with appropriate dimensions.

Linear time invariant discrete systems

In the discrete case, linear systems can be represented by the following difference equation:

$$\begin{cases} x[k+1] = Ax[k] + Bu[k] \\ y[k] = Cx[k] \end{cases} \tag{2.26}$$

where A, B and C are the matrices of the system's parameters with appropriate dimensions.

Linear time varying (LTV) discrete systems

Discrete time LTV systems can be described by the following equation:

$$\begin{cases} x[k+1] = A[k]x[k] + B[k]u[k] \\ y[k] = C[k]x[k] \end{cases} \tag{2.27}$$

Remark

As for continuous systems, time delays in discrete systems may occur in all parts of the discrete time systems described above. So Equation (2.21) can written in a discrete time case as follows:

$$\begin{cases} x\left[k+1\right] = \bar{A}x\left[k\right] + B_0 \Delta u_k + \bar{B}u\left[k\right] \\ y\left[k\right] = Cx\left[k\right] \end{cases} \tag{2.28}$$

where

$$\begin{aligned} \bar{A} &= e^{A(t_{k+1}-t_k)} \\ B_0 &= \int_{t_k}^{t_k+\tau(k)} e^{A(t-\tau(t_k))} d\tau B \\ B_1 &= \int_{t_k+\tau(k)}^{t_{k+1}} e^{A(t-\tau(t_k))} d\tau B \\ \bar{B} &= B_0 + B_1 \\ \Delta u_k &= u\left[k-1\right] - u\left[k\right] \end{aligned} \tag{2.29}$$

each sampling k occurs at time t_k and $\tau\left(k\right)$ represents the delay value associated to it.

Note, that for sake of simplicity, no delay has been considered on the output, but the same procedure can be applied to include it in Equation (2.28). Moreover, as it will be discussed in Chapter 6, a total delay which represents the sum of the delay on the input and the output can be considered.

2.3 Overview of fault-detection in systems

Fault detection is naturally the first step of fault diagnosis and also the most important element in fault tolerant systems (FTS). FTS are systems that can operate to a certain extent under considered faults scenarios. In most fault detection methods, the idea is to generate an information redundancy in order to evaluate the health of the system. Generally, information redundancy can be classified into two categories:

- Physical redundancy. A traditional approach is physical redundancy based fault diagnosis. The idea of this scheme is to generate information redundancy using at least two redundant physical devices and a typical example is that a suspension bridge's numerous cables are a form of physical redundancy. The main advantage of this approach is its high degree of reliability, but it faces, on the other hand, the problem of extra hardware costs and additional weight and space.

- Analytical redundancy. This method uses a mathematical model to describe the actual system's behavior. The model is the mathematical replicate of the physical model and is supposed to generate the same data as the actual system. In practice, unmodelled dynamics/modelling inaccuracies and initial conditions differences result in a difference between the system measured and the model estimated outputs, thus, the challenge that faces analytical redundancy fault detection is to distinguish between these differences and those induced by a fault occurring in the system.

Despite of extra hardware costs and additional weight and space, the physical redundancy is still applied in some fields such as nuclear power station, aircraft and safety-critical systems in general. In these methods, redundant physical devices are usually sensors, actuators or critical components. The fault detection is accomplished by a majority voting among redundant hardware. For example, fly-by-wire and hydraulic systems in aircraft and redundant emergency electrical systems in nuclear plant. Physical redundancy fault detection scheme have already possessed different mature technologies. However, in presence of its constraint, physical redundancy is difficult to implement in certain systems (For example, in micro-robots and some components of aircrafts).

In contrast to physical redundancy, analytical redundancy has its own advantages. One of these advantages is no additional redundant hardware. However, the analytical redundancy fault detection scheme is not yet a mature technology and there are still many open research topics in this area. Furthermore, several performances of fault detection system have to be improved like: sensitivity, robustness and rapidity of detection. As a result, more innovative methods are needed to meet the actual requirement.

2.3.1 Faults in dynamical systems

In the following, the definition, the categories and the nature of faults in dynamical systems are given.

Definition of fault

The term fault in the area of fault detection and diagnosis is generally defined as a departure from an acceptable range of an observed variable or a calculated parameter associated with a process [9].

Categories of faults

When developing a fault detection algorithm, it is important to know what kind of fault appears in the system and what are its possible effects on the system. Generally, fault can be classified into two categories [10]:

- Abrupt fault is fault whereby a system's component undergoes a sudden change in its value from normal into abnormal [11].
- An incipient fault is a slow growing fault which might eventually lead to a catastrophic situation in the system [12], [13], [14], [15], [16] and [17].

As it is well-known, the methods designed for the detection of abrupt faults may be not suitable for incipient faults, and vice versa [18]. Moreover, an abrupt fault is a discrete event and an incipient fault is a continuous event and both types of faults can be present at the same time.

Nature of faults

Due to the way how the faults affect the dynamical systems, in general, two ways of fault that can be dealt with are described below:

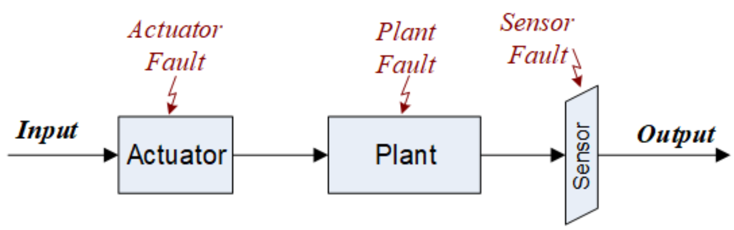

Figure 2.1: Additive faults and disturbances sources in model based fault detection systems

- Additive faults (unknown additive inputs)

The effect of additive faults can be represented as an additional unknown input vector acting on sensors, actuators or states of system as shown in Fig 2.1. These faults are usually due to a constant bias (positive or negative).

- Multiplicative faults (structural changes)

Multiplicative faults refer to changes in the process itself or malfunctions in the sensors and actuators. They occur due to hard failures in components of dynamical systems. Consequently, they often lead to changes in the parameters. It is important to note that multiplicative faults may influence the system stability [19].

Based on the characteristics of faults, the complexity of dynamical system and the performance requirements, different fault detection techniques will be applied. In the next subsection, a simple classification of fault detection approaches is given.

2.3.2 Fault detection approaches

In order to reduce human intervention and design reliable automated systems, a monitoring system become a crucial factor in industrial processes. Thanks to the advancement of modern control theory, FD has been developed extensively in different areas. Fault detection methodologies are mostly based on information redundancy. As mentioned above, two main categories of information redundancy can be distinguished: Physical redundancy and Analytical redundancy.

The traditional physical redundancy requiring the system to be equipped with redundant physical devices suffers the problem of extra hardware inconveniences. Thus analytical redundancy that uses causal relationships or mathematical relationships between the signals to verify whether a fault occurs in the system is more adapted to most of systems except those that require the highest degree of reliability and where extra cost and maintenance can be afforded while additional space and weight margin to accommodate the equipment is available. Analytical redundancy based FD can be classified into three main approaches:

Knowledge-based approach

The early knowledge-based fault detection systems were originally expert systems. The methods of expert systems are used when the knowledge of the system is heuristic. The expert system relies on its knowledge and experience. Moreover, expert system is capable of solving specific issues in the same way that a human expert. Because expert systems can manipulate a large number of non-homogeneous and independent context, this approach is very attractive for diagnosis. Over the last few years, the knowledge-based fault detection systems became more complex and also more intelligent. Then, the knowledge base of diagnosis systems became more and more sophisticated. Recently, more intelligent algorithms have been integrated into this approach (such as machine learning, data mining and pattern recognition etc...). Consequently, diagnosis systems not only receive knowledge from experts but also acquire knowledge themselves by receiving data of similar experiences. Generally, this approach is very attractive for large systems like nuclear power stations, chemical plants and some special functions of airplanes use this method to support operators to detect malfunctions. One of the reasons is that the model of these systems is difficult to be formalised by mathematical equations. But a large amount of historical process data is required.

Signal-based approach

Signal based approach assumes that certain measurements contain the information about a fault symptoms. In signal processing based fault detection approaches, some mathematical and statistical treatments are necessary to interpret this fault symptoms from measurements. The prerequisites for this method is a knowledge about the relationship between signal variations and faults. In practice, the fault effects are classified into two types:

- Time domain function such as: magnitudes, means, covariances, amplitude envelope, correlation coefficients and time domain reflectometry.
- Frequency domain function such as: spectral power densities, frequency spectral lines,

spectrums, etc..

Signal-based approaches have been widely applied to mechanical engineering (e.g., vibration monitoring), electric motors, etc. The signal-based approach is mainly designed for condition monitoring at the steady state.

Model-based approach

In this approach, a mathematical model is used to describe the behaviors of the system. Generally, an analytical model can be classified into two types: quantitative model and qualitative model. Quantitative model describes a system's behavior in quantitative mathematical terms. Qualitative model describes a system's behavior in qualitative terms such as causalities. In the following section, quantitative model-based methods will be investigated. The (quantitative) model-based fault detection was initialised in the early 1970s.

The basic idea of model based fault detection is to generate a residual between the mathematical model and the real system. In this context, the residual is defined as the difference between the actual measurements of the system and the values estimated by the mathematical model. Under ideal conditions, the mathematical model is assumed to describe exactly the behavior of the real system. In the presence of a fault, the residual is different from zero and the diagnosis system can detect the latter. However, it is unrealistic that the residual is different from zero only when a fault occurs. Sensor noise, disturbances, parametric variations, unmodelled dynamics and nonlinearities affect the amplitude of the residual. A lot of existing methods can reduce the impact of these effects on the residual.

Unfortunately, neither of them can give a perfect decoupling between these effects and the residual. As a result, the residual has to be compared with a threshold for decision-making.

2.4 Model based fault detection methods

In this section, a classification of the existing fault detection methodologies is given.

Four main methods have been developed in the literature:

- Parameter estimation;
- Frequency-based;
- Parity relation;
- Observer-based.

The typical works are globally summarised in the book [20] by Ding. During the past four decades, many works of model-based fault detection have been conducted. The representative survey papers have been written by Isermann, Patton, Frank and Ding such as [4], [21], [18] [22]. They have given a good state-of-the-art of modern model-based fault detection. Based on these four approaches, different papers have been published to tackle the robustness issues. Indeed, reduction in false alarms and non-detection become a critical issues in diagnosis schemes. The objective is to generate a residual that is robust against disturbances and sensitive to faults. Another strategy is an adaptive threshold scheme. In this scheme, the threshold is a time domain function and its value changes under different situations.

2.4.1 Parity space relation method

In the field of diagnostic, parity equations represent a mathematical tool for detecting and locating faults. In the review of parity space [23], an equation that generates a residual is called parity equation. The parity equation can be obtained from the model of the system. This method uses signal analysis to check the consistency between the measurements and values calculated by a model. In the literature, parity relation based residual generators are often called open-loop structured [20], because the signal of residual does not interfere with the residual generation.

2.4.2 Frequency-based method

Frequency domain fault detection relies on using frequency transformation methods such as Fourier transform to generate a signal in the frequency domain. Then condition monitoring of systems can be carried out by analysing the measured output and the system's generated signal. This approach had a particular success in domains such as mechanical structures by using vibrations analysis,and electrical machines rotors by analysing the stator's current [24].

2.4.3 Parameter estimation method

The diagnosis of an industrial system can be done by monitoring the evolution of its structural parameters. Parameter estimation-based fault detection has been applied in different areas of science and engineering. Consider the simple case of a linear first order system defined by its gain and its time constant. If the values of these two parameters remain constant over time, it indicates that the input-output relationship remains unchanged. In this situation, the system is in normal operation. If one of the parameters changes its value, the input-output has relationship changed. In this situation, the system has a fault. The magnitude and variation in these parameters are the useful indication to accomplish the diagnosis task.

When fault detection algorithm is based on the knowledge of system's parameters, it is necessary to calculate their values from the information of the input and the output of the system. Identification is one of the tools to acquire this knowledge.

In parameter estimation approach, there are a number of system identification techniques such as recursive Bayesian estimation, maximum likelihood estimation and least squares. These approaches execute a data processing procedures to achieve a non biased parameter estimation. The first works of parameter estimation-based fault detection were made by R. Isermann [25]. His survey paper summarises different parameter estimation based FD approaches. Observer-based on-line parameter identification was used in [26]. In this work, the estimation of critical parameters is done by a high-gain observer.

However, parameter estimation fault detection approach might be difficult to implement like an online real-time algorithm due to large amount of computations.

2.4.4 Observer-based method

One of the methods in model based FD is the observer-based fault detection, which is also the focus of this work. The observer-based fault detection approach is based directly on observer design and control theory. The very first observer is the so-called Luenberger observer designed by D. G. Luenberger in 1966 [27]. Because of certain constraints, certain states of system cannot be directly measured by a sensor. The observer is used as software sensors to estimate the unmeasured state of dynamical system. However, Luenberger observer can only deal with deterministic systems. In a fault detection scheme, it is important to note that the observer is used as output observer to reconstruct the output of a system.

In control theory community, different applications of observer in fault detection problem have been proposed since 1980's and a book on observer-based fault detection have been written by Patton and Frank [28]. This book summarises globally the observer-based fault detection techniques. For example, Luenberger observers for deterministic systems, Kalman filters for stochastic systems. Moreover, interval observers are used to deal with systems under uncertainty [29]. The observer-based residual generators is also called closed-loop structured [20], because the signal of residual is fed back to the residual generator.

In practice, the main tasks of observer-based FD are the generation of a residual and the calculation of a threshold value. To combat these two problems, the robustness techniques have become the main theme in the last two decades. The robustness techniques can be classified into two types such as active robustness and passive robustness. Active robustness is to generate a desirable residual which should be robust against disturbances, but sensitive to the faults. fault detection filter, unknown input observer [30], signal norms with LMI aided design [31] are usually used in active robustness scheme. Passive robustness is to choose a suitable threshold for fault detection. The methods like fuzzy logic [32], adaptive control [33], [34] and neural networks [35] are used to choose the threshold. Combining active robustness and passive robustness, the expected results are when the fault is present, the residual is greater than the threshold. When the fault is absent, the residual is less than the threshold.

Most of the literature in fault detection problems is focused on the active robustness. The most common method for fault detection is to treat model uncertainties as some unknown additive inputs. The idea is to decouple them from the residual. This makes the residual robust against unknown inputs. Eigenstructure assignment [36] and unknown input observer [37] are principally used to decouple the residual from unknown additive inputs. Eigenstructure assignment method parameterises the observer gain matrix and Unknown Input Observer method searches for the disturbance decoupling matrix. It is important to know that the non-unique solution allows the fault detection observer design in these two methods. Due to the lack of design freedom, the condition for complete decoupling is generally difficult to be satisfied. For this reason, different control techniques are applied to enable the optimal fault detection observer design like: LMI, signal and system norms and neural networks. The idea is to minimise the effect of disturbances to the residual and maximise the effect of faults to the residual. Thus, the fault detection design becomes an optimisation problem:

- Analysis of system constraints,
- Performance criterion selection,
- Optimisation techniques selection.

Nevertheless, parametric uncertainties are not resolved in the previous methods. As a result, the interval observer is used to deal with parametric uncertainties. In this approach, a model with parameters bounded in intervals is considered. Moreover, the evolution of estimated states (or output) at every time will not be described by a point in

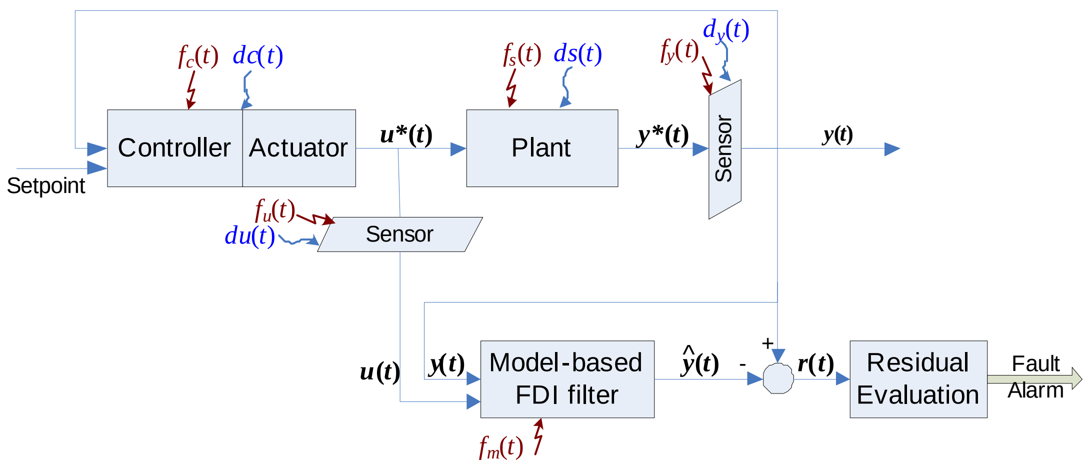

Figure 2.2: Observer based fault detector

the state space but by a region [29]. Due to the effect of the parametric uncertainties to the residuals, a suitable threshold has to be selected.

2.5 Observer based fault detection

The basic idea of observer-based fault detection is to estimate the outputs of the system and generate a residual as depicted in Fig 2.2. To calculate an estimated output $\hat{y}(t)$, the model uses $u^*(t)$ that is the plant actual input $u(t)$ subject to sensor's disturbances $d_u(t)$ and fault $f_u(t)$.and the plant measured output $y(t)$ that is the plant actual output $y^*(t)$ also subject to sensor's disturbances $d_y(t)$ and fault $f_y(t)$. The plant measured output $y(t)$ is also used by the controller to calculate the command signal value $u^*(t)$ which makes the system behavior meets with the setpoint value. For this, the controller needs to be robust to controller/actuator disturbances $d_c(t)$ and fault $f_c(t)$. Finally the residual $r(t)$ that is the difference between the measured and estimated output is generated. The residual should be close to zero in absence of fault and if a fault occurs, it should be detected by the embedded evaluation algorithm.

In this approach, the process model is replaced by an observer. Using the observer theory and mathematical optimisation tools, the designer can achieve the desired decoupling between different signals. The residual signal should be robust against disturbances and sensitive to faults. In practice, the residual has to be compared with a threshold for a decision-making. In this section, some existing faulty system modelling are given and two typical examples of observer-based fault detection are explained.

Nowadays, automatic control design possesses mainly a feedback loop. Because of the feedback structure, the controller gives a certain level of robustness to the system. But critical changes in the system cannot be simply covered by a controller. Consequently, Fault Diagnosis algorithm and Fault Tolerant Control (FTC) are often embedded in automatic control process. Before developing Fault Diagnosis algorithm, modelling of faults in automatic control system is a vital issue. Choosing a way of modelling the system depends on the signal availability. Two frameworks can be considered such as:

- Open loop fault detection

In this framework so-called open loop fault detection, the control input u and the output y are considered to be accessible. In order to access the command u, fault diagnosis algorithm has to be embedded locally in automatic control process. These two signals contain all the information to accomplish the fault detection task.

- Feedback loop fault detection

In this method, the so-called feedback loop fault detection framework, the input u and the output y are considered to be accessible. In practice, the command u is not always accessible. For example, if the control loop is a part of a large scaled system and located remotely from the supervision station, where the higher level controller and FD unit are located, the reference signal will be applied. The closed loop FD strategy is based on the closed loop model with u and y as input and output signals respectively [20].

Observer-based fault detection is normally based on feedback loop fault detection. It is important to note that a full state x estimation is not required in the FD case but rather an observer to estimate the output only or simply an output observer. In the literature, many papers on FD treat the observer-based fault detection problems by deriving a model in which the fault enters the system as an additive unknown input with no apparent relation with the system's dynamics [38], [39], [30], [11], [40], [41], [42]. In effect, these papers assume that the command vector u is available and the fault detection algorithm is based on open loop fault detection. More precisely, much of the academic research on observer-based FD is still based on the linear case and using the following model:

$$\begin{cases} \dot{x}\left(t\right) = Ax\left(t\right) + Bu\left(t\right) + E_d d\left(t\right) + E_f f\left(t\right) \\ y\left(t\right) = Cx\left(t\right) + G_d d\left(t\right) + G_f f\left(t\right) \end{cases} \tag{2.30}$$

where the vector d represents unknown inputs and the term f denotes the faults with matrices E_d, E_f, G_d, G_f derived from system's proprieties. Using the model as in Equation (2.30), different observer-based strategies are used to generate the residual; namely a standard Luenberger type observer or an unknown input observer.

2.5.1 Observability

The first step before designing an observer for a system is to determine wether the system is actually observable. Indeed, all systems don't posses the observability property required to build an observer although a non-observable system can posses a weaker property that is detectability if the non-observble modes of the systems are stable.

There exist many types of observabilities translating the possibility of reconstructing the entire state vector using the measured output and the system's input.

The system is said to observable if for two initial states $x(t_0) = x_1$ and $x(t_0) = x_2$, then their respective outputs $y_1(t)$ and $y_2(t)$, are not identical for all t.

For linear systems, the observability condition is written in a form of an algebraic matrix rank condition.

