Advanced fault diagnosis of dynamical systems using nonsmooth optimization

Jingwen Yang

Ecole Doctorale IAEM Lorraine
DFD Automatique et Production Automatisée

THÈSE

présentée pour l'obtention du grade de

Docteur de l'Université de Lorraine

(Spécialité Automatique, Traitement du Signal et Génie Informatique)

par

Jingwen YANG

Diagnostic de défauts des systèmes dynamiques par optimisation non lisse

Advanced fault diagnosis of dynamical systems using nonsmooth optimization

Soutenue publiquement le 16 Février 2015 devant la commission d'examen :

Rapporteurs :	Ping ZHANG	Professeure à l'Université de Kaiserslautern
	Dominique KNITTEL	Professeur à l'Université de Strasbourg
Examinateurs :	Pierre APKARIAN	Professeur à l'Université Paul Sabatier
	Guillaume DUCARD	Maitre de conférences à l'Université de Nice Sophia Antipolis
Directeurs de thèse :	Dominique SAUTER	Professeur à l'Université de Lorraine
	Frédéric HAMELIN	Professeur à l'Université de Lorraine

Centre de Recherche en Automatique de Nancy
CRAN – CNRS UMR 7039

dedicated to my family

Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisors Professor Dominique Sauter and Professor Frédéric Hamelin for their guidance and support. Dominique has always proposed some innovative ideas and explained to me so clearly. Frédéric can always understand my questions fast and propose some inspiring views on the questions that I was surprised that how can he obtain these constructive suggestions. I appreciate to have a chance to work with my supervisors, and the meetings with them always promote the process of the thesis research.

I also would like to thank Professor Pierre Apkarian and Dr Guillaume Ducard. The discussion with Professor Pierre about the application of nonsmooth optimization technology on FDI field is really interesting, which is the cornerstone of my thesis. Dr Guillaume Ducard introduced me to the CRAN (Centre de Recherche en Automatique de Nancy), which gives me the chance to do my research here.

My sincere gratitude also goes to Professor Ping Zhang and Professor Dominique Knittel for reading and reviewing my thesis. Professor Knittel gave some interesting comments in the view of optimization technology and Professor Zhang supplied some meaningful points in the view of FDI.

I am grateful to thank the members of CRAN. Particularly, I deeply thank to our secretary, Ms Sabine Huraux, for the kind help of many administrative tasks.

I would also deeply thank my fellow lab mates in CRAN: Tushar Jain, Mohammad Amine, Ahmed Faghraoui, Shaik Chitraganti, Fransisco Ronay, Manal Dakil, Ghassane Kabadi, Fengwei Chen, Lukasz Dziekan, Jean Carlo Salazar, Demaino Rotondo, Boyi Ni, Quan Liu, Marc-Abel Bisch, Maxim Doublet, Krishnan Srivisarengan, Tejaszini Darure for the interesting discussions and the fun. I also would like to extend my appreciation to my chinese friends in nancy: Fengwei Cheng, Xuanxuan Xiao, Yong Xu, Yong Hu, Xiaoyan Gu, Hengcai Tang, Jun Liu, Shanbin Li, Kun Liu, Yongxin Liao, Yuanzhao Chen, Xin Hu, Ning Yu, Yingfei Cai for the parties during these years. I am also grateful to all my friends in China: Xiangwen Yu, Fan Pang, Lei zuo, Shuai Tang ..., for the supports and encouragement.

I also give my thanks to China Scholarship Council (CSC) for the financial supports for my PhD.

I heartily thank Professor Zhiqiang Zheng for supporting me in these years with his kind encouragement and his kindly help in the application of China Scholarship Council (CSC) scholarship.

Last but not the lest, I appreciate my family for the selflessly supporting with their constant understanding and encouragement, which give me motivations to keep moving forward.

Abstract

This dissertation considers the application of nonsmooth optimization approach with mixed performance index on several FDI problems.