Consider the following time invarying linear system:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \tag{2.31}$$

Then, the observability of the system can be determined if there exist a unique solution to:

$$y(t) = Ce^{At}x(0) + C\int_0^{t_0} e^{A(t-t_0)}Bu(\tau)\,d\tau \tag{2.32}$$

which considering that C, B and $u(t)$ are known, to find the necessary observability condition, is equivalent to proving there exist a unique solution to:

$$y(t) = Ce^{At}x(0) \tag{2.33}$$

which leads to the following matrix being of full order:

$$\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n1} \end{pmatrix} \tag{2.34}$$

where n is the size of the state vector x.

For nonlinear systems, the types of observabilities can correspond to locally or globally defined approaches [43], [44].

First consider the unforced general nonlinear system:

$$\begin{cases} \dot{x}\left(t\right) = f(t,x) \\ y\left(t\right) = h(t,x) \end{cases} \tag{2.35}$$

then the observability of the system can be determined if the following matrix is of full order:

$$\begin{pmatrix} h\left(t,x\right) \\ L_f h\left(t,x\right) \\ \vdots \\ L_f^{(n-1)} h\left(t,x\right) \end{pmatrix} \tag{2.36}$$

where $L_f h\left(t,x\right)$ is the Lie derivative of $h\left(t,x\right)$ in the function f.

To determine the observability of the systems with an input, there exist many methods that have been introduced either to linearise these systems as in [45], [46], [47], where in most cases observabilty proprieties are determined locally. Particularly, the output injection technique to linearise nonlinear system had success and several works have been published as in [48], [49].

2.5.2 Observers design

Observers can be considered as software sensors as they use the knowledge on the system dynamics to estimate system's variables. As such, they allow reducing the number of physical sensors used on the system. The missing information due to sensors absence is reconstructed with the remaining sensors data and the knowledge of the system. In the following some of the observers design techniques are detailed

Sliding mode observer

Sliding mode observers have known a large success in variety of systems for their efficiency, relative ease of implementation and robustness against parameter variations [50], [51], [52]. The sliding modes technique consists of using a discontinuous function as correction term and which value depends on sign of the estimation error.

Take a system which dynamics are defined by the following general nonlinear function:

$$\dot{x}\,(t) = f(x, u) \tag{2.37}$$

where f is an analytical function, $x \in R^n$ is the state vector and $u \in R^m$ is the control input.

The sliding modes observer is designed as follows:

$$\begin{cases} \dot{\hat{x}}\left(t\right) = f(t, \hat{x}, u) + KSign(y\left(t\right) - \hat{y}\left(t\right)) \\ \hat{y}\left(t\right) = h(t, \hat{x}) \end{cases}$$

where K is the gain matrix and $sign$ is a discontinuous function in time described by:

$$sign(x) = \begin{cases} x & \text{if} \quad x > 0 \\ -x & \text{if} \quad x < 0 \\ \text{Undefined} & \text{if} \quad x = 0 \end{cases} \tag{2.38}$$

The technique consists of discontinuous constraining of the systems dynamics to converge towards the so called sliding surface. The attractivity and the invariance of the sliding surface are guaranteed by the defined sliding conditions. If the conditions are met, the observer convergence will be achieved. Once within the convergence surface, the systems dynamics can be calculated by methods such as equivalent control or equivalent vector.

Unknown Input Observer

In Unknown Input Observer (UIO) scheme, the disturbances are considered as unknown inputs so that the residual signals are decoupled from them. That means the residual is only sensitive to the faults and insensitive to the disturbances. Based on the model described by Equation (2.30), the residual generator whose inputs are input u and output y of the system can be constructed as follows:

$$\begin{cases} \dot{z}(t) = Nz(t) + Mu(t) + Gy(t) \\ r(t) = Pz(t) + Uu(t) + Ky(t) \end{cases} \tag{2.39}$$

The challenge is therefore to find the matrices N, M, L, P, U, K satisfying the following conditions:

- The residual generator is built so as to estimate a linear combination of the full state (or partial state) of the system like:

$$z(t) = Tx(t) \tag{2.40}$$

When only the partial state is rebuilt, the observer is so-called reduced order observer (the dimension of z is then less than x). Let e be the estimation error of Tx so that:

$$e(t) = Tx(t) - z(t) \tag{2.41}$$

- The second objective is to decouple the residual from the disturbances and make it sensitive to the faults. The observer should be built so that the estimation error in Equation (2.40) is canceled in the absence of fault, but also in presence of disturbances. The dynamic of the estimation error is:

$$\begin{cases} \dot{e}\left(t\right) = T(Ax\left(t\right) + Bu\left(t\right) + E_d d\left(t\right) + E_f f\left(t\right)) \\ -(Nz\left(t\right) + Mu\left(t\right) + Gy\left(t\right)) \\ y\left(t\right) = Cx\left(t\right) + F_d d\left(t\right) + F_f f\left(t\right) \end{cases} \tag{2.42}$$

Combining with Equations (2.39), (2.40) and (2.41), gives:

$$\begin{cases} \dot{e}\left(t\right) = NTe\left(t\right) + (TA - NT - GC)x\left(t\right) + (TB - M)u\left(t\right) \\ +(TE_d - GF_d)d\left(t\right) + (TE_f - GF_f)f\left(t\right) \\ r\left(t\right) = -Pe\left(t\right) + (PT + KC)x\left(t\right) + Uu\left(t\right) + KF_d d\left(t\right) + KF_f f\left(t\right) \end{cases} \tag{2.43}$$

First, the eigenvalues of N should be located in the left-half complex plan. Then in order to satisfy the objectives, the following conditions should be met:

$$TA - NT - GC = 0 \text{ and } PT + KC = 0 \tag{2.44}$$

$$M = TB \text{ and } U = 0 \tag{2.45}$$

$$TE_d - GF_d = 0 \text{ and } KF_d = 0 \tag{2.46}$$

$$TE_f - GF_f \neq 0 \text{ and } KF_f \neq 0 \tag{2.47}$$

The conditions (2.44) and (2.45) ensure that the estimation error convergence to 0 in the absence of fault ($f = 0$) and disturbance ($d = 0$). The condition (2.46) ensures the decouple of the residual r from the disturbances (unknown inputs). The condition (2.47) makes the residual r sensitive to the faults. Finally, as all conditions (2.44), (2.45), (2.46) and (2.47) are satisfied, the dynamic of the residual r becomes:

$$\begin{cases} \dot{e}(t) = Ne(t) + (TE_f(t) - GF_f)f(t) \\ r(t) = -Pe(t) + KF_f f(t) \end{cases} \tag{2.48}$$

Not only the transfer function between u and r is equal to 0, but also the transfer function between d and r is equal to 0 (decouple the residual from the disturbances). The transfer function between f and r is different from 0 (the residual is sensitive to faults).

The last step is to specify how to satisfy the conditions (2.44), (2.45), (2.46) and (2.47), and synthesise an unknown input observer. There are several approaches in literature such as [30].

The transfer function between f and r being different from zero is not a sufficient condition to detect faults. When evaluating the residual, due to measurement noise and modelling error, a relevant threshold has to be fixed. A disadvantage of this scheme is that the existence conditions of UIO (matching conditions) are difficult to be satisfied in practice.

Luenberger observer (eigenstructure assignment)

A typical Luenberger type observer for the system as described by Equation (2.30) is as follows:

$$\begin{cases} \dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - \hat{y}(t)) \\ \hat{y}(t) = C\hat{x}(t) \\ r(t) = y(t) - \hat{y}(t) \end{cases} \tag{2.49}$$

where $\hat{x}$ is the estimate of the state x, $\hat{y}$ is the estimate of the output y, K is the observer gain and the residual r is the difference between the real output y and the estimated output $\hat{y}$. The estimation error is $e = x - \hat{x}$. Using Equations (2.30) and (2.49) the following error dynamic and residual functions are obtained:

$$\begin{cases} \dot{e}(t) = (A - KC)e(t) + (E_d - KG_d)d(t) + (E_f - KG_f)f(t) \\ r(t) = Ce(t) + G_d d(t) + G_f f(t) \end{cases} \tag{2.50}$$

The residual r generally takes the following form in Laplace domain:

$$r(s) = H_d(s)d(s) + H_f(s)f(s) \tag{2.51}$$

and

$$\begin{cases} H_d(s) = C(sI - A + KC)^{-1}(E_d - KG_d) + G_d \\ H_f(s) = C(sI - A + KC)^{-1}(E_f - KG_f) + Gf \end{cases} \tag{2.52}$$

Where $H_d(s)$ gives the Laplace transfer function of disturbance $d(s)$ to the residual, $H_f(s)$ denotes Laplace transfer function of faults $f(s)$ to the residual. In this case, the objective is to minimise the influences of the disturbance to the residual $r(s)$ and maximise the one of the faults. Consequently, if disturbance $d(s)$ and of faults $f(s)$ are bounded, the observer gain has to be correctly chosen to minimised the following index:

$$J = \frac{\|H_d(s)\|}{\|H_f(s)\|} \tag{2.53}$$

Where the operator $\|\bullet\|$ represents the norm function. Depending on different fault detection strategies, $H_d(s)$ and $H_f(s)$ can be described by different system norms such as:

H_∞ norm, H_2 norm, 1 norm, ∞ norm and Frobenius norm. In order to minimise the index J, the subsequent conditions should be achieved:

- Stability: The matrix $A - KC$ should be stable. It means that The eigenvalues of $A - KC$ should be located in the left-half complex plane.
- Robustness: Improve the robustness against the disturbances by minimising $\|H_d(s)\|$.
- Sensitivity: Improve the sensitivity to the faults by maximising $\|H_f(s)\|$.

2.6 Conclusion

In this chapter, a classification of dynamical systems that are relevant to the thesis topic has been given. Both continuous time and discrete time systems have been presented although this research work mainly focuses on the continuous time domain. Then, model based fault detection have been introduced and the current state of the art discussed..Also, some techniques relevant to this thesis work have been detailed. A particular focus has been given to the model based fault detection and particularly the observer based approach and a attempt to highlight the benefit and challenges of this method. In the next chapters, the bulk of the thesis research work on observer design for fault detection will be presented.

Chapter 3

Proportional Integral (PI) observer gain optimisation for fault detection with disturbance attenuation

3.1 Introduction

One of the widely used methods in observer-based fault detection is to use the knowledge on faults and disturbances dynamics to set a frequency based criterion. Then, the criterion is used to find the observer tuning that optimises the fault detector performances by signifying the fault impact in the residual while reducing the disturbances effect. This method has given good results although disturbances attenuation and fault signal amplification in the residual being optimised using the same criterion requires making a trade-off regarding the effectiveness of the FD scheme. Moreover, this method being based on the frequency knowledge of faults and disturbances, one can notice that whereas fault frequency analysis can give conclusive results considering that a FD scheme can be designed for an identified type of faults, it is more difficult to be as straight-out when it comes to disturbances analysis considering the variety of sources and its randomness over a large frequency range.

In this work, the frequency based criterion is mainly considered for fault signal amplification in the residual. The proportional integral (PI) observer method presented in [53], [54], [55], [56] is used for disturbances effect attenuation considering the originality and effectiveness it has shown to achieve the desired purpose. The technique relies on integrating the output signal to design an augmented model that allows having an additional degree of freedom when designing the observer.

In the following proposed solution, a systematic approach to tune the observer for fault detection purposes is detailed. To optimise the observer gain, an identified fault characteristic frequency is used in the criterion, while no constraints have been considered on the disturbances.

3.2 PI observer design for fault detection

In this section, a dynamic system subject to both external disturbances and possible faults is considered and a PI observer is proposed as a fault detection mechanism in an event of fault occurring in the system.

3.2.1 System model

Consider a class of linear multi input multi output (MIMO) systems under disturbances with sensor and/or actuator faults described by:

$$\begin{cases} \dot{\bar{x}}\left(t\right) = A_0\bar{x}\left(t\right) + B_0 u\left(t\right) + E_{d0} d\left(t\right) + E_{f0} f\left(t\right) \\ \bar{y}\left(t\right) = C_0\bar{x}\left(t\right) + G_{d0} d\left(t\right) + G_{f0} f\left(t\right) \end{cases} \tag{3.1}$$

where $\bar{x}(t) \in R^n$ is the state vector, $u(t) \in R^m$ the input signal and $\bar{y}(t) \in R^p$ the measured output. Signal $d(t) \in R^j$ denotes the unknown disturbance and $f(t) \in R^k$ represents the fault signal. A_0, B_0, E_{d0}, E_{f0}, G_{d0}, G_{f0} are matrices of appropriate dimensions describing the dynamics of the input and output signals in the state's $\bar{x}\left(t\right)$ dynamics equation.

First, for an accurate estimation of the state, disturbance effect needs to be attenuated in the observer. Motivated by the results that PI observer gives regarding disturbance attenuation [54], the proposed solution for disturbance attenuation in fault detection relies on integrating the output signal in order to build an augmented system in which better fault detection performances can be achieved.

3.2.2 System augmentation with integral action

In [54], it has been demonstrated that some desired robustness and disturbance attenuation performances can be achieved by augmenting the state space model. The added component to the he state vector is proportional to the integral of the measured output signal. In fact, the addition of an integral term also gives an additional degree of freedom and thus an opportunity to improve the fault detection performance. Similar to [54], an augmented system is first developed to include the additional integral action into the model. The additional state vector component $x_0(t) \in R^p$ which represents the integration of the measured output is introduced as follows:

$$x_0(t) = \int_0^t \bar{y}(\tau)\, d\tau \tag{3.2}$$

It is worth noting that the measured output $\bar{y}(t)$ contains the actual output under disturbances signal, as well as possible fault occurring on the system signature.

Indeed, substituting $\bar{y}(t)$ from (3.1) in (3.2) gives:

$$\dot{x}_0(t) = C_0\bar{x}(t) + G_{d0}d(t) + G_{f0}f(t) \tag{3.3}$$

The additional variable $x_0(t)$ that represents the integral of noisy, and possibly faulty, output is combined with the state variable $\bar{x}(t)$ to form an augmented system. The augmented state vector $x(t) \in R^{(n+p)}$ is defined as:

$$x(t) = \begin{pmatrix} x_0(t) \\ \bar{x}(t) \end{pmatrix} \tag{3.4}$$

An additional output, denoted by $y_0(t)$, is also introduced into the augmented system. The output $y_0(t) \in R^p$ is defined as:

$$y_0(t) = x_0(t) \tag{3.5}$$

It can be seen that $y_0(t)$ is associated to $x_0(t)$ and it works like a 'virtual' output to reflect the integral of the output.

Now the augmented system output $y(t) \in R_{2p}$ is:

$$y(t) = \begin{pmatrix} y_0(t) \\ \bar{y}(t) \end{pmatrix} \tag{3.6}$$

Finally, the augmented system can be written as follows:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + E_d d(t) + E_f f(t) \\ y(t) = Cx(t) + G_d d(t) + G_f f(t) \end{cases} \tag{3.7}$$

where matrices A, B, C, E_d, E_f, G_d and G_f can be derived straightforward from the matrices in the original system state representation model (3.1), the additional state (3.3) and output Equations (3.4) & (3.6):

$$\begin{gathered} A = \begin{bmatrix} 0_{p\times p} & C_0 \\ 0_{n\times p} & A_0 \end{bmatrix} \; B = \begin{bmatrix} 0_{p\times m} \\ B_0 \end{bmatrix} \; C = \begin{bmatrix} I_p & 0_{p\times n} \\ 0_{p\times p} & C_0 \end{bmatrix} \\ E_d = \begin{bmatrix} G_{d0} \\ E_{d0} \end{bmatrix} \; E_f = \begin{bmatrix} G_{f0} \\ E_{f0} \end{bmatrix} \; G_d = \begin{bmatrix} 0_{p\times j} \\ G_{d0} \end{bmatrix} \; G_f = \begin{bmatrix} 0_{p\times j} \\ G_{f0} \end{bmatrix} \end{gathered} \tag{3.8}$$

For the following, the pair (A,C) in the augmented system (3.7) is assumed to be observable (see chapter 2 for observabiliy conditions).

3.2.3 Fault detector observer design

In this section, an observer of the augmented system is designed. Since the augmented system contains the integral terms of the noisy output measurements, a classic proportional observer can be used in the proposed design, but still maintains the main feature of PI observer. Without loss of generality, the Luenberger observer is chosen for sake of optimisation feasibility, which is beneficial in reducing the complexity of fault detector and tuning computation time. The proposed fault detection observer is described by:

$$\begin{cases} \dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - \hat{y}(t)) \\ \hat{y}(t) = C\hat{x}(t) \end{cases} \tag{3.9}$$

where $\dot{\hat{x}}(t) \in R^{(n+p)}$ is the estimated state vector and $\hat{y}(t) \in R^{2p}$ is the estimated output. The gain matrix K associated to the difference between the actual and estimated output $\begin{pmatrix} y(t) - \hat{y}(t) \\ y_0(t) - \hat{y}_0(t) \end{pmatrix}$, represents the correction term. The correction gain is selected in a way that observer estimate $\hat{x}(t)$ converges to the actual system values $x(t)$ despite the differences in initial state and the disturbances occurring in the system.

For the purpose of fault detection, a residual, which is the difference between the measured and estimated output, is used to detect the occurrence of faults. The residual is defined as:

$$r(t) = y(t) - \hat{y}(t) \tag{3.10}$$

Then, the estimation error $e(t) \in R^{(n+p)}$is defined by:

$$e(t) = x(t) - \hat{x}(t) \tag{3.11}$$

Thus the dynamic of the estimation error for system (3.7) is governed by:

$$\dot{e}(t) = (A - KC)e(t) + (E_d - KG_d)d(t) + (E_f - KG_f)f(t) \tag{3.12}$$

It can be seen that the residual $r(t)$ is linked to the estimation errors as below:

$$r(t) = y(t) - \hat{y}(t) = Ce(t) + G_d d(t) + G_f f(t) \tag{3.13}$$

It can be seen in Equation (3.12) that the gain matrix K is associated to the error, disturbance and fault signals. So when calculating its value, one should make sure that the matrix $(A - KC)$ is Hurwitz so the error is convergent. On the other hand, the fault effect can be amplified in the residual through the value of the matrix $(E_f - KG_f)$ when the disturbance effect can be reduced by minimising the value of $(E_d d - KG_d)$.

3.3 Observer gain calculation for disturbance attenuation

As shown in previous section, given a dynamic system defined in (3.1) with state vector $x(t) \in R^n$ and the output vector $y(t) \in R^p$, a PI observer with an augmented state vector $x(t) \in R^{(n+p)}$ and augmented output vector $y(t) \in R^{2p}$ can be designed. In the proposed PI observer (3.9), the gain matrix K is a $(n+p)$ by $2p$ matrix, i.e. $K \in R^{(n+p)\times 2p}$.

Furthermore, in order to facilitate the discussion on disturbance attenuation and fault detection, $K = \begin{bmatrix} K_I & K_p \end{bmatrix}$ can be expressed in a block-matrix form with four blocks as follows:

$$K_I = \begin{bmatrix} K_{11} \\ K_{21} \end{bmatrix} \; K_p = \begin{bmatrix} K_{12} \\ K_{22} \end{bmatrix} \tag{3.14}$$

where $K_{11} \in R$, $K_{12} \in R^{(p \times p)}$, $K_{21} \in R^{(n \times p)}$ and $K_{22} \in R_{(n \times p)}$

$$\begin{array}{c} K_{11} = \begin{bmatrix} k_{1,1} & \dots & k_{1,p} \\ \vdots & \ddots & \vdots \\ k_{p,1} & \dots & k_{p,p} \end{bmatrix} \; K_{12} = \begin{bmatrix} k_{1,p+1} & \dots & k_{1,2p} \\ \vdots & \ddots & \vdots \\ k_{p,p+1} & \dots & k_{p,2p} \end{bmatrix} \\ K_{21} = \begin{bmatrix} k_{p+1,1} & \dots & k_{p+1,p} \\ \vdots & \ddots & \vdots \\ k_{p+n,1} & \dots & k_{p+n,p} \end{bmatrix} \; K_{12} = \begin{bmatrix} k_{p+1,p+1} & \dots & k_{p+1,2p} \\ \vdots & \ddots & \vdots \\ k_{p+n,p+1} & \dots & k_{p+n,2p} \end{bmatrix} \end{array} \tag{3.15}$$

Submitting the block-matrix (3.14) into (3.12), the dynamics of the estimation errors can be rewritten as:

$$\begin{aligned} \dot{e}(t) \;&=\; \begin{pmatrix} -K_{11} & (I_{p-K_{12}})\,C_0 \\ -K_{21} & A_0 - K_{22}C_0 \end{pmatrix} e(t) \\ &+ \begin{pmatrix} (I_{p-K_{12}})\,G_{d0} \\ E_{d0} - K_{22}G_{d0} \end{pmatrix} d\,(t) + \begin{pmatrix} (I_{p-K_{12}})\,G_{f0} \\ E_{f0} - K_{22}G_{f0} \end{pmatrix} f\,(t) \end{aligned} \tag{3.16}$$

It can be seen that the disturbance $d(t)$ has an influence on the state estimation error through the gain matrix $\begin{pmatrix} (I_{p-K_{12}})\,G_{d0} \\ E_{d0} - K_{22}G_{df0} \end{pmatrix}$. So setting this matrix to zero would allow decoupling the state estimation from the disturbance or in another words, the disturbance will be completely attenuated.

While setting $(E_{d0} - K_{22}G_{d0})$ to zero can be done by finding an appropriate value of K_{22} such that:

$$(E_{d0} - K_{22}G_{d0}) = 0 \tag{3.17}$$

where $E_{d0} \in R^{(n*j)}$, is the input disturbance matrix and $G_{d0} \in R^{(p*j)}$ the output disturbance matrix, respectively.

Setting $(I_p - K_{12})$ to zero may result on losing the observability of the system since the same expression can be found in the matrix $\begin{pmatrix} -K_{11} & (I_{p-K_{12}})\,C_0 \\ -K_{21} & A_0 - K_{22}C_0 \end{pmatrix}$ which governs the error dynamic. So one has to be careful when calculating the value of K_{12} and need to set it such that $(I_p - K_{12})$ is minimised to reduce its effect on the error dynamic and while keeping the observability property of the system.

Let H denote the solution to Equation (3.17), that is:

$$E_{d0} = HG_{d0} \tag{3.18}$$

So, setting K_{22} to H results in the disturbance $d(t)$ being partially attenuated in the state estimation vector $\hat{x}(t)$. Therefore, K is:

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & H \end{bmatrix} \tag{3.19}$$

and the error dynamics in (3.16) is now given by:

$$\begin{aligned} \dot{e}(t) &= \begin{pmatrix} -K_{11} & (I_{p-K_{12}})\, C_0 \\ -K_{21} & A_0 - HC_0 \end{pmatrix} e(t) \\ &+ \begin{pmatrix} (I_{p-K_{12}})\, G_{d0} \\ E_{d0} - HG_{d0} \end{pmatrix} d\,(t) + \begin{pmatrix} (I_{p-K_{12}})\, G_{f0} \\ E_{f0} - HG_{f0} \end{pmatrix} f\,(t) \end{aligned} \tag{3.20}$$

As the $n+p$-dimensional state vector $x(t)$ consists of two parts, the original state variable $x(t)$ and the integral output variable $x_0(t)$, respectively, setting $(E_{d0} - K_{22}G_{d0})$ to zero, means attenuating the disturbance direct transfer in the original state $x(t)$. On the other hand the indirect transfer in the original state through the additional one, is reduced in the next section by minimising the value of $(I_{p-K_{12}})$.

3.4 Observer gain optimisation for fault detection

Since the objective of this work is to reduce the disturbance effect for fault detection, and the disturbance effect being already considerably reduced using the PI observer, this section focus is on finding appropriate values for the rest parts $\{K_{11}, K_{12}, K_{21}\}$ of the observer gain matrix so that the fault detection performance of the proposed PI observer can be optimised.

3.4.1 Residual analysis and eigen-decomposition

The approach used here is similar to the method developed in [57] based on [58], [59] and [60]. It relies on using the Laplace-transform of the residual to set an optimisation criterion. So, the Laplace-transform of (3.13) is calculated:

$$r(s) = M_d(s)d(s) + M_f(s)f(s) \tag{3.21}$$

where:

$$M_d(s) = C(sI - A + KC)^{-1}(E_d - KG_d) + G_d \tag{3.22}$$

and:

$$M_f(s) = C(sI - A + KC)^{-1}(E_f - KG_f) + G_f \tag{3.23}$$

As it has been shown in ([58]), ([59]) and ([60]), the observer gain matrix can be expressed as:

$$K = L^{-1}Q \tag{3.24}$$

where $L \in R^{(n+p)*(n+p)}$:

$$L = \begin{bmatrix} l_1^T \\ \vdots \\ l_{n+p}^T \end{bmatrix} = \begin{bmatrix} q_1^T C\left(A - \lambda_1 I\right)^{-1} \\ \vdots \\ q_{n+p}^T C\left(A - \lambda_{n+p} I\right)^{-1} \end{bmatrix} \tag{3.25}$$

and $Q \in R^{(n+p)*2p}$ is the matrix of free parameters:

$$Q^T = [q_1 \; q_2 \; \cdots \; q_{n+p}] \tag{3.26}$$

Then, based on this method, it has been shown in [57] that:

$$(sI - A + KC)^{-1} = R\Psi(s)L \tag{3.27}$$

where:

$$R = L^{-1} = [r_1 \; r_2 \; \cdots \; r_{n+p}] \tag{3.28}$$

and $\Psi(s) \in R^{(p+n)*(p+n)}$ is a diagonal matrix:

$$\Psi(s) = \begin{bmatrix} \frac{1}{s-\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{s-\lambda_{n+p}} \end{bmatrix} \tag{3.29}$$

Thus $(sI - A + KC)^{-1}$ can be written as:

$$(sI - A + KC)^{-1} = \frac{r_1 l_1^T}{s - \lambda_1} + \frac{r_2 l_2^T}{s - \lambda_2} + \ldots + \frac{r_{n+p} l_{n+p}^T}{s - \lambda_{n+p}} \tag{3.30}$$

Using this method to calculate the gain is very interesting as not only the values of the vector of free parameters Q can be optimised but it also can be done for the eigenvalues.

3.4.2 Gain matrix optimisation

As shown in (3.20), residual $r(t)$ is affected by both the fault and the disturbance. Therefore the gain matrix optimisation for fault detection is to find an appropriate value of K such that the residual is sensitive to the faults, but robust (non-sensitive) to the disturbance. In another words, the disturbance should be attenuated but the fault is signified in the residual [57]. Intuitively, the criterion for optimisation is two-fold to reduce the disturbance and signify the fault in the residual.

As it has been shown in [57], the criterion is set based on the signals frequency range analysis.

$$J = \frac{\|CR\Psi(s)L(E_d - L^{-1}QG_d) + G_f\|_{(s=jw_f)}}{\|CR\Psi(s)L(E_d - L^{-1}QG_d) + G_f\|_{(s=jwd)}} \tag{3.31}$$

where $|w_f| \in [0, \pi]$ and $|w_d| \in [0, \pi]$.are respectively the fault and disturbance frequencies to be considered

It is worth noting that since K_{22} has already been set so that the disturbance is attenuated, the denominator part of the criterion is added just to ensure that $(I_{p-K_{12}})$ is not set to a big value and thus finding the right value of w_d is less critical that for w_f. So maximising J will allow to considerably amplifying the fault and having good fault detection sensitivity

Using this criterion, various optimisation algorithms can be used to find the optimal gain value which allows having a stable observer i.e. the matrix $(A - KC)$ is Hurwitz, and optimising the fault detection in the residual.

3.5 Simulation results

In this section, simulation of fault detection in a DC motor is presented to demonstrate the performance of the proposed PI observer design method.

The state space model of the DC motor is as follows:

$$\begin{cases} \begin{pmatrix} \dot{x}_{0,1}(t) \\ \dot{x}_{0,2}(t) \end{pmatrix} = \begin{pmatrix} \frac{-b}{J} & \frac{K_t}{L} \\ \frac{-K_e}{L} & \frac{-R}{L} \end{pmatrix} \begin{pmatrix} x_{0,1}(t) \\ x_{0,2}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{L} \end{pmatrix} u(t) \\ y_0(t) = \begin{pmatrix} 0 & 1 \end{pmatrix} \end{cases} \tag{3.32}$$

where

J_w	Moment of inertia of the motor
b	Damping ratio
K_t	Torque constant
K_e	Voltage constant
R	Electric resistance
L	Electric inductance

The input of the system u is the voltage source and the output $y_0 = x_{0,2}$ is the rotational speed of the motor shaft. $x_{0,2}$ is the armature current.

The considered faults and disturbances are system and measurement faults. As the purpose of this work is to be able to signify the fault effect in the residual while reducing the disturbances effect, the considered disturbances are affecting the system's dynamic and the measured output. The fault to noise ratio is equal to 1.

The simulated fault is an incipient fault that develops gradually in the system until it reaches a certain final value. Indeed, it can be seen that, before $t = 40s$ the system is fault free and the fault develops during $t = 40s$ to $60s$ and maintains its value at $60s$ and onward.

The fault is simulated by the function below:

$$f(t) \begin{cases} 0 \ (t < 40 \\ 0.01 \ t(40 \leq t < 60) \\ 0.2 \ (t \geq 60)) \end{cases} \tag{3.33}$$

The disturbance signal $d(t)$ is set to be a Gaussian noise with mean value $\mu = 0$ variance $\sigma^2 = 0.2$.

In the objective function (3.31), the angular frequency w_f is set to 0 to reflect the fact that main components of the fault signal $f(t)$ is constant offset that affects the system.

Solving Equation (3.17) and using the optimisation algorithm based on the developed eigenstructure optimisation method to maximise the value of J gives the following optimised gain matrix K:

$$K = \begin{bmatrix} 7.84 & 0.23 \\ -11.56 & 1 \\ -95.58 & -2 \end{bmatrix} \tag{3.34}$$

For the purpose of comparison, a traditional proportional observer (denoted by P) is designed by using the place command in Matlab place $(A_0', C_0', [-4, -5])$. A traditional PI observer (denoted by PI) is also designed using the method used in [54]. The fault detection scheme simulation results are depicted in Fig. 3.1.

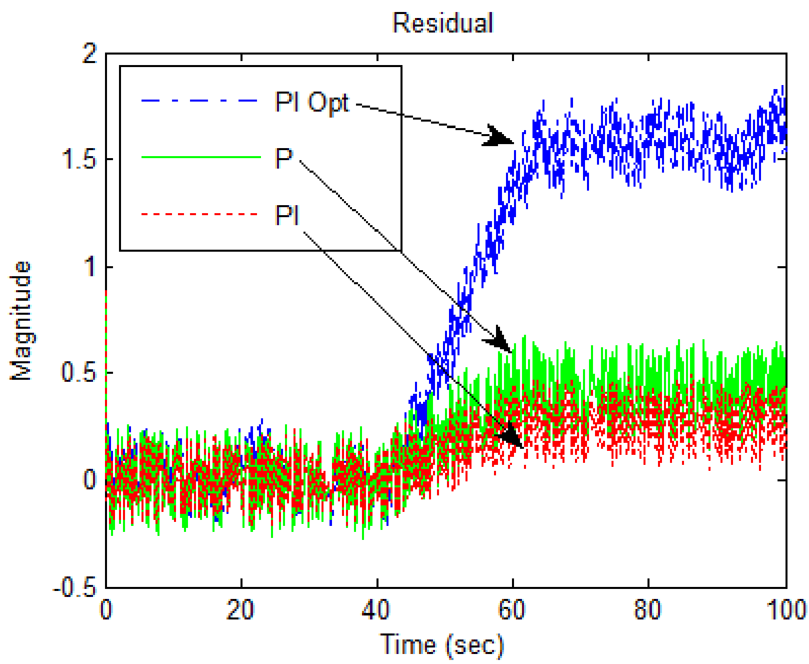

Figure 3.1: Residuals of Proportional (P), Proportional Integral (PI) and the Optimized Proportional Integral (PI-Opt) observers.

Fig. 3.1 shows that regarding disturbances attenuation, best results are given by the traditional PI observer, when both standard Proportional and optimised PI give acceptable results. This result was expected considering the good impact on reducing disturbances that PI observer has showed [54]. Besides, since optimised PI observer has been designed toward fault detection purpose it was also expected that its disturbances attenuation performances have been reduced but are still sufficient for the purpose of the FD scheme. Indeed, the deviation from zero of the residual in absence of fault is small enough not to cause flagging false alarms. Finally, the traditional P was simulated here for the purpose of comparison, validates its limited performances in the considered case scenario.

Now for the main purpose of scheme that is fault amplification impact in the residual, it is clear that the optimised PI observer is the only scheme that gives results that can be used for fault detection. Indeed, as shown in Fig. 3.1, the optimised PI residual value deviates considerably when the fault appears at time $t = 40s$. Furthermore, considering the gap between the signal in absence and in presence of fault, a threshold value can straightforward be found. For the traditional P and PI observers, simulation results show that it has insufficient results to be used for FD as the residual value isn't enough distinguishable in absence and presence of fault and my raise false alarms/miss actual alarms if used for FD.

3.6 Conclusion

In this chapter, a proportional integral (PI) observer is proposed for model based fault detection in systems. For the purpose of the design method, the considered MIMO model was first augmented using the integral of the output . Then, an approach based on eigenstructure optimisation to determine the PI observer gain to attenuate disturbances impact while signifying the fault effect in the residual was proposed. In particular, the frequency signatures of faults are taken into account to simplify an objective function and reduce the computation cost for better performances optimisation. Simulation results on a DC motor have verified the fault detection capabilities of the proposed scheme .

Chapter 4

Output observer design and application to nonlinear systems

4.1 Introduction

The concept of output observer design was first proposed in [61] for linear systems. It was shown that the output-observer was most appropriately designed by using an input-output model of the system rather than a state space representation of the latter. The technique consists of modelling the system in an input/output representation by using its input and output signals, as well as its parameters to build the link between the different parts of the differential equation that describes the system. So given that, only the input and output signals are used in the observer while in traditional observers, all the estimated states are required. As using less information may affect the observer performances, in this chapter it is shown that using respective integrals of the input and the output signals will allow to overcome the missing information impact and guarantees the observer estimation error convergence to zero.

Furthermore, as the majority of physical systems are nonlinear in nature, and their analytical solution being usually complex to find, in this chapter, instead of trying to find a solution for those systems, it will be taken advantage of using the output observer proprieties to propose a systematic way to tune the observer and determine its performances and stability proprieties. Indeed, the key feature of the output observer being the output injection, it is used to deign an observer for a class of nonlinear systems in which the estimation error dynamic will be independent from the nonlinearities. This allows the stability analysis to be straight forward and the observer performance to be more effectively enhanced.

This chapter is organised as follows. First the input/output representation development method is detailed. Then, the concept of output observer is introduced and its design procedure technique is given. Finally, based on its output injection feature, output observer technique is applied to nonlinear systems.

4.2 Output observer design for linear systems

Fig.4.1. shows, the structure of an output observer for linear systems. It can be seen that in the proposed observer, there is no need to use additional states to model the system. Instead, the key element of the observer here is the use of the input and output signals, and their respective integrals to duplicate the system. The tuning of the observer gain will then be the key to attenuate the estimation error and thus enhance its estimation

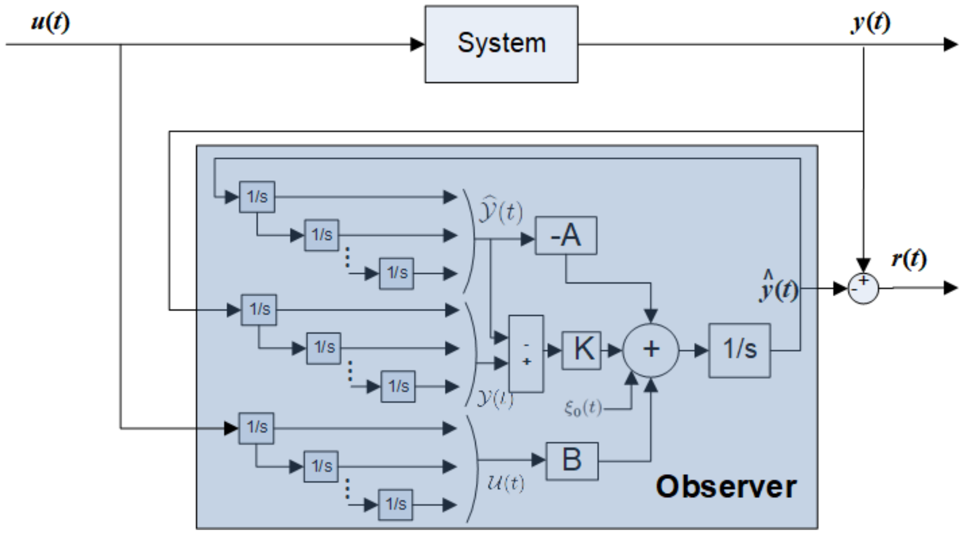

Figure 4.1: Output observer design for linear systems

performances.

4.2.1 System model - Linear Time Invariant (LTI) systems

For the following, consider the single input single output continuous linear time invariant systems that are described by the following input-output relationship

$$
\begin{aligned}
& y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) \\
= \; & b_m u^{(m)}(t) + \cdots + b_1\dot{u}(t) + b_0 u(t) \qquad (4.1)
\end{aligned}
$$

where $y(t)$ is the system's output and $u(t)$ the system's input and the system its transfer function is assumed to be proper, i.e. $n \geq m$. It is also assumed that $y(0)$ and $u(0)$ are known/measured.

The above system can be written in compact form as

$$y^{(n)}(t) = -A^T Y(t) + B^T U(t) \tag{4.2}$$

where

$$\begin{aligned} A &= \begin{pmatrix} a_{n-1} & \dots & a_1 & a_0 \end{pmatrix}^T \\ B &= \begin{pmatrix} b_m & \dots & b_1 & b_0 \end{pmatrix}^T \\ Y(t) &= \begin{pmatrix} y^{(n-1)}(t) & \dots & \dot{y}(t) & y(t) \end{pmatrix}^T \\ U(t) &= \begin{pmatrix} u^{(m)}(t) & \dots & \dot{u}(t) & u(t) \end{pmatrix}^T \end{aligned} \tag{4.3}$$

Notations: For $k \geq 0$, one shall denote by $\mathcal{I}_k\{f(t)\}$ the k integrations of the function $f(t)$ with respect to time; that is $\mathcal{I}_k\{f(t)\} = \underbrace{\int_0^t \cdots \int_0^t}_{k \text{ times}} f(\tau)d\tau ... d\tau$ and in particular, $\mathcal{I}_0\{f(t)\} = f(t)$. Also, $I_{-k}\{f(t)\} = \dfrac{d^k}{dt^k} f(t)$ and in particular $I_{-1}\{f(t)\} = \dfrac{df(t)}{dt}$.

By integrating system (4.2) $n-1$ times with respect to time, the following is obtained

$$\mathcal{I}_{n-1}\left\{y^{(n)}(t)\right\} = -A^T \mathcal{I}_{n-1}\left\{Y(t)\right\} + B^T \mathcal{I}_{n-1}\left\{U(t)\right\} \tag{4.4}$$

This yields

$$\dot{y}(t) = -A^T \mathcal{Y}(t) + B^T \mathcal{U}(t) + \xi_0(t) + \eta_\theta(t) \tag{4.5}$$

where

$$\mathcal{Y}(t) = \mathcal{I}_{n-1}\{Y(t)\} = \begin{pmatrix} y(t) & \ldots & \mathcal{I}_{n-2}\{y(t)\} & \mathcal{I}_{n-1}\{y(t)\} \end{pmatrix}^T$$
$$\mathcal{U}(t) = \mathcal{I}_{n-1}\{U(t)\} = \begin{pmatrix} \mathcal{I}_{n-1-m}\{u(t)\} & \ldots & \mathcal{I}_{n-2}\{u(t)\} & \mathcal{I}_{n-1}\{u(t)\} \end{pmatrix}^T \tag{4.6}$$

The function $\xi_0(t)$ is a known (polynomial) function of the known initial conditions $\{y(0), u(0)\}$, of the system, while $\eta_\theta(t)$ is an unknown (polynomial) function that is dependent on the unknown initial conditions of the system with $\theta = \{\dot{y}(0), \ddot{y}(0)..., \dot{u}(0), \ddot{u}(0)...\}$.

4.2.2 Development procedure of output/input representation

In this section, the detailed procedure of developing system (4.4) is given, and in particular, initial conditions polynomials $\xi_0(t)$ and $\eta_\theta(t)$ are detailed.

First, for $p \leq k$, the p-th order integration $y^{(k)}(t)$ can be expressed as

$$\mathcal{I}_p\{y^{(k)}(t)\} = y^{(k-p)}(t) - \sum_{i=0}^{p-1} y^{(k-p+i)}(0)\frac{t^i}{i!} \tag{4.7}$$

where $y^{(i)}$ denotes the $i-th$ derivative initial condition.

For the sake of simply notation, a polynomial $R_k(t)$ is introduced to represent the polynomial of initial conditions. That is

$$R_k(t) = \sum_{i=0}^{k-1} y^{(i)}(0)\frac{t^i}{i!} \tag{4.8}$$

Now, in particular, when $p = k$, it is

$$\mathcal{I}_p\{y^{(k)}(t)\} = \mathcal{I}_k\{y^{(k)}(t)\} = y(t) - \sum_{i=0}^{k-1} y^{(i)}(0)\frac{t^i}{i!} \tag{4.9}$$

Hence, the $k - th$ order integration of $y^{(k)}(t)$ can be written as

$$\mathcal{I}_k\{y^{(k)}(t)\} = y(t) - R_k(t) \tag{4.10}$$

Similarly, the $(k-1) - th$ order integration of $y^{(k)}(t)$ can be written as

$$\mathcal{I}_k\{y^{(k-1)}(t)\} = \dot{y}(t) - R_{k-1}(t) \tag{4.11}$$

In the scenario of $p > k$

$$\begin{aligned} \mathcal{I}_p\left\{y^{(k)}(t)\right\} &= \mathcal{I}_{p-k}\left\{\mathcal{I}_k\left\{y^{(k)}(t)\right\}\right\} \\ &= \mathcal{I}_{p-k}\left\{y(t)\right\} - \mathcal{I}_{p-k}\left\{R_k(t)\right\} \end{aligned} \tag{4.12}$$

Now

$$\begin{aligned} \mathcal{I}_{n-1}\left\{Y(t)\right\} &= \begin{pmatrix} \mathcal{I}_{n-1}\left\{y^{(n-1)}(t)\right\} \\ \vdots \\ \mathcal{I}_{n-1}\left\{\dot{y}(t)\right\} \\ \mathcal{I}_{n-1}\left\{y(t)\right\} \end{pmatrix} = \begin{pmatrix} y(t) - R_{n-2}(t) \\ \vdots \\ \mathcal{I}_{n-2}\left\{y(t)\right\} - \mathcal{I}_{n-2}\left\{R_0(t)\right\} \\ \mathcal{I}_{n-1}\left\{y(t)\right\} \end{pmatrix} \\ &= \mathcal{Y}(t) - \Gamma(t) \end{aligned} \tag{4.13}$$

where

$$\mathcal{Y}(t) = \begin{pmatrix} Y(t) \\ \vdots \\ \mathcal{I}_{n-1}\left\{Y(t)\right\} \end{pmatrix} \; \Gamma(t) = \begin{pmatrix} R_{n-2}(t) \\ \mathcal{I}_1\left\{R_{n-3}(t)\right\} \\ \vdots \\ \mathcal{I}_{n-2}\left\{R_0(t)\right\} \\ 0 \end{pmatrix} \tag{4.14}$$

with $R_{-1}(t) = 0$.