Firstly, to overcome the drawback of classical methods, like LMI, a nonsmooth optimization approach is proposed to solve a multiobjective fault detection problem in the worst case, where $\|\cdot\|_\infty$ is used to evaluate the maximum effects of the disturbances on the residual, while $\|\cdot\|_-$ is utilized to measure the minimum influences of the fault on the residual. An additional constraint of fast transients of residual responses could be added into the design, which could be solved by nonsmooth optimization approach. A framework of designing a unique observer gain and a unique residual weighting matrix is proposed for a system with multiple models. For the case that the exact model is unknown, a new framework of robust fault detection filter and an unchanged threshold are proposed, which are independent on the exact model of the system. Both designs are transformed as a min-max formulation, which could be solved by the proposed nonsmooh optimization method.

Secondly, a method is proposed to design an integrated fault detection observer for general case (unknown $\mathcal{L}_2$ bounded faults and disturbances) and specific case (some specific faults) with specifications in frequency and time domain, where the design for general case is evaluated with mixed specifications H_-/H_∞. Different from the traditional frequency design method, lower and upper bound envelopes in time domain are proposed to adjust the transient of the residual for the specific faults directly. Using the lower bound envelope, the first application designs a fast fault detection observer for the specific faults with a guaranteed ability of fault detection in the worst case. By contrast, to decrease false alarms when fault disappearing, a constraint of an upper bound envelope is added into the design.

Thirdly, in some cases, the effects of fault are masked by the operations of controller, which is difficult to be detected with the passive fault diagnosis method. A new framework of active diagnosis with auxiliary signal is proposed. Different from the classical specifications used in the literature, a criterion of peak amplitude is proposed to evaluate the worst effects from the auxiliary signal on the system. The effects of auxiliary signal on the outputs and control signals are considered into the design. The design method is firstly shown with a case of two models, which could be considered as model detection. Furthermore, for the multiple models case (model isolation), the framework is improved with a decision logic to discriminate all the models.

The techniques developed in this dissertation are well illustrated using either academic or practical examples (A vehicle lateral dynamics system and DC motor control system DR300) and the results show the effectiveness of the proposed techniques.

Résumé

Cette thèse consiste à appliquer des méthodes d'optimisation non lisse pour la résolution de différents problèmes de diagnostic impliquant des indices de performances multi-critères (fréquentiels et temporels).

Dans un premier temps, afin de surmonter les inconvénients des méthodes traditionnelles basées sur la résolution d'inégalités matricielles (LMI), une approche de l'optimisation non lisse est proposée pour résoudre le problème de détection de défauts multi-objectif dans le pire des cas, dans lequel la norme $\|\cdot\|_\infty$ est utilisée pour évaluer les effets maximaux que produisent des perturbations sur le résidu, alors que l'indice $\|\cdot\|_-$ est utilisé pour évaluer l'influence minimale du défaut sur le résidu. Une contrainte supplémentaire relative à la rapidité de la réponse résiduelle est également ajoutée dans la conception et le problème d'optimisation est résolu grâce à une approche non lisse. Avec comme objectif de concevoir un générateur de résidu présentant un gain et une pondération de la sortie uniques, une démarche est ensuite proposée pour le diagnostic de système à commutation multi-modèles. Dans le cas où le modèle exact est inconnu, une construction d'un filtre robuste de détection de défaut est également étudiée.

Dans la seconde partie de la thèse, une méthode est proposée pour concevoir un observateur intégré de détection de défauts dans le cas général (défauts et perturbations $\mathcal{L}_2$ borné et inconnu) et pour un cas particulier (défaut spécifique). Dans le cas général, la synthèse est réalisée dans les domaines temporels et fréquentiels, et la conception est réalisée en considérant des spécifications H_-/H_∞ mixtes. Dans le domaine temporel, une enveloppe des bornes inférieure et supérieure du résidu est proposée pour ajuster le transitoire du résidu en réponse à des défauts spécifiques; l'enveloppe inférieure sert à régler la rapidité de la réponse alors que l'enveloppe supérieure permet d'ajuster le taux de fausses alarmes.

Dans certaines situations, les effets de défauts peuvent être dissimulés par les actions du contrôleur, qui réduit la capacité du système de diagnostic à détecter les défauts par un moyen dit passif. Nous proposons dans la dernière partie de la thèse d'étudier une approche active du diagnostic, consistant à injecter des extra-signaux sur les commandes du système de manière à révéler au mieux la présence de défauts. A la différence des spécifications traditionnelles relevées dans la littérature, un critère d'amplitude de crête est proposé pour évaluer les effets du signal auxiliaire sur le système. Les effets du signal auxiliaire sur les sorties et les signaux de contrôle sont introduits dans la conception du signal auxiliaire et du post-filtre. Tout d'abord, la conception est présentée dans le cas de deux modèles caractérisant respectivement le fonctionnement normale et le fonctionnement défaillant du système. En outre, dans le cas où plusieurs défauts peuvent cohabiter, une démarche logique proposée pour l'isolation du défaut.