Similarly,

$$\begin{aligned} \mathcal{I}_{n-1}\left\{U(t)\right\} &= \begin{pmatrix} \mathcal{I}_{n-1}\left\{u^{(m)}(t)\right\} \\ \vdots \\ \mathcal{I}_{n-1}\left\{\dot{u}(t)\right\} \\ \mathcal{I}_{n-1}\left\{u(t)\right\} \end{pmatrix} = \begin{pmatrix} u(t) - S_{m-1}(t) \\ \vdots \\ \mathcal{I}_{n-2}\left\{u(t)\right\} - \mathcal{I}_{n-2}\left\{S_0(t)\right\} \\ \mathcal{I}_{n-1}\left\{u(t)\right\} \end{pmatrix} \\ &= \mathcal{U} - \Upsilon \end{aligned} \tag{4.15}$$

where

$$S_k(t) = \sum_{i=0}^{k} u^{(i)}(0)\frac{t^i}{i!} \tag{4.16}$$

with

$$\Upsilon(t) = \begin{pmatrix} S_{m-1}(t) \\ \vdots \\ \mathcal{I}_{n-2}\left\{S_0(t)\right\} \\ 0 \end{pmatrix} \tag{4.17}$$

Then, using the previously developed calculations, one can obtain the following expression of the output dynamic:

$$\begin{aligned} \mathcal{I}_{n-1}\left\{y^{(n)}(t)\right\} &= -A^T\mathcal{I}_{n-1}\left\{Y(t)\right\} + B^T\mathcal{I}_{n-1}\left\{U(t)\right\} \\ \dot{y}(t) - \sum_{i=0}^{n-2} y^{(i+1)}(0)\frac{t^i}{i!} &= -A^T\left(\mathcal{Y} - \Gamma\right) + B^T\left(\mathcal{U} - \Upsilon\right) \end{aligned} \tag{4.18}$$

That is

$$\begin{aligned} \dot{y}(t) &= -A^T\left(\mathcal{Y} - \Gamma\right) + B^T\left(\mathcal{U} - \Upsilon\right) + \sum_{i=0}^{n-2} y^{(i+1)}(0)\frac{t^i}{i!} \\ &= -A^T\mathcal{Y} + B^T\mathcal{U} + \xi_\theta(t) \end{aligned} \tag{4.19}$$

where

$$\xi_0(t) = A^T\Gamma - B^T\Upsilon + \sum_{i=0}^{n-2} y^{(i+1)}(0)\frac{t^i}{i!} \tag{4.20}$$

is a parametrised function with the initial conditions of the input and the output, i.e. $\theta = (y(0), \dot{y}(0), ..., u(0), \dot{u}(0), ...)$.

Note that

$$R_k(t) = \sum_{i=0}^{k} y^{(i)}(0)\frac{t^i}{i!} = y(0) + \bar{R}_k(t) \tag{4.21}$$

and

$$S_k(t) = u(0) + \sum_{i=1}^{k} u^{(i)}(0)\frac{t^i}{i!} = u(0) + \bar{S}_k(t) \tag{4.22}$$

Therefore,

$$\begin{aligned}
\xi_\theta(t) &= A^T\Gamma - B^T\Upsilon + \sum_{i=0}^{n-2} y^{(i+1)}(0)\frac{t^i}{i!} \\
&= A^T\bar{\Gamma} - B^T\bar{\Upsilon} + \eta_\theta(t) \\
&= \xi_0(t) + \eta_\theta(t) \qquad (4.23)
\end{aligned}$$

where

$$\xi_0(t) = A^T \begin{pmatrix} y(0) \\ \vdots \\ \mathcal{I}_{n-2}\{y(0)\} \\ 0 \end{pmatrix} - B^T \begin{pmatrix} u(0) \\ \vdots \\ \mathcal{I}_{n-2}\{u(0)\} \\ 0 \end{pmatrix} \qquad (4.24)$$

$$\eta_\theta(t) = A^T\bar{\Gamma} - B^T\bar{\Upsilon} + \sum_{i=0}^{n-2} y^{(i+1)}(0)\frac{t^i}{i!} \qquad (4.25)$$

and

$$\bar{\Gamma} = \begin{pmatrix} \bar{R}_{n-2}(t) \\ \vdots \\ \mathcal{I}_{n-3}\left\{\bar{R}_1(t)\right\} \\ 0 \\ 0 \end{pmatrix} \quad \bar{\Upsilon} = \begin{pmatrix} \bar{S}_{m-1}(t) \\ \vdots \\ \mathcal{I}_{n-3}\left\{\bar{S}_1(t)\right\} \\ 0 \\ 0 \end{pmatrix} \qquad (4.26)$$

Thus $\xi_0(t)$ and $\eta(\theta, t)$, are of a polynomial form which exact expression is known. This will allow to find the final value of the estimation error of the observer.

The interest of this development is that it shows the procedure to follow in order to obtain the output of the signal dynamics as function of its $n-1$ integrals, the input and its $p-1$ integrals. Thus the system can be studied using only $y\,(t)$ and $u\,(t)$.

4.2.3 Output observer design for LTI systems

The following result is obtained.

Theorem 1

The observer described by the following expression:

$$\dot{\hat{y}}(t) = -A^T \widehat{\mathcal{Y}}(t) + B^T \mathcal{U}(t) + K^T \left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t) \right) + \xi_0(t) \tag{4.27}$$

is an asymptotic observer for the system described by Equation (4.5) provided that the gain $K^T = \begin{pmatrix} k_{n-1} & \dots & k_1 & k_0 \end{pmatrix}$ *is chosen such that the polynomial* $D(s) = s^n + \sum_{i=0}^{n-1} (a_i + k_i)\, s^i$ *is stable; i.e. the roots of* $D(s) = 0$ *lie in the left-half complex plane, with* $a_0 + k_0 \neq 0$.

where

$$\widehat{\mathcal{Y}}(t) = \begin{pmatrix} \hat{y}^{(n-1)}(t) & \dots & \dot{\hat{y}}(t) & \hat{y}(t) \end{pmatrix}^T \tag{4.28}$$

Proof

By defining $\varepsilon(t) = y(t) - \hat{y}(t)$, to represent the estimation errors of the proposed observer, the error dynamics of the observer can be written as

$$\dot{\varepsilon}(t) = -\left(A^T + K^T\right) \left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t) \right) + \eta_\theta(t) \tag{4.29}$$

Now, by taking the Laplace transform of Equation (4.29), one obtains

$$s\varepsilon(s) - \varepsilon(0) = -\left(A^T + K^T\right)\Delta(s)\varepsilon(s) + \eta_\theta(s) \tag{4.30}$$

where $\varepsilon(s)$ and $\eta_\theta(s)$ are the Laplace transform of the functions $\varepsilon(t)$ and $\eta_\theta(t)$ respectively and

$$\Delta^T(s) = \begin{pmatrix} 1 & \frac{1}{s} & \cdots & \frac{1}{s^{n-1}} \end{pmatrix} \tag{4.31}$$

Consequently,

$$\varepsilon(s) = \frac{\eta_\theta(s) + \varepsilon(0)}{s + \left(A^T + K^T\right)\Delta(s)} \tag{4.32}$$

Note that

$$\left(A^T + K^T\right)\Delta(s) = \sum_{i=0}^{n-1} \left(a_i + k_i\right)\frac{1}{s^{n-1-i}} \tag{4.33}$$

From the expression of the function $\eta_\theta(t)$ given in (4.25), it can be noticed that it is a polynomial function of order $n-2$. Hence, $\eta_\theta(s)$ is also a polynomial in $1/s$ of order $n-1$.

As a result,

$$\varepsilon(s) = \frac{s^{n-1}\eta_\theta(s) + s^{n-1}\varepsilon(0)}{s^n + s^{n-1}\left(A^T + K^T\right)\Delta(s)} = \frac{N(s)}{D(s)} \tag{4.34}$$

Using the expression (4.33), the polynomial $D(s)$ is explicitly given by

$$\begin{aligned} D(s) &= s^n + s^{n-1}\left(A^T + K^T\right)\Delta(s) \\ &= s^n + \sum_{i=0}^{n-1}\left(a_i + k_i\right)s^i \end{aligned} \tag{4.35}$$

One must choose, K such that this polynomial is stable so final value theorem can apply [62].

Finally, applying the final value theorem to gives

$$\begin{aligned} \lim_{s\to 0} s\varepsilon(s) &= \frac{s^n\eta_\theta(s) + s^n\varepsilon(0)}{s^n + \sum_{i=0}^{n-1}\left(a_i + k_i\right)s^n} \\ &= \frac{s^n\eta_\theta(s) + s^n\varepsilon(0)}{s^n + \sum_{i=1}^{n-1}\left(a_i + k_i\right)s^i + (a_0 + k_0)} \\ &= \frac{0}{(a_0 + k_0)} \end{aligned} \tag{4.36}$$

As one can see, the final value of the residual is equal to zero if k_0 is chosen such that $(a_0 + k_0) \neq 0$. This proves that if the polynomial $\xi_0(t)$ is known, the proposed output observer is asymptotically stable and its estimation error converges asymptotically to zero.

4.3 Output observer with output injection

Consider the observer described in the previous section, one can notice, since the input/output representation of the system is being used, when duplicating the system model in the observer, the actual system output can be used instead of using the estimated output.

The benefit of such a technique is to use the actual system output in the observer when possible which will result in reducing the estimation error. This will inevitably lead to increase the observer performances (shorter response time, more stability margin and lesser estimation error value). So the key difference in the observer with output injection structure is that the measured output is not only used for error estimation correction, but also in the system duplication part. Thus, it is the estimated output that is only used for the estimation error correction.

Theorem 2

The following system:

$$\dot{y}(t) = -A^T \mathcal{Y}(t) + B^T \mathcal{U}(t) + K^T \left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t) \right) + \xi_0(t) \tag{4.37}$$

is an asymptotic observer for system (4.5) provided that the gain $K = \begin{pmatrix} k_{n-1} & \dots & k_1 & k_0 \end{pmatrix}^T$ *is chosen such that the polynomial* $D(s) = s^n + \sum_{i=0}^{n-1} k_i s^i$ *is stable; i.e. the roots of* $D(s) = 0$ *lie in the left-half complex plan, with* $k_0 \neq 0$.

Proof

The proof follows along the same lines as that of Theorem 1. Indeed, defining $\varepsilon(t)$ and $\varepsilon(s)$ as above, one can show that:

$$\dot{\varepsilon}(t) = -K^T \left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t) \right) + \eta_\theta(t) \tag{4.38}$$

Similar to previous section, the following is obtained

$$\varepsilon(s) = \frac{\eta_\theta(s) + \varepsilon(0)}{s + K^T\Delta(s)} \tag{4.39}$$

thus, applying the final value theorem to gives

$$\begin{aligned} \lim_{s\to 0} s\varepsilon(s) &= \frac{s^n\eta_\theta(s) + s^n\varepsilon(0)}{s^n + \sum_{i=0}^{n-1}(k_i)\, s^n} \\ &= \frac{s^n\eta_\theta(s) + s^n\varepsilon(0)}{s^n + \sum_{i=1}^{n-1}(k_i)\, s^i + (k_0)} \\ &= \frac{0}{k_0} \end{aligned} \tag{4.40}$$

The output injection technique will remove the dependency of the observer error convergence on the system parameters vector A provided that the roots of $s^n + \sum_{i=0}^{n-1}(a_i)\, s^i$ are stable. So choosing K such that the roots $s + K^T\Delta(s)$ are on the left half open plan and $k_0 \neq 0$ will ensure that the estimation error will converge asymptotically to zero.

4.4 Output observer for a class of nonlinear systems

To deal with nonlinear systems, output observer with output injection technique will allow to eliminate the nonlinearity in the error estimation dynamic. This means that for the considered class of nonlinear systems, using this technique will allow to build an observer which performances can be proved without having to deal with the nonlinearities.

4.4.1 System model - a class of nonlinear system

Consider the following class of nonlinear systems:

$$
\begin{aligned}
\dot{x}_1(t) &= x_2(t) + \varphi_1(y(t), u(t)) \\
\dot{x}_2(t) &= x_3(t) + \varphi_2(y(t), u(t)) \\
\dot{x}_3(t) &= x_4(t) + \varphi_3(y(t), u(t)) \\
& \dots \\
\dot{x}_{n-2}(t) &= x_{n-1}(t) + \varphi_{n-2}(y(t), u(t)) \\
\dot{x}_{n-1}(t) &= x_n(t) + \varphi_{n-1}(y(t), u(t)) \\
\dot{x}_n(t) &= \varphi_n(y(t), u(t))
\end{aligned}
\tag{4.41}
$$

$$
y(t) = x_1(t) \tag{4.42}
$$

where the nonlinear functions $\varphi_{1 \dots n}(t)$ are expressed only in term of input and output signals. This method can be applied to systems with strong nonlinearities where only states that bring nonlinearities need to be measured in order to build an observer for the nonlinear system.

In this case, for sake of simplicity, consider a single output case but knowing that the approach developed here also applies to multiple output systems.

Rearranging the system described by (4.41) gives

$$\begin{aligned}
x_2(t) &= \dot{x}_1(t) - \varphi_1(y(t), u(t)) \\
x_3(t) &= \ddot{x}_1(t) - \dot{\varphi}_1(y(t), u(t)) - \varphi_2(y(t), u(t)) \\
x_4(t) &= \dddot{x}_1(t) - \ddot{\varphi}_1(y(t), u(t)) - \dot{\varphi}_2(y(t), u(t)) - \varphi_3(y(t), u(t)) \\
&\dots \\
x_{n-1}(t) &= x_1^{(n-2)}(t) - \varphi_1^{(n-3)}(y(t), u(t) - \varphi_2^{(n-4)}(y(t), u(t)) - \dots - \varphi_{n-2}(y(t), u(t)) \\
x_n(t) &= x_1^{(n-1)}(t) - \varphi_1^{(n-2)}(y(t), u(t)) - \varphi_2^{(n-3)}(y(t), u(t)) - \dots - \varphi_{n-1}(y(t), u(t)) \\
x_1^{(n)}(t) &= \varphi_1^{(n-1)}(y(t), u(t)) + \varphi_2^{(n-2)}(y(t), u(t)) \\
&+ \dots + \dot{\varphi}_{n-1}(y(t), u(t)) + \varphi_n(y(t), u(t)) \qquad (4.43)
\end{aligned}$$

which leads by using the output Equation (4.42) to the input/output representation of the system

$$\begin{aligned}
y^{(n)}(t) &= \varphi_1^{(n-1)}(y(t), u(t)) + \varphi_2^{(n-2)}(y(t), u(t)) \\
&+ \dots + \dot{\varphi}_{n-1}(y(t), u(t)) + \varphi_n(y(t), u(t)) \qquad (4.44)
\end{aligned}$$

In order to have the output dynamic expression $\dot{y}(t)$, it is only needed to integrate Equation 4.44 $n-1$ times

$$\begin{aligned}
\dot{y}(t) &= \varphi_1(y(t), u(t)) + \mathcal{I}_1 \{\bar{\varphi}_2(y(t), u(t))\} + \ldots + \mathcal{I}_{n-2} \left\{\bar{\varphi}_{n-1}(y(t), u(t))\right\} \\
&+ \mathcal{I}_{n-1} \{\bar{\varphi}_n(y(t), u(t))\} + \xi_\theta(t) \qquad (4.45)
\end{aligned}$$

It can be seen here that the output dynamic now depends only on the input and output signals. This result is necessary as to design the output observer, no estimated states can be used in the model.

4.4.2 Output observer design for a class of nonlinear systems

Theorem 3

The following system:

$$\begin{aligned}\dot{\hat{y}}(t) \quad &= \quad \varphi_1(y(t),u(t)) + \mathcal{I}_1\left\{\bar{\varphi}_2(y(t),u(t))\right\} + \ldots + \\ &\mathcal{I}_{n-2}\left\{\bar{\varphi}_{n-1}(y(t),u(t))\right\} + \mathcal{I}_{n-1}\left\{\bar{\varphi}_n(y(t),u(t))\right\} \qquad (4.46)\\ &+K^T\left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t)\right) + \xi_0(t) \qquad (4.47)\end{aligned}$$

is an asymptotic observer for system (4.45) provided that the gain $K = \begin{pmatrix} k_{n-1} & \ldots & k_1 & k_0 \end{pmatrix}^T$ *with* $k_0 \neq 0$ *is chosen such that the polynomial* $D(s) = s^n + \sum_{i=0}^{n-1} k_i s^i$ *is stable; i.e. the roots of* $D(s) = 0$ *lie in the left-half complex plan.*

Proof

By defining $\varepsilon(t) = y(t) - \hat{y}(t)$, to represent the estimation errors of the proposed observer, the error dynamics of the observer can be written as

$$\dot{\varepsilon}(t) = -K^T\left(\mathcal{Y}(t) - \widehat{\mathcal{Y}}(t)\right) + \eta_\theta(t) \qquad (4.48)$$

It can be seen here that all the nonlinearities have been eliminated in the estimation error dynamic, which means that the observer estimation error convergence can be proved in a linear form. To do so, the final value theorem that is used with the Laplace transform of the error dynamic expression.

So, by taking the Laplace transform of Equation (4.48), one obtains

$$s\varepsilon(s) - \varepsilon(0) = -K^T\Delta(s)\varepsilon(s) + \eta_\theta(s) \tag{4.49}$$

where $\varepsilon(s)$ and $\eta(\theta, s)$ are the Laplace transform of the functions $\varepsilon(t)$ and $\eta_\theta(s)$ respectively and

$$\Delta^T(s) = \begin{pmatrix} 1 & \frac{1}{s} & \cdots & \frac{1}{s^{n-1}} \end{pmatrix} \tag{4.50}$$

Consequently,

$$\varepsilon(s) = \frac{\eta_\theta(s) + \varepsilon(0)}{s + K^T\Delta(s)} \tag{4.51}$$

Note that

$$K^T\Delta(s) = \sum_{i=0}^{n-1} k_i \frac{1}{s^{n-1-i}} \tag{4.52}$$

It can be seen here that the proof procedure provided for LTI systems applies here as the nonlinearities have been removed in the estimation error expression. Thus it can be established that for this class of nonlinear systems, the output observer with output injection is asymptotically stable.

4.5 Extension to general case of nonlinear systems

In this section, the case where the nonlinearity in the system will depend not only on the measured states of the systems, but on all the states is considered.

First, as proved in [63], provided that the system is stable, it can then be rewritten in a way that the nonlinearities are expressed in terms of the states in a triangular form as shown below

$$\begin{aligned}
\dot{x}_1(t) &= x_2(t) + \varphi_1(x_1, u(t)) \\
\dot{x}_2(t) &= x_3(t) + \varphi_2(x_1, x_2, u(t)) \\
\dot{x}_3(t) &= x_4(t) + \varphi_3(x_1, x_2, x_3, u(t)) \\
& ... \\
\dot{x}_{n-2}(t) &= x_{n-1}(t) + \varphi_{n-2}(x_1, x_2, ...x_{n-2}, u(t)) \\
\dot{x}_{n-1}(t) &= x_n(t) + \varphi_{n-1}(x_1, x_2, ...x_{n-1}, u(t)) \\
\dot{x}_n(t) &= \varphi_n(x(t), u(t))
\end{aligned} \tag{4.53}$$

$$y(t) = x_1(t) \tag{4.54}$$

Then system 4.53 can be developed as below

$$
\begin{aligned}
x_2(t) &= \dot{x}_1(t) - \varphi_1(x_1, u(t)) \\
x_3(t) &= \ddot{x}_1(t) - \dot{\varphi}_1(x_1, u(t)) - \bar{\varphi}_2(x_1, \dot{x}_1(t), u(t)) \\
x_4(t) &= \dddot{x}_1(t) - \ddot{\varphi}_1(x_1, u(t)) - \dot{\bar{\varphi}}_2(x_1, \dot{x}_1(t), u(t)) - \bar{\varphi}_3(x_1, \dot{x}_1(t), \ddot{x}_1(t), u(t), \dot{u}(t)) \\
&\ldots \\
x_{n-1}(t) &= x_1^{(n-2)}(t) - \varphi_1^{(n-3)}(x_1(t), u(t)) - \bar{\varphi}_2^{(n-4)}(x_1(t), \dot{x}_1(t), u(t)) \\
&-\bar{\varphi}_3^{(n-5)}(x_1(t), \dot{x}_1(t), \ddot{x}_1(t), u(t), \dot{u}(t)) \\
&- \ldots - \dot{\bar{\varphi}}_{n-3}^{(1)}(x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-4)}(t), u(t), \ldots, u(t)^{(n-5)}) \\
&-\bar{\varphi}_{n-2}(x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-3)}(t), u(t), \ldots, u(t)^{(n-4)}) \\
x_n(t) &= x_1^{(n-1)}(t) - \varphi_1^{(n-2)}(x_1(t), u(t)) - \bar{\varphi}_2^{(n-3)}(x_1(t), \dot{x}_1(t), u(t)) \\
&-\bar{\varphi}_3^{(n-4)}(x_1(t), \dot{x}_1(t), \ddot{x}_1(t), u(t), \dot{u}(t)) \\
&- \ldots - \dot{\bar{\varphi}}_{n-2}(x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-3)}(t), u(t), \ldots, u(t)^{(n-4)}) \\
&-\varphi_{n-1}((x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-2)}(t), u(t), \ldots, u(t)^{(n-3)}) \\
x_1^{(n)}(t) &= \varphi_1^{(n-1)}(x_1(t), u(t)) + \bar{\varphi}_2^{(n-2)}(x_1(t), \dot{x}_1(t), u(t)) \\
&+\bar{\varphi}_3^{(n-3)}(x_1(t), \dot{x}_1(t), \ddot{x}_1(t), u(t), \dot{u}(t)) \\
&+ \ldots + \dot{\bar{\varphi}}_{n-1}((x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-2)}(t), u(t), u(t), \ldots, u(t)^{(n-3)}) \\
&+\bar{\varphi}_n((x_1(t), \dot{x}_1(t), \ldots, x_1^{(n-1)}(t), u(t), u(t), \ldots, u(t)^{(n-2)}) \qquad (4.55)
\end{aligned}
$$

And using the same procedure than in previous subsections gives

$$\begin{aligned} y^{(n)}(t) &= \varphi_1^{(n-1)}(x_1(t),u(t)) + \bar{\varphi}_2^{(n-2)}(x_1(t),\dot{x}_1(t),u(t)) \\ &+\bar{\varphi}_3^{(n-3)}(x_1(t),\dot{x}_1(t),\ddot{x}_1(t),u(t),\dot{u}(t)) + \ldots + \\ &\dot{\bar{\varphi}}_{n-1}((x_1(t),\dot{x}_1(t),\ldots,x_1^{(n-2)}(t),u(t),u(t),\ldots,u(t)^{(n-3)}) \\ &+\bar{\varphi}_n((x_1(t),\dot{x}_1(t),\ldots,x_1^{(n-1)}(t),u(t),u(t),\ldots,u(t)^{(n-2)}) \end{aligned} \tag{4.56}$$

Finally, integrating the above equation $n-1$ times gives

$$\begin{aligned} \dot{y}(t) &= \varphi_1(y(t),u(t)) + \mathcal{I}_1\left\{\bar{\varphi}_2(y(t),\dot{y}(t),u(t))\right\} \\ &+\mathcal{I}_2\left\{\bar{\varphi}_3(y(t),\dot{y}(t),\ddot{y}(t),u(t),\dot{u}(t))\right\} \\ &+\ldots+\mathcal{I}_{n-2}\left\{\dot{\bar{\varphi}}_{n-1}((y(t),\dot{y}(t),\ldots,y^{(n-2)}(t),u(t),\ldots,u(t)^{(n-3)})\right\} \\ &+\mathcal{I}_{n-1}\left\{\bar{\varphi}_n((y(t),\dot{y}(t),\ldots,y^{(n-1)}(t),u(t),\ldots,u(t)^{(n-2)})\right\} + \eta_\theta(s) \end{aligned} \tag{4.57}$$

Now, as one can see, the nonlinearities contain derivatives of the input and the output which can not be considered in the observer. For this, since the degree of integration associated to every nonlinear functions $\bar{\varphi}_{1\ldots n}$ is equal or greater to the degree of derivation of the input and the output it depends on, are considered only the class of nonlinear functions where after solving those integration functions, the nonlinear functions will not depend on the derivatives of the input and the output anymore.

Provided that, an observer for the above system using the output injection technique is designed:

$$\begin{aligned}\dot{\hat{y}}(t) &= \varphi_1(y(t),u(t)) + \mathcal{I}_1\left\{\bar{\varphi}_2(\hat{y}(t),\dot{y}(t),u(t))\right\} \\ &+\mathcal{I}_2\left\{\bar{\varphi}_3(\hat{y}(t),\dot{y}(t),\ddot{y}(t),u(t),\dot{u}(t))\right\} + \ldots + \\ &\mathcal{I}_{n-2}\left\{\dot{\bar{\varphi}}_{n-1}((\hat{y}(t),\dot{y}(t),\ldots,y^{(n-2)}(t),u(t),\ldots,u(t)^{(n-3)})\right\} \\ &+\mathcal{I}_{n-1}\left\{\bar{\varphi}_n((\hat{y}(t),\dot{y}(t),\ldots,y^{(n-1)}(t),u(t),\ldots,u(t)^{(n-2)})\right\} \\ &+K_{n-1}\mathcal{I}_{n-1}\left\{e(\tau)d\tau\right\} + \eta_0(s) \qquad (4.58)\end{aligned}$$

So, the error dynamics is given by:

$$\dot{e}(t) = \bar{\eta}_\theta(t) - \ldots - K_{n-1}\mathcal{I}_{n-1}\left\{e(\tau)d\tau\right\} \qquad (4.59)$$

Then using the same procedure than in previous subsection, the observer convergence can be proved.

This shows that for general class of nonlinear functions which fall under the assumption considered here, the proposed output observer design calculation is feasible and its convergence proven without having to deal with the nonlinearities.

4.6 Conclusion

In this chapter, a design of the output observer was proposed and its stability conditions proved, first for a special class of nonlinear systems then for general class of nonlinear systems in which a constraint was considered. It was showed its benefits to deal with a special class of nonlinear systems and general nonlinear systems under a given assumption. Indeed, in this approach, using the output injection technique, it was possible to eliminate the nonlinearity from the residual then using the proof that was developed for output observer for LTI systems, it was showed that its stability can be proved. Consequently, the observer gain calculation can be done toward observer performances in a systematic given way.

Chapter 5

Robust output observer for fault detection

5.1 Introduction

Model based fault detection relies on using a mathematical model that describes the behavior of the system. The basic idea is to generate a residual in order to make a diagnosis of the system. The residual is the difference between the actual measurements of the monitored system and the values estimated using the mathematical model based observer. Under healthy conditions, the mathematical model matches the real system and the residual will be small enough (i.e. close to zero) to be ignorable. In the presence of fault, the residual of a well-designed fault detector should be apparently different from zero, thus a fault detection alarm can be raised. One can make the observation that as far as fault-detection is concerned, it is unnecessary to estimate all the states variables as long as an estimation of the output is obtained. Consequently, rather than a using a state observer, an output observer is sufficient.

Further, the performance criterion of fault detection system includes robustness. Indeed, in practice, even in healthy conditions, the residual deviates from zero and it is impossible to have a perfect zero residual, because the residual is not only determined by the faults, but also affected by measurement noise, disturbances, parametric variations, unmodelled dynamics and nonlinearities presented to the system. A non-zero residual will easily lead to false alarms. A lot of existing methods can reduce the impact of these effects on the residual and avoid false alarms. Unfortunately, neither of them can give a perfect decoupling between these effects and the residual. Furthermore, the residual has to be compared with a threshold for decision-making. In model based fault detection, one of the main challenges is the generation of a residual that is as small as possible under healthy conditions and the calculation of threshold value that avoid raising false alarms or missing faults occurring in the system.

In this chapter, the previously developed output observer technique is applied to fault detection in LTI systems. It will be first applied it to single-output systems, then the approach will be extended to multi-output systems. Also, a solution to robust fault-detection issue is proposed and a systematic method to calculate the observer gain based on an optimisation method is introduced.

5.2 Output observer for fault detection

In Chapter 4, it was proved that under the assumed constraints, using the output observer approach for nonlinear leads to dealing with a residual that nonlinearities free. Thus, the attention in this chapter is focused on linear systems. First, the output observer approach is applied to LTI systems subject to faults.

Consider the model of a faulty system, where fault dynamics parameters are known, and described as follows

$$\begin{aligned} y^{(n)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) &= \left[b_m u^{(m)}(t) + \cdots + b_1\dot{u}(t) + b_0 u(t)\right] \\ &+ \left[g_q f^{(q)}(t) + \cdots + g_1\dot{f}(t) + g_0 f(t)\right] \end{aligned} \tag{5.1}$$

where $f(t)$ represents the signal of the fault acting on the system through the fault input parameters $\begin{pmatrix} g_q & \dots & g_1 & g_0 \end{pmatrix}$. Similar to the compact form of the fault-free system model, the system subject to fault $f(t)$ can be rewritten as

$$y^{(n)}(t) = -A^T Y(t) + B^T U(t) + G^T F(t) \tag{5.2}$$

Denoting Y^T and U^T as in (4.6). $G^T = \begin{pmatrix} g_q & \dots & g_1 & g_0 \end{pmatrix}$ is a q-th order input vector and $F^T = \begin{pmatrix} f^{(q)}(t) & \dots & \dot{f}(t) & f(t) \end{pmatrix}$is the vector of fault derivatives. For sake of simplicity, it is assumed that the system is free of fault at $t = 0$. That is $f(0) = \dot{f}(0) = \ldots = f^{(q)}(0) = 0$.

As previously done, by integrating system (5.2), $n-1$ times with respect to time, one obtains

$$\dot{y}(t) = -A^T \mathcal{Y} + B^T \mathcal{U} + G^T \mathcal{F} + \xi_0(t) + \eta_\theta(t) \tag{5.3}$$

where

$$\mathcal{F}^T = \mathcal{I}_{n-1}\{F(t)\} = \begin{pmatrix} \mathcal{I}_{n-q-1}\{f(t)\} & \dots & \mathcal{I}_{n-2}\{f(t)\} & \mathcal{I}_{n-1}\{f(t)\} \end{pmatrix} \tag{5.4}$$

The output observer design for fault detection relies on using the same approach used for the free system. Only, in the case of faulty systems, in addition to estimation error convergence despite initial conditions deviations, we will prove that under certain conditions, the fault can be detected with minimal conditions on the gain value.

Theorem 4

Consider the system described by (5.3) and its following observer

$$\dot{\hat{y}}(t) = -A^T \widehat{\mathcal{Y}} + B^T U + K^T \left(\mathcal{Y} - \widehat{\mathcal{Y}}\right) + \xi_0(t) \tag{5.5}$$

where $K^T = \begin{pmatrix} k_{n-1} & \dots & k_1 & k_0 \end{pmatrix}$ *with* $a_0 + k_0 \neq 0$ *is chosen such that the polynomial* $D(s) = s^n + \sum_{i=0}^{n-1} (a_i + k_i)\, s^i$ *is stable. Let the fault* $f(t)$ *be a polynomial of order at most equal to* $q - 1$.

Then,

$$\lim_{t\to\infty} (y(t) - \hat{y}(t)) \neq 0 \tag{5.6}$$

where f *is the identified parameter index and* c_f *is a constant.*

Proof

Setting $\varepsilon(t) = y(t) - \hat{y}(t)$, the error dynamics can be written as

$$\dot{\varepsilon}(t) = -\left(A^T + K^T\right)\left(\mathcal{Y} - \widehat{\mathcal{Y}}\right) + G^T\mathcal{F} + \eta_\theta(t)) \tag{5.7}$$

Then, by taking the Laplace transform of Equation (5.7)

$$s\varepsilon(s) - \varepsilon(0) = -\left(A^T + K^T\right)\Delta(s)\varepsilon(s) + G^T\Lambda(s)f(s) + \eta_\theta(s) \tag{5.8}$$

where $f(s)$ is the Laplace transform of $f(t)$.

$\Delta(s)$ is defined as in (4.31) and

$$\Lambda(s)^T = \begin{pmatrix} \frac{1}{s^{n-q-1}} & \dots & \frac{1}{s^{n-2}} & \frac{1}{s^{n-1}} \end{pmatrix} \tag{5.9}$$

Consequently,

$$\begin{aligned} \varepsilon(s) &= \frac{G^T\Lambda(s)f(s)}{s + (A^T + K^T)\Delta(s)} + \frac{\eta_\theta(s) + \varepsilon(0)}{s + (A^T + K^T)\Delta(s)} \\ &= \frac{s^{n-1}G^T\Lambda(s)f(s)}{s^n + s^{n-1}(A^T + K^T)\Delta(s)} + \frac{s^{n-1}\eta_\theta(s) + s^{n-1}\varepsilon(0)}{s^n + s^{n-1}(A^T + K^T)\Delta(s)} \\ &= \frac{s^{n-1}G^T\Lambda(s)f(s)}{D(s)} + \frac{N(s)}{D(s)} \end{aligned} \tag{5.10}$$

where $N(s)$ and $D(s)$ are defined as in (4.34).

Using the final value theorem one obtains

$$\lim_{s\to 0} s\varepsilon(s) = \lim_{s\to 0}\left(\frac{s^n G^T\Lambda(s)f(s)}{D(s)} + \frac{sN(s)}{D(s)}\right) \tag{5.11}$$

In chapter 4, it was proved that $\lim_{s\to 0}\left(\frac{sN(s)}{D(s)}\right) = \frac{0}{a_0+k_0}$. So in this chapter, the focus will be on proving that the value of $\lim_{s\to 0}\frac{s^n G^T \Lambda(s) f(s)}{D(s)}$ allows the error estimation error to converge while the condition established in chapter 4 are still fulfilled so the results are still valid.

First, note that:

$$
\begin{aligned}
s^n G^T \Lambda(s) f(s) &= \left(\frac{s^n g_q}{s^{n-q-1}} + \cdots + \frac{s^n g_1}{s^{n-2}} + \frac{s^n g_0}{s^{n-1}}\right) f(s) \\
&= s\left(g_q s^q + \cdots + g_1 s + g_0\right) f(s) \qquad (5.12)
\end{aligned}
$$

Therefore,

$$
\lim_{s\to 0}\frac{s^n G^T \Lambda(s) f(s)}{D(s)} = \lim_{s\to 0}\left(\frac{\left(g_q s^q + \cdots + g_1 s^2 + g_0\right) s f(s)}{s^n + \displaystyle\sum_{i=1}^{n-1}\left(a_i + k_i\right) s^i + \left(a_0 + k_0\right)}\right) \qquad (5.13)
$$

If $f(t)$ is approximated by a polynomial of order $l \leq q$ then the fault detectability condition can be set following the first non-zero element of the fault dynamics input matrix G. Indeed, if for instance $g_0 \neq 0$ then $\lim\limits_{s\to 0} s\varepsilon(t) \neq 0$, which means that all type of faults can be detected. When if $g_0 = 0$ and $g_1 \neq 0$ then $\lim\limits_{s\to 0} s\varepsilon(t) \neq 0$ if $f(t)$ is at least of order 1. Thus, constant type fault signals are not detectable while detection of ramp type and higher order faults signals is achievable. Following this logic, if the first non zero element of G is g_f then the fault signal type that can be detected has to be at least of order f. Thus, the fault detectability condition depends on both the dynamics brought by the fault and the fault signal.

Now, assuming that the fault type and signal fulfill the fault detectability condition, and setting k_0 such that $a_0 + k_0 \neq 0$ then the residual will always contain a fault signature i.e. $\lim_{s\to 0} s\varepsilon(s) \neq 0$.

The above result shows that only k_0 is needed to insure that the fault signature will be present and furthermore signified in the residual steady state. Indeed k_0 value should be calculated such that the fault signal to residual ratio is high enough so a fault occurrence can be distinguished in the residual.

5.3 Observer gain optimisation for robust fault detector design

In this section is considered the system model used in previous section to which is added disturbances dynamics

$$\begin{aligned} y^{(n)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) &= \left[b_m u^{(m)}(t) + \cdots + b_1\dot{u}(t) + b_0 u(t)\right] \\ &\quad + \left[g_q f^{(q)}(t) + \cdots + g_1\dot{f}(t) + g_0 f(t)\right] \\ &\quad + \left[e_p d^{(p)}(t) + \cdots + e_1\dot{d}(t) + e_0 d(t)\right] \end{aligned} \tag{5.14}$$

where $e(t)$ represents the considered disturbance acting on the system through its disturbance input parameters $\begin{pmatrix} e_q & \ldots & e_1 & e_0 \end{pmatrix}$.

Note that the general form of dynamics brought by disturbances that can effect the system, the actuator or the sensor is used here. For instance,in the case of a Gaussian noise, the successive derivatives of the disturbance will not be needed thus only e_0 will be nonzero.

First, Equation (5.14) is rewritten as

$$y^{(n)}(t) = -A^T Y + B^T U + G^T F + E^T D \tag{5.15}$$

where

$$\begin{aligned} E^T &= \begin{pmatrix} e_p & e_{p-1} & \dots & e_1 & e_0 \end{pmatrix} \\ D^T &= \begin{pmatrix} d^{(p)} & d^{(p-1)} & \dots & \dot{d} & d \end{pmatrix} \end{aligned} \tag{5.16}$$

Following the same steps than in previous section, the output dynamic expression is

$$\dot{y}(t) = -A^T \mathcal{Y} + B^T \mathcal{U} + G^T \mathcal{F} + E^T \mathcal{D}\,(t) + \bar{\xi}_0(t) + \bar{\eta}_\theta(t) \tag{5.17}$$

denote

$$\mathcal{D}^T\,(t) = \mathcal{I}_{n-1}\left\{D(t)\right\} = \begin{pmatrix} \mathcal{I}_{n-p-1}\left\{d(t)\right\} & \dots & \mathcal{I}_{n-2}\left\{d(t)\right\} & \mathcal{I}_{n-1}\left\{d(t)\right\} \end{pmatrix} \tag{5.18}$$

and $\bar{\xi}_0(t)$ is the term containing the known initial conditions of $y(t)$, $u(t)$ and $d(t)$ while $\bar{\eta}_\theta(t)$ is a unknown (polynomial) function that is dependent on the unknown initial conditions of the system with $\theta = \left\{\dot{y}(0)..., \dot{u}(0)..., d(0), \dot{d}(0)...\right\}$.

Then, the observer is defined by

$$\dot{\hat{y}}(t) = -A^T \widehat{\mathcal{Y}} + B^T U + K^T \left(\mathcal{Y} - \widehat{\mathcal{Y}}\right) + \bar{\xi}_0(t) \tag{5.19}$$

Now, the residual dynamic expression is calculated as

$$\dot{\varepsilon}(t) = -\left(A^T + K^T\right)\left(\mathcal{Y} - \widehat{\mathcal{Y}}\right) + G^T\mathcal{F} + E^T\mathcal{D}\left(t\right) + \bar{\eta}_\theta(t) \tag{5.20}$$

Applying Laplace Transform to (5.20) gives

$$s\varepsilon(s) - \varepsilon(0) = -\left(A^T + K^T\right)\Delta(s)\varepsilon(s) + G^T\Lambda(s)f(s) + E^T\Pi\left(s\right)d\left(s\right) + \bar{\eta}_\theta(s) \tag{5.21}$$

where $d\left(s\right)$ is the Laplace transform of $d(t)$ and

$$\Pi\left(s\right)^T = \begin{pmatrix} \frac{1}{s^{n-p-1}} & \ldots & \frac{1}{s^{n-2}} & \frac{1}{s^{n-1}} \end{pmatrix} \tag{5.22}$$

Finally, one obtains

$$\varepsilon(s) = \frac{s^{n-1}G^T\Lambda(s)f(s)}{D(s)} + \frac{s^{n-1}E^T\Pi\left(s\right)d\left(s\right)}{D(s)} + \frac{N(s)}{D(s)} \tag{5.23}$$

It is already known from previous results that $\lim_{s\longrightarrow 0}\left(\frac{sN(s)}{D(s)}\right) = \frac{0}{(a_0+k_0)}$ and $\lim_{s\longrightarrow 0}\left(\frac{s^nG^T\Lambda(s)f(s)}{D(s)}\right) \neq 0$. Now, one needs to attenuate the disturbance contribution in the residual brought by $\frac{s^nG^T\Lambda(s)f(s)}{D(s)}$.

The method used here is based on finding the gain K value that minimises the disturbance to residual dynamic $\frac{E^T\Pi(s)d(s)}{s+(A^T+K^T)\Delta(s)}$ while the latter remains sensitive to the faults.

As shown in previous section, only k_f needs to be used to achieve good fault detection performances, so in this section, an optimisation method will be used to calculate the gain $K' = \begin{pmatrix} k_{n-1} & \dots & k_{f+1} & k_{f-1} & \dots & k_0 \end{pmatrix}^T$ value toward disturbance attenuation while k_f remains fixed as set in previous section.

The gain optimisation method used for finding the optimal K' for disturbances attenuation was proposed in [64] where a criterion to be minimised to reduce the disturbances effect in the residual is set. Indeed, the criterion is defined as $\min_{K'} \left\| \frac{E^T \Pi(s) d(s)}{s+(A^T+K^T)\Delta(s)} \right\|_{s=jw_d}$ with $w_d = 2\pi f_d$ and f_d is the identified frequency of the disturbance. For a white Gaussian, as the noise is spread equally over the range of frequency, one can use the infinity norm $\min_{K'} \left\| \frac{E^T \Pi(s) d(s)}{s+(A^T+K^T)\Delta(s)} \right\|_{\infty}$ instead of using and identified fixed value.

5.3.1 Extension to the multi-output case

The above output observer design can be extended to the multi-output case for systems of the form. Indeed, consider the system as described in (5.24)

$$\begin{aligned} A_n Y^{(n)}(t) &= -A_{n-1} Y^{(n-1)}(t) - \cdots - A_0 Y(t) + B_m U^{(m)}(t) + \cdots + B_0 U(t) \\ &\quad + G_l F^{(l)}(t) + \cdots + G_0 F(t) \end{aligned} \tag{5.24}$$

where

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}, U = \begin{pmatrix} u_1 \\ \vdots \\ u_q \end{pmatrix} \text{ and } F = \begin{pmatrix} f_1 \\ \vdots \\ f_r \end{pmatrix} \tag{5.25}$$

where $A_i \in R^{p \times p}$, $B_i \in R^{p \times q}$ and $G_i \in R^{p \times r}$.

By integrating system (5.24) $n-1$ times with respect to time, one obtains

$$\begin{aligned} A_n\dot{Y}(t) &= \begin{pmatrix} -A_{n-1} & \dots & -A_0 \end{pmatrix} \begin{pmatrix} Y(t) \\ \vdots \\ \mathcal{I}_{n-1}\left\{Y(t)\right\} \end{pmatrix} \\ &+ \begin{pmatrix} B_m & \dots & B_0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_{n-m-1}\left\{U(t)\right\} \\ \vdots \\ \mathcal{I}_{n-1}\left\{U(t)\right\} \end{pmatrix} \\ &+ \begin{pmatrix} G_l & \dots & G_0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_{n-m-1}\left\{F(t)\right\} \\ \vdots \\ \mathcal{I}_{n-1}\left\{F(t)\right\} \end{pmatrix} + \Xi_0\left(t\right) + \Omega(\theta, t) \end{aligned} \tag{5.26}$$

where $\Xi(t,\theta)$ and $\Omega(\theta,t)$ are polynomials functions containing respectively the known and unknown initial conditions, .

Then, the output observer can be designed similarly to the approach used for SISO systems as follows

$$A_n\dot{\hat{Y}}(t) = \begin{pmatrix} -A_{n-1} & \dots & -A_0 \end{pmatrix} \begin{pmatrix} Y(t) \\ \vdots \\ \mathcal{I}_{n-1}\{Y(t)\} \end{pmatrix} + \begin{pmatrix} B_m & \dots & B_0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_{n-m-1}\{U(t)\} \\ \vdots \\ \mathcal{I}_{n-1}\{U(t)\} \end{pmatrix} + \begin{pmatrix} K_n & \dots & K_0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_{n-m-1}\left\{Y(t)-\hat{Y}(t)\right\} \\ \vdots \\ \mathcal{I}_{n-1}\left\{Y(t)-\hat{Y}(t)\right\} \end{pmatrix} + \Xi_0(t) \quad (5.27)$$

In order to evaluate the observer performances, the estimation error $\varepsilon(t)$ is set defined by $\varepsilon(t) = Y(t) - \hat{Y}(t)$, and thus the error dynamic is calculated as

$$A_n\dot{\varepsilon}(t) = \begin{pmatrix} -A_{n-1}-K_{n-1} & \dots & -A_0-K_0 \end{pmatrix} \begin{pmatrix} \varepsilon(t) \\ \vdots \\ \mathcal{I}_{n-1}\{\varepsilon(t)\} \end{pmatrix} + \begin{pmatrix} G_l & \dots & G_0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_{n-m-1}\{F(t)\} \\ \vdots \\ \mathcal{I}_{n-1}\{F(t)\} \end{pmatrix} + \Omega(\theta, t) \quad (5.28)$$

Now, calculating the Laplace transform of the error dynamic gives

$$
\begin{aligned}
A_n s\varepsilon(s) &= \begin{pmatrix} -A_{n-1}-K_{n-1} & \ldots & -A_0-K_0 \end{pmatrix} \begin{pmatrix} \varepsilon(s) \\ \vdots \\ \frac{1}{s^{n-1}}\{\varepsilon(s)\} \end{pmatrix} \\
&\quad + \begin{pmatrix} G_l & \ldots & G_0 \end{pmatrix} \begin{pmatrix} \frac{1}{s^{n-m-1}}\{F(s)\} \\ \vdots \\ \frac{1}{s^{n-1}}\{F(s)\} \end{pmatrix} + \Omega(\theta, s) \qquad (5.29)
\end{aligned}
$$

thus

$$
\begin{aligned}
&\varepsilon(s) \begin{pmatrix} A_n & A_{n-1}+K_{n-1} & \ldots & A_0+K_0 \end{pmatrix} \begin{pmatrix} sI_n \\ I_n \\ \vdots \\ \frac{1}{s^{n-1}}I_n \end{pmatrix} \varepsilon(s) \\
&= \begin{pmatrix} G_l & \ldots & G_0 \end{pmatrix} \begin{pmatrix} \frac{1}{s^{n-m-1}}\{F(s)\} \\ \vdots \\ \frac{1}{s^{n-1}}\{F(s)\} \end{pmatrix} + \Omega(\theta, s) \qquad (5.30)
\end{aligned}
$$

where $I_p \in R^{n\times n}$ is the n-dimensional identity matrix.

Finally, is obtained

$$s\varepsilon\left(s\right) = \left(\begin{pmatrix} A_n & A_{n-1}+K_{n-1} & \dots & A_0+K_0 \end{pmatrix}\begin{pmatrix} sI_n \\ I_n \\ \vdots \\ \frac{1}{s^{n-1}}I_n \end{pmatrix}\right)^{-1}$$

$$\times\left(\begin{pmatrix} G_l & \dots & G_0 \end{pmatrix}\begin{pmatrix} \frac{1}{s^{n-m-1}}\left\{F(s)\right\} \\ \vdots \\ \frac{1}{s^{n-1}}\left\{F(s)\right\} \end{pmatrix} + \xi_1(t,\theta)\right) \quad (5.31)$$

The same approach used for the single output case can be used to determine the gains matrix values for fault detection as instead of dealing with simple fractions, it requires handling matrices. It is worth noting though that in order to continue the analysis of the fault detection scheme and calculating the gain matrix to achieve good fault detection performances, the admissible values of the gain matrix $\begin{pmatrix} K_n & \dots & K_0 \end{pmatrix}$ are those that make the following matrix $\begin{pmatrix} A_n & A_{n-1}+K_{n-1} & \dots & A_0+K_0 \end{pmatrix}$ of full order so it is indeed invertible. This method can also be extended to systems under disturbances as in previous section.

5.4 Modelling of three wheeled robot model

A three wheeled robot model has been built to validate the fault detection scheme performances. In the modelled robot, the rear axle which connected to two wheels is driven by two DC motors connected to each wheel. The system input is thus, a two dimensioned

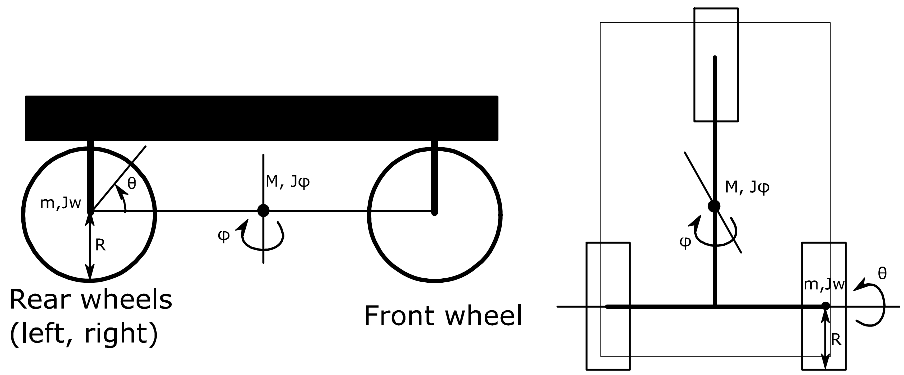

Figure 5.1: Three wheeled robot model

vector that contains both the motors voltages v_l and v_r. The system output is also a two dimensional vector which contains each of the rear axle wheels positions y_l and y_r.

The simplified mechanical model of the three wheeled robot is shown in Fig. 5.1. The system generalised dynamics are given by

$$\left[(2m+M)R^2 + mR^2 + 2n^2 J_m\right]\ddot{\theta} = \frac{\alpha}{2}(v_l + v_r) - (\beta + f_w)\dot{\theta} \tag{5.32}$$

$$\begin{aligned} &\left[\frac{1}{2}mW^2 + \frac{M\left(W^2+L^2\right)}{12} + \frac{W^2}{2R^2}\left(J_w + n^2 J_m\right)\right]\ddot{\phi} \\ = \ & \frac{R}{W}\alpha(v_l - v_r) - \left(\beta + \frac{W}{R}f_w\right)\dot{\phi} \end{aligned} \tag{5.33}$$

where θ is the robot translational motion and ϕ is the robot yaw angle

M	Body mass	J_w	Wheel inertia moment
m	Wheel mass	n	Gear ratio
R	Wheel radius	f_w	Wheel/road friction coefficient
W	Body width	α, β	Electrical motor coefficients
L	Body length	J_m	Electrical motor inertia

Now, consider a fault occurring in the system and which is affecting the robot mass such that: $M \to M + \Delta M$ and a disturbance d acting on the system. Equations (5.32) and (5.33) become

$$\begin{aligned} & \left[(2m+M)R^2 + mR^2 + 2n^2 J_m\right]\ddot{\theta} + \Delta M R^2 \ddot{\theta} \\ = \; & \frac{\alpha}{2}(v_l + v_r) - (\beta + f_w)\dot{\theta} + d; \end{aligned} \tag{5.34}$$

$$\begin{aligned} & \left[\frac{1}{2} m W^2 + \frac{M\left(W^2 + L^2\right)}{12} + \frac{W^2}{2R^2}\left(J_w + n^2 J_m\right)\right]\ddot{\phi} + \frac{\Delta M\left(W^2 + L^2\right)}{12}\ddot{\phi} \\ = \; & \frac{R}{W}\alpha(v_l - v_r) - \left(\beta + \frac{W}{R} f_w\right)\dot{\phi} + d \end{aligned} \tag{5.35}$$

The output equations are

$$y_r = \frac{180}{\pi}\left(\theta - \frac{W}{2R}\phi\right) \tag{5.36}$$

$$y_l = \frac{180}{\pi}\left(\theta + \frac{W}{2R}\phi\right) \tag{5.37}$$

By solving simultaneously (5.37) and (5.36), one gets

$$\begin{cases} \theta = \frac{\pi}{360}\left(y_l + y_r\right) \\ \phi = \frac{\pi R}{180W}\left(y_l - y_r\right) \end{cases} \tag{5.38}$$

Consequently

$$c_1\left(\ddot{y}_l + \ddot{y}_r\right)\frac{\pi}{360} = \frac{\alpha}{2}\left(v_l + v_r\right) - \frac{\pi}{360}\left(\beta + f_w\right)\left(\dot{y}_l + \dot{y}_r\right) - R^2\ddot{f}_\theta \tag{5.39}$$

$$c_2\left(\ddot{y}_l - \ddot{y}_r\right)\frac{\pi R}{180W} = \frac{R}{W}\alpha\left(v_l - v_r\right) - \frac{\pi R}{180W}\left(\beta + \frac{W}{R}f_w\right)\left(\dot{y}_l - \dot{y}_r\right) - \frac{\left(W^2 + L^2\right)}{12}\ddot{f}_\phi \tag{5.40}$$

where

$$c_1 = \left(2m + M\right)R^2 + mR^2 + 2n^2 J_m, \tag{5.41}$$

$$c_2 = \left[\frac{1}{2}mW^2 + \frac{M\left(W^2 + D^2\right)}{12} + \frac{W^2}{2R^2}\left(J_w + n^2 J_m\right)\right] \tag{5.42}$$

Now, by setting the fault signals

$$\begin{cases} f_\theta = \Delta M\theta \\ f_\phi = \Delta M\phi \end{cases} \tag{5.43}$$

the system is written in the following input/output representation

$$
\begin{aligned}
&\begin{pmatrix} c_1 \frac{\pi}{360} & c_1 \frac{\pi}{360} \\ c_2 \frac{\pi R}{180W} & -c_2 \frac{\pi R}{180W} \end{pmatrix} \begin{pmatrix} \ddot{y}_l \\ \ddot{y}_r \end{pmatrix} \\
= &\begin{pmatrix} -\frac{\pi}{360} \left(\beta + f_w\right) & -\frac{\pi}{360} \left(\beta + f_w\right) \\ -\frac{\pi R}{180W} \left(\beta + \frac{W}{R} f_w\right) & \frac{\pi R}{180W} \left(\beta + \frac{W}{R} f_w\right) \end{pmatrix} \begin{pmatrix} \dot{y}_l \\ \dot{y}_r \end{pmatrix} \\
&+ \begin{pmatrix} \frac{\alpha}{2} & \frac{\alpha}{2} \\ \frac{R}{W}\alpha & -\frac{R}{W}\alpha \end{pmatrix} \begin{pmatrix} v_l \\ v_r \end{pmatrix} \\
&+ \begin{pmatrix} -R^2 & 0 \\ 0 & \frac{-\left(W^2+D^2\right)}{12} \end{pmatrix} \begin{pmatrix} \ddot{f}_\theta \\ \ddot{f}_\phi \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} d
\end{aligned} \tag{5.44}
$$

and finally, is obtained

$$
\begin{aligned}
&\begin{pmatrix} \ddot{y}_l \\ \ddot{y}_r \end{pmatrix} \\
= &\begin{pmatrix} -\frac{1}{2c_1} \left(\beta + f_w\right) - \frac{1}{2c_2} \left(\beta + \frac{1}{R} W f_w\right) & -\frac{1}{2c_1} \left(\beta + f_w\right) + \frac{1}{2c_2} \left(\beta + \frac{1}{R} W f_w\right) \\ -\frac{1}{2c_1} \left(\beta + f_w\right) + \frac{1}{2c_2} \left(\beta + \frac{1}{R} W f_w\right) & -\frac{1}{2c_1} \left(\beta + f_w\right) - \frac{1}{2c_2} \left(\beta + \frac{1}{R} W f_w\right) \end{pmatrix} \begin{pmatrix} \dot{y}_l \\ \dot{y}_r \end{pmatrix} \\
&+ \begin{pmatrix} \frac{90}{\pi} \frac{\alpha}{c_1} + \frac{90}{\pi} \frac{\alpha}{c_2} & \frac{90}{\pi} \frac{\alpha}{c_1} - \frac{90}{\pi} \frac{\alpha}{c_2} \\ \frac{90}{\pi} \frac{\alpha}{c_1} - \frac{90}{\pi} \frac{\alpha}{c_2} & \frac{90}{\pi} \frac{\alpha}{c_1} + \frac{90}{\pi} \frac{\alpha}{c_2} \end{pmatrix} \begin{pmatrix} v_l \\ v_r \end{pmatrix} \\
&+ \begin{pmatrix} -\frac{180}{\pi} \frac{R^2}{c_1} & \frac{90}{\pi R} \frac{W}{c_2} \left(-\frac{1}{12} L^2 - \frac{1}{12} W^2\right) \\ -\frac{180}{\pi} \frac{R^2}{c_1} & -\frac{90}{\pi R} \frac{W}{c_2} \left(-\frac{1}{12} L^2 - \frac{1}{12} W^2\right) \end{pmatrix} \begin{pmatrix} \ddot{f}_\theta \\ \ddot{f}_\phi \end{pmatrix} + \begin{pmatrix} \frac{180}{\pi c_1} + \frac{90}{\pi R} \frac{W}{c_2} \\ \frac{180}{\pi c_1} - \frac{90}{\pi R} \frac{W}{c_2} \end{pmatrix} d
\end{aligned} \tag{5.45}
$$

Finally replacing the systems parameters by their numerical values give

$$\begin{pmatrix} \ddot{y}_l \\ \ddot{y}_r \end{pmatrix} = \begin{pmatrix} -18.6448 & -3.0418 \\ -3.0418 & -18.6448 \end{pmatrix} \begin{pmatrix} \dot{y}_l \\ \dot{y}_r \end{pmatrix} + \begin{pmatrix} 2076.6 & 338.8 \\ 338.8 & 2076.6 \end{pmatrix} \begin{pmatrix} v_l \\ v_r \end{pmatrix} + \begin{pmatrix} 75.1420 & 106.1828 \\ 75.1420 & 106.1828 \end{pmatrix} \begin{pmatrix} \ddot{f}_\theta \\ \ddot{f}_\phi \end{pmatrix} + \begin{pmatrix} 77.581 \\ 77.581 \end{pmatrix} d \tag{5.46}$$

It can be seen here that by considering a mass fault in the system, both wheel rotational angle and yaw rate are affected. Thus, a fault can be detected if there is motion following one of these axes so that fault detectability conditions are fulfilled.

5.5 Simulation results

To validate the fault detection scheme, for sake of simplicity, the case where the robot moves following a straight line path (i.e. turning angle equal to zero) is considered. In this case as the yaw angles is always equal to zero, only f_θ can be detected which is sufficient for the purpose of this work.

First the observer simulations in absence of fault are carried out to validate the performances of the observer. Fig.5.2. shows that the system desired and achieved output (robot wheels speed) fairly match. Indeed, the observer 5% response time (i.e. output within 5% of the final value) is less than 10s, the system response overshoot is about 10 % and the static error is close to zero. As the purpose of this work is to evaluate FD

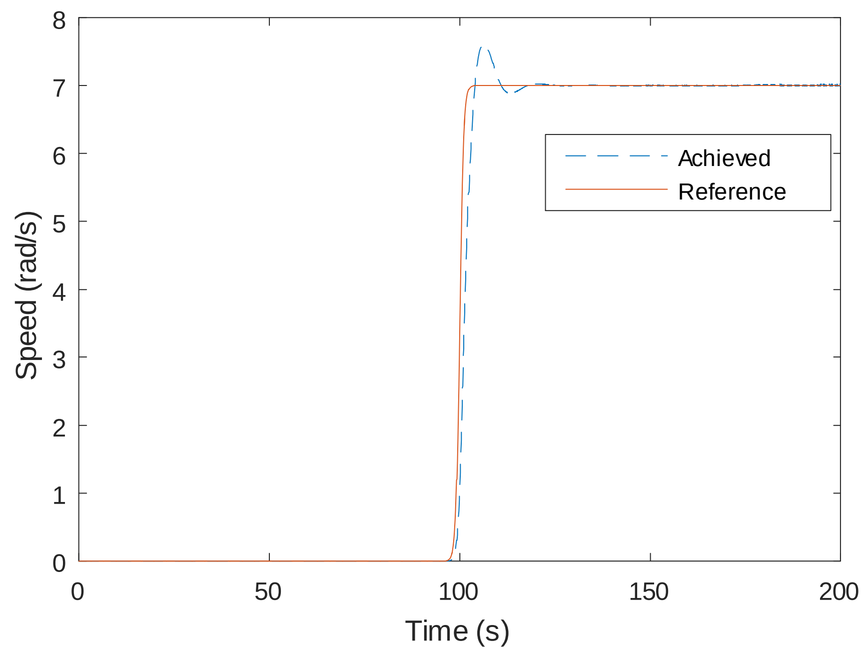

Figure 5.2: Three robot wheels desired and achieved speed

performances, these results are sufficient to proceed with simulation analysis.

The fault considered here is a system fault (change of robot's mass) that is set to 10% less than the original. The value of the fault over time is described by the following formula

$$f(t) = \begin{cases} 0 & (t < 30) \\ 0.1t & (30 \leq t < 40) \\ 1 & t \geq 40 \end{cases} \tag{5.47}$$

and the disturbance signal $d(t)$ set to be a Gaussian noise with mean value $\mu = 0$ and variance $\sigma^2 = 1$. This means that the disturbances and the faults are in the same range. It is then, up to the fault detection scheme to show the robust fault detection performances.

Now, using simulations results, the spectrum analysis of the residual is given, in

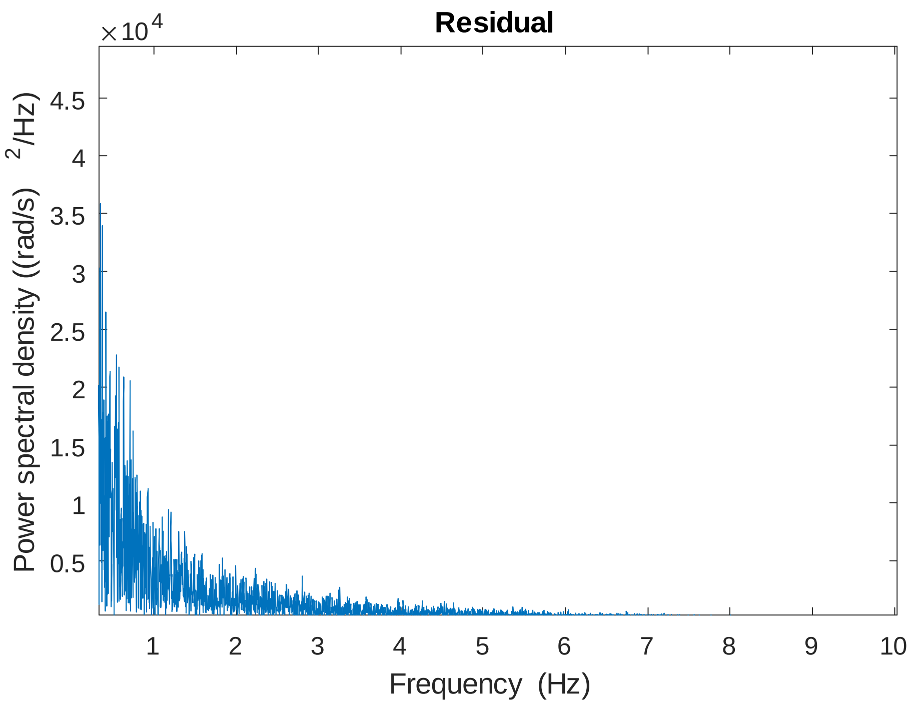

Figure 5.3: Residual power spectrum - Fault free system

absence of fault Fig. 5.3 and in presence of fault Fig. 5.4.

By analysing the residual spectrums, it can be seen that a significant difference between the fault free and the fault free cases lies between 0 and $1Hz$. Thus, it can be concluded that the fault contribution in the residual is mainly among low frequencies.. So, the frequency to be used in the optimisation criterion is set as $f_f = 0.2Hz$.

To validate the performance of the FD scheme, it can be seen in simulation results of the residual in absence of fault as shown in Fig. 5.5 and, in presence of fault as shown in Fig. 5.6, that when the fault occurs at time $30s$, the residual value deviates considerably from zero. Indeed, the fault signal is signified and a threshold value can be chosen such that false alarms triggering is avoided. Besides, the residual is different from zero for a long

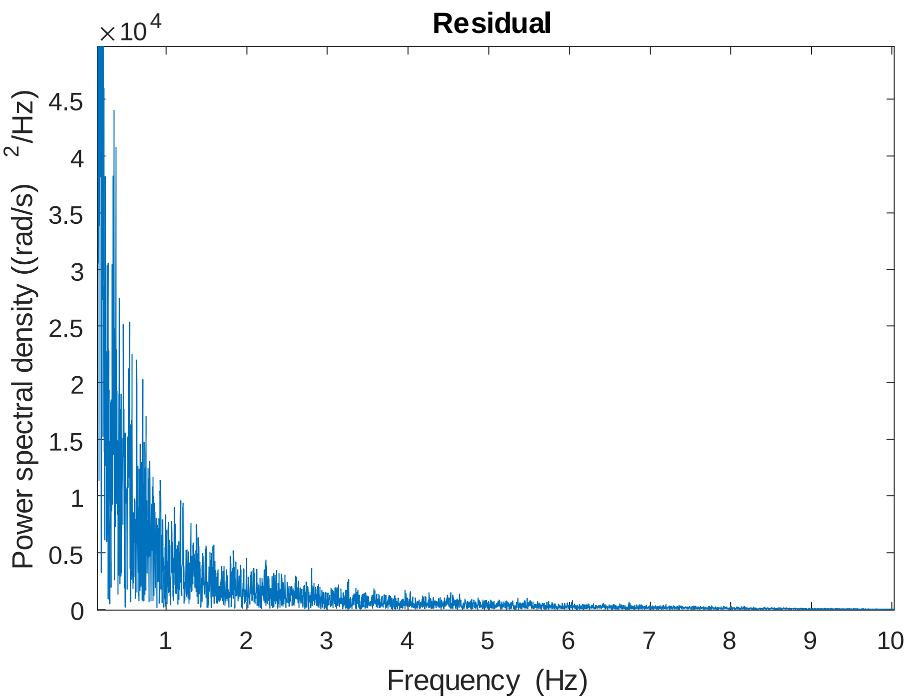

Figure 5.4: Residual power spectrum - Faulty system

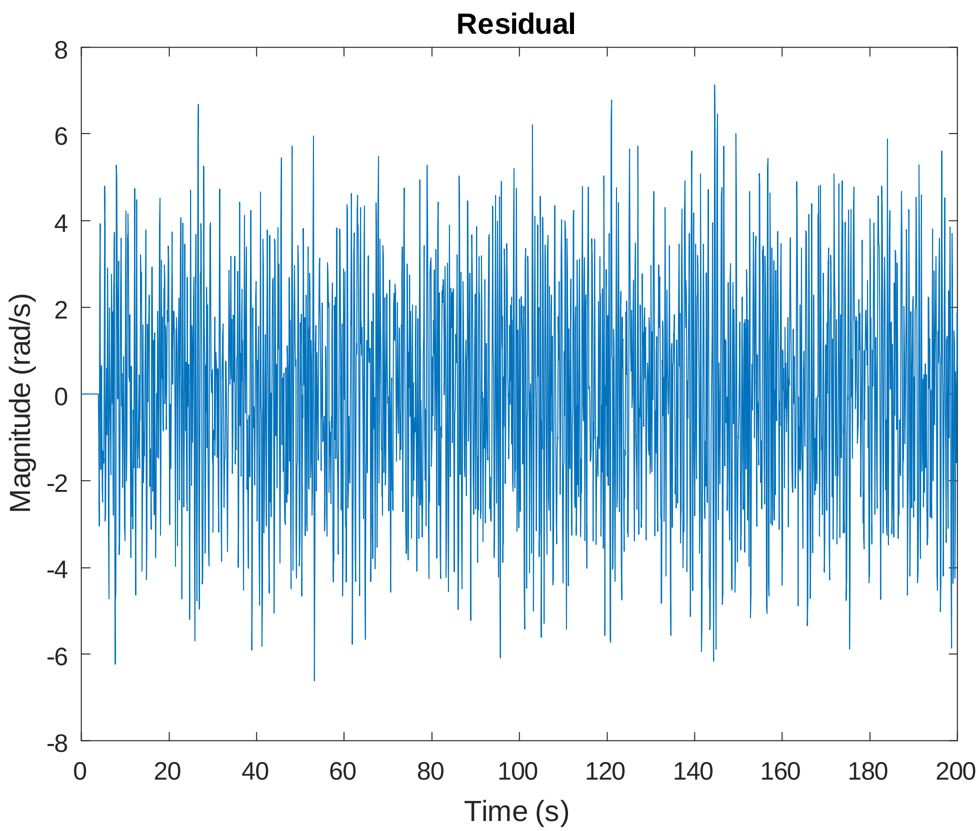

Figure 5.5: Residual - Fault free system

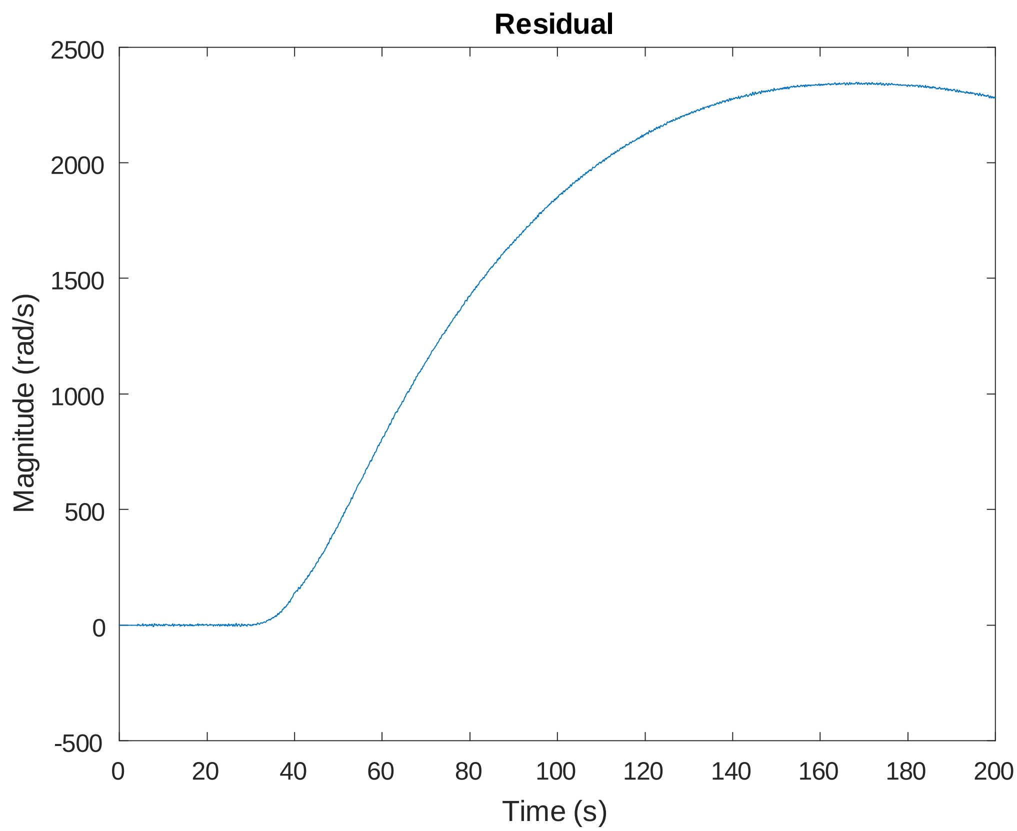

Figure 5.6: Residual - Faulty system

period after the first occurs in the systems, which means that the fault can be detected over a long period of time although its value can not maintained at such high value and may start decreasing after 180s due to the observer compensating for the estimation error.

5.6 Conclusion

In this chapter, the fault detection technique using the output observer has been presented and its application to fault detection in systems under disturbances has been shown. Indeed, a systematic method to calculate the observer gain, using an optimisation criterion, that ensures having best performances for both attenuating the disturbance and signifying the fault signature in the residual is proposed. This approach is also extended to the MIMO case. Simulations results using a three wheeled robot, which model has been built in a MIMO form, have been presented to validate the proposed method performances.

Chapter 6

Output observer for time delayed systems

6.1 Introduction

Time delay is a challenge that needs to be considered when studying controlled dynamical systems, where it mainly occurs due to actuation, measurement, data processing and transfer delays, and may jeopardize the stability of the system.

Furthermore when working on fault detectors design, delays occurring in the system need to be considered as their effect can disturb the behavior of the system. If not included in the fault detector design, its effect may result in increasing the estimation error which could result in rising false alarms.

As delays can occur in several parts of the systems (actuator, plant, sensor, controller...), an efficient way to deal with such differet sources of delays is to fix a time reference point in the system at which the delay will be assumed to be the sum of all the delays occurring in the system. This allows to consider the delay only once in the system's model which will help simplifying the model expression. In the considered case, the reference point is chosen to be the moment the measurement is received from the sensors.

In this chapter and for sake of simplicity, in the model expression, faults nor disturbances occurring in the system are considered. Indeed, considering that it is believed that these two topics have been thoroughly studied in previous chapters and that it is straight forward to combine the previously proposed techniques with the one proposed here for time delays in systems.

6.2 System representation

In Fig. 6.1. the system considered in this work, is displayed. Two parts of the systems can be distinguished, one being the controller/observer side, and the other one being the physical system which is composed of the actuator, plant and the sensors. The two parts exchange information through a communication network.

As it is believed that the main source of delays is the time it takes the data to travel through the communication network, the two delays that are considered are the input delay which is the time it takes to the information to travel from the controller/observer to the physical system and the output delay which is the time for the measurements to arrive to the controller/observer.

So the controller calculated input is subject to a delay τ^a which results on the actuator using the delayed input. Similarly, the measured output is subject to a delay τ^s.

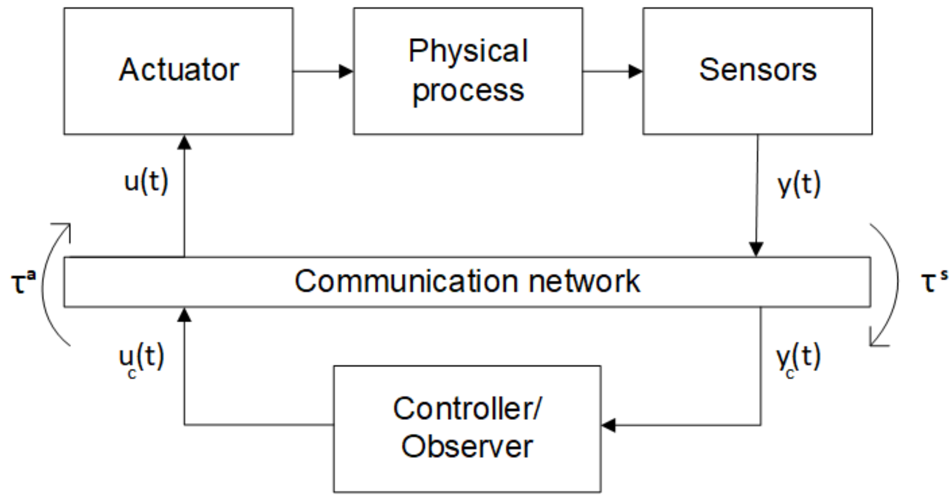

Figure 6.1: Time delay in control systems

As it is chosen to set the time reference point to be the moment the measurement arrives to the controller/observer, the total delay τ considered in the system model will be the sum of the two delays $\tau = \tau^a + \tau^s$.

6.2.1 Model in absence of delay

The proposed model to be used in this section in absence of delay for the considered nonlinear system is

$$\frac{dy(t)}{dt} = g(u, y, \check{I}_{n-1}(y,u)) + f(w,t) \tag{6.1}$$

where, similarly to the previously developed input/output representation, $\check{I}_{n-1}(y,u) = (\mathcal{I}_1(y), \mathcal{I}_1(u), ..., \mathcal{I}_{n-1}(y), \mathcal{I}_{n-1}(u))$ and $f(w,t) = w^T v$ with $w = (y^{(1)}(0), ..., y^{(n-1)}(0)) \in R^{n-1}$ is a vector of unknown initial conditions and $v^T = \left(1, t, ..., \frac{t^{n-1}}{(n-1)!}\right)$. this means, $f(w,t)$ is a polynomial of order $n-1$ in t with coefficients $y^{(i)}(0)$. $\mathcal{I}_k\{y(t)\}$ and

$i \in \{0, 1, ..., n-1\}$ denotes the k th integrations of the function $y(t)$ with respect to time; that is $\mathcal{I}_k\{y(t)\} = \underbrace{\int_0^t \cdots \int_0^t}_{k \text{ times}} y(\tau)d\tau...d\tau$ for $k \geq 0$. In particular, $\mathcal{I}_0\{y(t)\} = y(t)$.

In Chapter 5, a fault detection scheme has been discussed thus the proof also applies for the system considered in this chapter. So here, it is chosen to rather focus on dealing with the difficulties that considering time delay in the system brings. In other words, a solution where the delay will be dealt with in a way that the observer built in Chapter 4 still works, is proposed. In order to achieve this goal, the stability analysis will be needed to be take further than previously.

6.2.2 System model with delay

As mentioned above, it is chosen to use only one delay in the model which is $\tau = \tau^a + \tau^s$. To show the proposed solution for delay in systems, the delay in the previously proposed model is introduced as follows:

$$\begin{cases} \frac{dy(t)}{dt} = f(w,t) + g(u, y, \check{I}_{n-1}(y,u)) \\ y_m(t) = y(t-\tau) \end{cases} \tag{6.2}$$

where $y_m(t)$ is the measured output and τ is the measurement delay.

6.3 Delay effect attenuation techniques

In the following, two cases of delays are considered. First a simpler case where the delay is a priori known/estimated is considered, then a solution for a general case where the delay is unknown will be is proposed.

6.3.1 Case i): known delay

Consider the case where the delay is known and can be included in the observer design. This case is possible whether the delay information is already available in the system or can be estimated using a separate tool.

So, rewriting the output observer for the above system gives:

$$\begin{aligned}\frac{d\hat{y}(t)}{dt} &= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) + K^T\left(\mathcal{Y}_m(t) - \widehat{\mathcal{Y}}(t-\tau)\right) \\ &= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) + K^T\left(\mathcal{Y}(t-\tau) - \widehat{\mathcal{Y}}(t-\tau)\right) \end{aligned} \tag{6.3}$$

In this case the error dynamics is given by:

$$\begin{aligned}\frac{d\varepsilon(t)}{dt} &= f(w,t) - f(\hat{w},t) - K^T\left(\mathcal{Y}(t-\tau) - \widehat{\mathcal{Y}}(t-\tau)\right) \\ &= \left(w^T - \hat{w}^T\right)v(t) - K^T\left(\mathcal{Y}(t-\tau) - \widehat{\mathcal{Y}}(t-\tau)\right) \end{aligned} \tag{6.4}$$

As the error expression here is the same as in Chapter 4 with the difference that both measured and estimated output are delayed, thus the same proof using output injection will apply. So in the following, it will be proceeded by following the same steps.

Taking the Laplace transform of (6.4), gives

$$\begin{aligned} sE(s) - \varepsilon(0) &= \left(w^T - \hat{w}^T\right)\mathcal{L}\{v(t)\} - K_0E(s)e^{-\tau s} \\ &\quad -\frac{1}{s}K_1E(s)e^{-\tau s} + \ldots + \frac{1}{s^{n-1}}K_{n-1}E(s)e^{-\tau s} \end{aligned} \tag{6.5}$$

In other words,

$$\begin{aligned}
&\left[s + e^{-\tau s} K_1 + \frac{1}{s} K_2 e^{-\tau s} + \ldots + \frac{1}{s^{n-1}} K_{n-1} E(s) e^{-\tau s}\right] E(s) \\
= \ & \varepsilon(0) + \left(w^T - \hat{w}^T\right) \mathcal{L}\{v(t)\} \qquad (6.6)
\end{aligned}$$

Hence,

$$\begin{aligned}
&\left[s + e^{-\tau s} K_1 + \frac{1}{s} K_2 e^{-\tau s} + \ldots + \frac{1}{s^{n-1}} K_{n-1} E(s) e^{-\tau s}\right] E(s) \\
= \ & \varepsilon(0) + \sum_{i=1}^{n} \left(y^{(i)}(0) - \hat{y}^{(i)}(0)\right) \frac{1}{s^i} \qquad (6.7)
\end{aligned}$$

That is,

$$E(s) = \frac{s\varepsilon(0) + s \sum_{i=1}^{n} \left(y^{(i)}(0) - \hat{y}^{(i)}(0)\right) \frac{1}{s^i}}{\left[s + e^{-\tau s} K_1 + \frac{1}{s} K_2 e^{-\tau s} + \ldots + \frac{1}{s^{n-1}} K_{n-1} E(s) e^{-\tau s}\right]} \qquad (6.8)$$

Using the final value theorem, is obtained

$$\begin{aligned}
\lim_{t \to +\infty} \varepsilon(t) &= \lim_{s \to 0} \{sE(s)\} \\
&= \frac{s^2\varepsilon(0) + s^2 \sum_{i=1}^{n} \left(y^{(i)}(0) - \hat{y}^{(i)}(0)\right) \frac{1}{s^i}}{\left[s + e^{-\tau s} K_1 + \frac{1}{s} K_2 e^{-\tau s} + \ldots + \frac{1}{s^{n-1}} K_{n-1} E(s) e^{-\tau s}\right]} \\
&= \frac{0}{k_0} \qquad (6.9)
\end{aligned}$$

It can be seen that when the delay is known, its effect can be attenuated and the system stability is proved. Furthermore, it can be considered that the delay is not exactly known, thus the difference between the estimated and actual delay will result on having an additional term in the expression of the error dynamic that will be treated as a disturbance. Thus using the approach presented in chapter 5 for systems under disturbances, its effect can be attenuated.

6.3.2 Case ii): unknown delay

The general way to deal with delays is to assume that it is unknown so no constraining assumption is made. This is different from the previous section in the sense that the delay is considered too random to be estimated with delay estimation error being considered as a disturbance.

In the previous section, assuming the delay to be known simplifies the expression of the residual as the delay is used in the observer which allows the solution to be straight forward. But when dealing with an unknown delay, that is not possible which means that as the delay will be present in the residual expression, when calculating the observer gain, the delay impact should be reduced.

The following observer for the considered system is proposed:

$$\begin{aligned}
\frac{d\hat{y}(t)}{dt} &= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) \\
&\quad + K_0(y_m(t) - \hat{y}(t)) + K_1 \int_0^t (y_m(\tau) - \hat{y}(\tau))d\tau \\
&= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) \\
&\quad + K_0(y(t-\tau) - \hat{y}(t)) + K_1 \int_0^t (y(t-\tau) - \hat{y}(\tau))d\tau
\end{aligned} \tag{6.10}$$

To rewrite the above equation in a way that will help in calculating the gain $K_{0,1}$, it is reformulated using $\varepsilon(t) = y(t) - \hat{y}(t)$; that is, $\hat{y}(t) = y(t) - \varepsilon(t)$, so one has

$$\begin{aligned}
\frac{d\hat{y}(t)}{dt} &= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) \\
&\quad +K_0(y(t-\tau) - y(t) + \varepsilon(t)) + K_1 \int_0^t (y(\lambda-\tau) - y(\lambda) + \varepsilon(\lambda))d\lambda \\
&= f(\hat{w},t) + g(u,y,\check{I}_{n-1}(y,u)) + K_0\varepsilon(t) + K_1 \int_0^t \varepsilon(\lambda)d\lambda \\
&\quad +K_0(y(t-\tau) - y(t)) + K_1 \int_0^t (y(\lambda-\tau) - y(\lambda))d\lambda \qquad (6.11)
\end{aligned}$$

When dealing with the equation above, the challenge is to rewrite it in a way that the delay can be extracted from the expression of $y(t-\tau)$ so that the gains $K_{0,1}$ can be calculated to enhance the performances of the observer. To do so, it is chosen to use the mean value theorem, as it allows to extract the delay without using any approximation. To use this method, $y(t)$ must be continuous within the segment $[t, t-\tau]$ which is a fair assumption to make if it is considered that for the considered continuous time discrete measurement system, the delay is not bigger than the sampling time.

Thus, the following is obtained,

$$y(t) - y(t-\tau) = \frac{dy}{dt}(\gamma)\tau \qquad (6.12)$$

for some (unknown) γ inside the segment $[t, t-\tau]$. The delay is now extracted from the expression $y(t-\tau)$ and is multiplied by $\frac{dy}{dt}(\gamma)$ which is a constant value that represents the evaluation of the derivative of $y(t)$ at γ.

Then,

$$\begin{aligned}\frac{d\hat{y}(t)}{dt} &= f(\hat{w},t)+g(u,y,\check{I}_{n-1}(y,u)) \\ &\quad +K_0\varepsilon(t)+K_1\int_0^t \varepsilon(\lambda)d\lambda \\ &\quad +K_1\frac{dy}{dt}(\gamma)\tau+K_2\frac{dy}{dt}(\gamma)\tau\int_0^t \lambda d\lambda \\ &= f(\hat{w},t)+g(u,y,\check{I}_{n-1}(y,u)) \\ &\quad +K_0\varepsilon(t)+K_1\int_0^t \varepsilon(\lambda)d\lambda \\ &\quad +K_0\frac{dy}{dt}(\gamma)\tau+K_1\frac{dy}{dt}(\gamma)\tau t \end{aligned} \tag{6.13}$$

Now, the error dynamic $\frac{d\varepsilon(t)}{dt}$ expression can be given by:

$$\begin{aligned}\frac{d\varepsilon(t)}{dt} &= \left(w^T-\hat{w}^T\right)v(t)-K_0\varepsilon(t)-K_1\int_0^t \varepsilon(\lambda)d\lambda \\ &\quad -K_0\frac{dy}{dt}(\gamma)\tau-K_1\frac{dy}{dt}(\gamma)\tau t \end{aligned} \tag{6.14}$$

In the following, a methodology to reduce the delay effect in the residual as well as proving the observer stability will be presented.

First, one must make the assumption that the output signal is Lipshitz which means that there exists some constant M such that $\left|\frac{dy}{dt}(\gamma)\right| < M$.

Then by setting $\xi(t)=\int_0^t e(t)$,it gives

$$\begin{cases} \dot{\xi}=e \\ \dot{e}=\left(w^T-\hat{w}^T\right)v(t)-k_0e-k_1\xi-K_1\frac{dy}{dt}(\gamma)\tau-K_2\frac{dy}{dt}(\gamma)\tau t \end{cases} \tag{6.15}$$

which can be rewritten as

$$\begin{pmatrix} \dot{\xi} \\ \dot{e} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ k_1 & k_0 \end{pmatrix} \begin{pmatrix} \xi \\ e \end{pmatrix} + \begin{pmatrix} 0 \\ \varphi \end{pmatrix} \tag{6.16}$$

where $\varphi = \left(w^T - \hat{w}^T\right) v(t) - K_0 \frac{dy}{dt}(\gamma)\tau - \frac{1}{2} K_1 \frac{dy}{dt}(\gamma)\tau t$ is a polynomial in (t, τ)

Set $k_0 = -2\theta$, $k_1 = -\theta^2$ where $\theta > 0$ is some constant and define Δ_θ as:

$$\Delta_\theta = \begin{pmatrix} \frac{1}{\theta} & 0 \\ 0 & \frac{1}{\theta^2} \end{pmatrix} \tag{6.17}$$

then, let define ϵ as

$$\epsilon = \Delta\varepsilon \tag{6.18}$$

Then, system described by Equation (6.15) is transformed into:

$$\begin{aligned} \dot{\epsilon} &= \begin{pmatrix} \dot{\bar{\xi}} \\ \dot{\bar{e}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\theta} & 0 \\ 0 & \frac{1}{\theta^2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -\theta^2 & -2\theta \end{pmatrix} \begin{pmatrix} \theta & 0 \\ 0 & \theta^2 \end{pmatrix} \begin{pmatrix} \bar{\xi} \\ \bar{e} \end{pmatrix} \\ &\quad + \begin{pmatrix} \frac{1}{\theta} & 0 \\ 0 & \frac{1}{\theta^2} \end{pmatrix} \begin{pmatrix} 0 \\ \varphi \end{pmatrix} \\ &= \theta \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} \bar{\xi} \\ \bar{e} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \\ &= \theta F \begin{pmatrix} \bar{\xi} \\ \bar{e} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{\theta^2}\varphi \end{pmatrix} = \theta F \epsilon + \begin{pmatrix} 0 \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \end{aligned} \tag{6.19}$$

In the new system, the eigenvalue is $-\theta$ of order 2. which means that by choosing $\theta > 0$ the free systems will be stable. But as φ is a polynomial in (t, τ), the stability analysis needs to be studied by including φ.

A key feature of the proposed observer though, is as its can be noticed, φ being multiplied by $\frac{1}{\theta^2}$. This means that considering a high gain observer case, the bigger θ is chosen, the more it will bring the free system toward a more stable state as φ is lowered to power two.

Now consider the Lyapunov function $V = \epsilon^T P \epsilon$, then since F is stable there exists a positive definite symmetric matrix P such that

$$PF + F^T P = -I \tag{6.20}$$

Derivating Equation (6.20) with respect to time gives

$$\dot{V} = 2\epsilon^T P \dot{\epsilon} \tag{6.21}$$

then replacing ϵ by its expression as in Equation (6.14) gives

$$\begin{aligned} \dot{V} &= 2\epsilon^T P \left(\theta F \epsilon + \begin{pmatrix} 0 \\ \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \right) \\ &= 2\theta\epsilon^T P F \epsilon + 2\epsilon^T P \begin{pmatrix} 0 \\ \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \\ &= -\theta\epsilon^T\epsilon + 2\epsilon^T P \begin{pmatrix} 0 \\ \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \end{aligned} \tag{6.22}$$

Finally

$$\begin{aligned} \dot{V} &= -\theta\epsilon^T\epsilon + 2\epsilon^T P \begin{pmatrix} 0 \\ \\ \frac{1}{\theta^2}\varphi \end{pmatrix} \\ &= -\theta \left\|\epsilon\right\|^2 + 2\frac{1}{\theta^2} \left\|\epsilon\right\| \left\|P\right\| \left|\varphi\right| \end{aligned} \tag{6.23}$$

So the condition for the system to be stable, i.e. $\dot{V} < 0$, is that θ is chosen such that

$$\left\|\epsilon\right\| > \frac{2}{\theta^3} \left\|P\right\| \left|\varphi\right| when \left\|\epsilon\right\| \neq 0. \tag{6.24}$$

In plain words, this result means that the amplitude contribution in the residual brought by the initial conditions and the delay multiplied by $\frac{2}{\theta^3}$, knowing that in the considered case θ is chosen to be of high value, is less than the error value.

With the proposed observer, it is proved that the system stability is met if the observer gain value insures that the condition (6.24) set here is fulfilled. more importantly, as the purpose of this work is to deal with delay in the system, it was showed that using this observer, its effect can effectively be reduced.

Finally, it was showed that with this approach, it can be managed to get the observer design working on the direction that works best to increase the observer performances. Indeed, the high gain value desirable to improve the systems stability, also allows to attenuate the undesirable effects of delay best.

6.4 Conclusion

In this chapter a technique to attenuate known and unknown time delay impact in nonlinear systems has been proposed. It relays on considering all the delays occurring in the system at one reference point, chosen here to be the measurement moment. Then using the output observer with output injection, it was showed that for known delays, the solution is straight forward as the delay effect can be removed in the residual and thus the observer performance analysis can be done without dealing with delay. For unknown time delays, the same solution is proposed in which the observer performances analysis uses mean value theorem and relies on choosing a Lyapunov function to prove the efficiency of the output observer design, while establishing the stability proprieties of the system.

Chapter 7

Conclusions and Future work

7.1 Conclusions

The main motivation of this research work has been to explore model based fault detection techniques and strategies for linear and nonlinear dynamical systems. In effect, the thesis focus has been on developing novel observers design techniques for fault detection purposes. For this purpose, several observer design methodologies has been proposed whereby each of them has specific features and properties.

First, a proportional integral observer has been proposed which allows to attenuate disturbances with respect to faults occurring in the systems. This has the advantage of reducing false alarms triggering. The proposed proportional integral observer based fault detection scheme relies on using an augmented state vector that is the result of integrating the original system output. Then, the disturbance effect has been reduced using the disturbances input gain matrix directly in the observer gain, and which, at the same time, allows to have more freedom in the observer gain optimisation process. The optimisation process uses a frequency based criterion that has the property of increasing the faults sensitivity while reducing that of the distrurbance.

Next, an output observer design has been proposed for the purpose of reducing computation time as well as for the ease of implementation. The output observer solution relies on using an input/output representation of the system. This representation is shown to be more adapted for output estimation than using state space representation for dynamical systems, since only an output estimation is needed for fault detection. The considered output observer fault detection scheme design method have been detailed and its performances proved for linear and general nonlinear systems under disturbances. Indeed output observer output injection technique feature helped simplifying the study process of nonlinear systems by eliminating the nonlinearities in the residual. Finally, output observer technique has been proposed in a solution to deal with time delay impact in systems. In the proposed solution, the observer performances have been proved using Lyapunov function which showed the effectiveness of the proposed method. The novelty this work consist on developing observers designs for fault detection rather than adapting existing estimation techniques to the fault detection case.

7.2 Future work

The main task that remains to be done regarding this thesis work is concerned with the practical implementation of the results obtained. Indeed, as the focus of the thesis was on developing the theory and validating results with simulation only, the focus next is to implement it in the modelled three wheeled robot. In order to do so, the research work carried out during the PhD studies needs to be adapted to discrete time systems.

Regarding the output observer development, another focus should be to design reduced order output observers so as to further reduce the observer estimation computation time. Also, as the observer have been developed for a general class of nonlinear systems, its adaptation to specific classes such state affine and control affine systems may result in performances enhancement. Another topic of interest can be output observer design for hybrid dynamical systems since these types of systems are omnipresent in many applications.

Finally, only time delay was considered among possible challenges in interconnected systems. Thus, more challenges such as packet dropout in networked systems, limited communication bandwidth and non-uniform sampling, need to be addressed.

References

[1] Y. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annual reviews in control*, vol. 32, no. 2, pp. 229-252, 2008.

[2] P.M. Frank and D. Xianchun, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *Journal of process control*, vol. 7, no. 6,. pp. 403-424, 1997.

[3] R. Isermann and B. Peter, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control engineering practice*, vol. 5, no.5, pp. 709-719, 1997.

[4] R. Isermann, "Supervision, fault-detection and fault-diagnosis methods-an introduction," *Control engineering practice*, vol. 5, no. 5, pp. 639-652, 1997.

[5] R. Kalman, "On the general theory of control systems," *IRE Transactions on Automatic Control*, vol. 4, no. 3, pp. 110-110, 1959.

[6] R. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics Series A Control*, vol. 1, no. 2, pp. 152-192, 1963.

[7] D.G. Luenberg, *Introduction to Dynamic Systems: Theory, Models and Applications*, New York: Wiley, 1979.

[8] T. Kailath, *Linear systems*, New Jersey: Prentice-Hall, 1980.

[9] D.M. Himmelblau, *Fault detection and diagnosis in chemical and petrochemical processes*, Amsterdam: Elsevier, 1978.

[10] W. Borutzky, *Bond Graph Model-based Fault Diagnosis of Hybrid Systems*, New York: Springer, 2015.

[11] R. Isermann, "Model-based fault-detection and diagnosis–status and applications," *Annual Reviews in control*, vol. 29, no. 1, pp. 71-85, 2005.

[12] X. Zhang, M.P. Marios and T. Parisini, "Fault diagnosis of a class of nonlinear uncertain systems with Lipschitz nonlinearities using adaptive estimation," *Automatica*, vol. 6, no. 2, pp. 290-299, 2010.

[13] X. Zhang, T. Parisini and M. M. Polycarpou, "Adaptive fault-tolerant control of nonlinear uncertain systems: an information-based diagnostic approach," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1259-1274, 2004.

[14] X, Zhang, M.M. Polycarpou and T. Parisini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 576-593, 2002.

[15] S. Bhatnagar, V. Rajagopalan and A. Ray, "Incipient fault detection in mechanical power transmission systems," in American contol conference, pp. 472-477, 2005.

[16] R. Ferrari, "Distributed fault detection and isolation of large-scale nonlinear systems: an adaptive approximation approach," Ph.D. dissertation, 2009.

[17] A. Saxena, K. Goebel, D.Simon and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in International Conference on Prognostics and Health Management, 2008.

[18] R. Isermann, "Process fault detection based on modeling and estimation methods-a survey," *Automatica*, vol. 20, no. 4, pp. 387-404, 1984.

[19] H.P. Huang, C.C. Li, and J.C. Jeng, "Multiple multiplicative fault diagnosis for dynamic processes via parameter similarity measures," *Industrial & engineering chemistry research*, vol. 46, no.13, pp. 4517-4530, 2007.

[20] S. Ding, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*, London: Springer, 2013.

[21] P.M.Frank and D. Xianchun, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *Journal of process control*, vol. 7, no 6, pp. 403-424, 1997.

[22] I. Hwang, S. Kim, Y. Kim and C. E. Seah, "A Survey of Fault Detection, Isolation, and Reconfiguration Methods," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636-653, 2010.

[23] R.J. Patton and J. Chen, "A review of parity space approaches to fault diagnosis," *Journal of Guidance Control Dynamics*, vol. 17, pp. 278-285, 1994.

[24] Z. Gao, C. Cecati and S. X. Ding, "A Survey of Fault Diagnosis and Fault-Tolerant Techniques-Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757-3767, 2015.

[25] R. Isermann, "Parameter adaptive control algorithms-A tutorial," *Automatica*, vol. 18,.no. 5, pp. 513-528, 1982.

[26] X. Dai, "Observer-based parameter estimation and fault detection", Ph.D. dissertation, 2008.

[27] D. Luenberger, "Observers for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 190-197, 1966.

[28] R.J. Patton, P.M. Frank and R.N. Clarke, *Fault diagnosis in dynamic systems: theory and application*, New Jersey: Prentice-Hall, 1989.

[29] V. Puig, A. Stancu, T. Escobet, F. Nejjari, J. Quevedo and R.J. Patton, "Passive robust fault detection using interval observers: Application to the DAMADICS benchmark problem," *Control engineering practice*, vol. 14, no. 6, pp. 621-633, 2006.

[30] 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, pp. 85-105, 2007.

[31] M. Hou and R. J. Patton, "An LMI approach to H-/H∞ fault detection observers," in UKACC International Conference, vol. 1, pp. 305-310, 1996.

[32] H. Sneider and P. M. Frank, "Observer-based supervision and fault detection in robots

using nonlinear and fuzzy logic residual evaluation," *IEEE Transactions on Control Systems Technology*, vol. 4, no. 3, pp. 274-282, 1996.

[33] K. Zhang, B. Jiang, and V. Cocquempot, "Adaptive observer-based fast fault estimation," *International Journal of Control, Automation and Systems*, vol. 6, no. 3, pp. 320-326, 2008.

[34] Z. Shi, F. Gu, B. Lennox and A.D. Ball, "The development of an adaptive threshold for model-based fault detection of a nonlinear electro-hydraulic system," *Control Engineering Practice*, vol. 13, no. 11, pp. 1357-1367, 2005.

[35] M. Markou and S. Singh. "Novelty detection: a review-part 2:: neural network based approaches," *Signal processing*, vol. 83, no. 12, pp. 2499-2521, 2003.

[36] R.J. Patton and J. Chen, "Robust fault detection using eigenstructure assignment: a tutorial consideration and some new results," in Conference on Decision and Control, vol..3, pp. 2242-2247, 1991.

[37] 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.

[38] R.J. Patton and J. Chen, "Observer-based fault detection and isolation: robustness and applications," *Control Engineering Practice*, vol. 5, no. 5, pp. 671-682, 1997.

[39] C. Edwards, S.K. Spurgeon and R.J. Patton, "Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, no. 4, pp. 541-553, 2000.

[40] P.M. Frank. and X. Ding, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *Journal of Process Control*, vol. 7, pp. 403-424, 1997.

[41] E.A. Garcia, and P.M. Frank, "Deterministic nonlinear observer-based approaches to fault diagnosis: a survey," *Control Engineering Practice*, vol. 5,.no. 5, pp. 663-670, 1997.

[42] P. Zhang and S.X. Ding, "An integrated trade-off design of observer based fault detection systems," *Automatica*, vol. 44, pp. 1886-1894, 2008.

[43] H. Khalil, *Nonlinear Control*, London: Pearson Education, 2014.

[44] R. Hermann and A.J. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728-740, 1977.

[45] D. Boutat and D. Liu, "Observer canonical form for a class of multi-outputs nonlinear systems," in 33rd Chinese Control Conference, pp. 2547-2552, 2014.

[46] A.J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *System & Control Letters*, vol. 3, no. 1, pp. 47-52, 1983.

[47] K. Busawon and M. Djemaï, "Algorithms for Transformation into the Extended Jordan Controllable and observable forms," in IEEE Conference on Decision and Control/Chinese Control Conference, pp. 1764-1769, 2009.

[48] A.J. Krener and A. Isidori, "Linearization by Output Injection and Nonlinear Observers," *Systems & Control Letters*, vol. 3, pp. 47-52, 1983.

[49] A.J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *Journal on Control and Optimization*, vol. 23, no. 2, pp. 197-216, 1985.

[50] S.V. Drakunov, "Sliding-mode observers based on equivalent control method," in Conference on Decision and Control, vol. 2, pp. 2368-2369, 1992.

[51] S. Drakunov and V. Utkin, "Sliding-mode observers. Tutorial", in Conference on Decision and Control, vol. 4, pp. 3376-3378, 1995.

[52] W. Perruquetti, and J.P. Barbot, *Sliding mode control in engineering*, New York: Marcel Dekker, 2002.

[53] P. Kabore and K. Busawon,"On the design of integral and proportional integral observers," in American Control Conference, vol. 6, pp. 3725-3729, 2000.

[54] P. Kabore and K. Busawon, " Disturbance attenuation using proportional integral observers," *International Journal of Control*, vol. 74, no. 6, pp. 618-627, 2001.

[55] D. Koenig and S. Mammar, "Design of proportional-integral observer for unknown input descriptor systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 12, pp. 2057-2062, 2002.

[56] C. Hua and X. Guan, "Synchronization of chaotic systems based on PI observer design," *Physics Letters*, vol. 334, no. 5-6, pp. 382-389, 2005.

[57] X. Dai, Z. Gao, T. Breikin and H. Wang, "Disturbance Attenuation in Fault Detection of Gas Turbine Engines: A Discrete Robust Observer Design," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 39, no. 2, pp. 234-239, 2009.

[58] G. Liu and R. J. Patton, *Eigenstructure Assignment for Control System Design*, Wiley, 1998.

[59] G. Roppenecker, "On parametric state feedback design," *International Journal of Control*, vol. 43, no 3, pp. 793-804, 1986.

[60] M. Fahmy, and J. O'Reilly, "Eigenstructure assignment in linear multivariable systems–A parametric solution," *IEEE Transactions on Automatic Control*, vol 28, no 10, pp. 990-994, 1983.

[61] K. Busawon, "Output observer design for linear systems: Application to filtering and fault detection," in European Control Conference, pp.318-323, 2014.

[62] K. Ogata, *Modern Control Engineering*, 4th ed, New Jersey: Prentice Hall, 2001.

[63] J.P. Gauthier and I.A.K.Kupka, "Observability and observers for nonlinear systems," *SIAM Journal of Control Optimization*, vol. 32, no. 4, pp. 975-994, 1994.

[64] X. Dai, Z. Gao, T. Breikin and H. Wang, "Disturbance Attenuation in Fault Detection of Gas Turbine Engines: A Discrete Robust Observer Design," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 39, no. 2, pp. 234-239, 2009.