Les techniques développées dans cette thèse sont bien illustrées par des exemples académiques et les résultats obtenus montrent l'efficacité des techniques proposées.

Contents

List of Figures

List of Tables

Notations and Acronyms

Matrices and vectors

$\mathbb{R}^n$	n dimensional real Euclidean space
A^T	The transpose of a matrix A
A^{-1}	The inverse of the matrix A
A^+	The Moore-Penrose pseudo inverse of the matrix A
$\| \, . \, \|$	The standard vector norm inR^n
$\| \, A \, \|$	An induced norm of the matrix A
$\| \, A \, \|_2$	H_2 norm of A
$\| \, A \, \|_\infty$	H_∞ norm of A
$\| \, A \, \|_-$	H_- index of A
$\|A\|_{\text{peak}}$	The peak index of A
$\|A\|_{\text{rms}}$	The RMS norm of the matrix A
$\Psi \left(x \left(t \right) \right)$	Natural response with state $x(t)$
$\mathfrak{L}^{-1}$	The inverse Laplace transform
$\mathfrak{L}$	The Laplace transform

Acronyms

FD	Fault Detection
FDI	Fault Detection and Isolation
LTI	Linear Time Invariant
SISO	Single Input Single Output
MIMO	Multiple Inputs Multiple Outputs
LMI	Linear Matrix Inequality
ILMI	Iterative Linear Matrix Inequality

GSVD	Generalized Singular Vaule Decomposion
RMS	Root Mean Square
RFDF	Robust Fault Detection Filter
DC	Direct Current

1. Introduction

1.1. Motivation

The research interests of this dissertation are focused on improving the effects of fault diagnosis for dynamical system with the aid of performance indices in time and frequency domain. The motivating objectives of this research are explained in the following lines:

Why fault detection?

With the development of more and more integrated and complex systems during the past several decades, the critical issues of system, like, higher performance, cost efficiency, product quality, dependability and reliability, are received much attention. In order to improve the above objectives of system, one direct way is developing the reliability of individual component and unit like controllers, sensors, actuators and plants. However, it is unrealistic to guarantee that all the components work in the perfect states when the conditions or environments change. Faults in the process arise due to loss of efficiency of actuators or sensor malfunctions, etc, which may move the states of the process far away from the designed operating points. As a result, the performance of process is degraded. The degradation of process will in turn affects the cost of operation and the quality of products. In a worst case, if faults are not considered in the early stages, the consequences will be not only a significant performance degradation, but also serious damage of system, catastrophic disaster or even human casualties, such as in aircrafts, nuclear systems.

Within the different objectives to deal with faults, three essential tasks in fault diagnosis (FDI) should be considered [27, 35]:

- Fault detection: to make a decision whether there is the occurrence of faults in the process, which lead to undesired or intolerable behaviors of whole system;
- Fault isolation: to determine the location of fault;
- Fault identification: to estimate the size, type and cause of fault.

These three tasks are comparably important, and fault detection is the basis of the other two tasks. Without the fault detection, fault isolation and fault identification are hard to achieve directly.

Why model based fault detection?

A traditional approach is to install an extra hardware in parallel with the process, called as hardware redundancy. A fault will be detected if the output of process is different from the

output of the redundancy. Although this method has the advantages of high reliability and direct fault isolation, it has at least three problems: require additional cost, space and weight to install the extra hardware. What's more, the additional components increase the complexity of system, which in turn may introduce some additional problems.

One intuitive method is replacing the hardware redundancy by an analytical model (mathematical model), named as model based fault detection [110]. By contrast, an alternative idea is to detect faults without using models (neither hardware redundancy nor analytical model), which called as model-free fault detection [108, 110, 109]. Model free fault detection needs a great amount of system measurements as training data to develop fault detection logic. When the model information is not available, data driven fault detection can be effective. It has relatively high computational complexity and it is suitable to be applied to large scale industrial complex systems [29]. By contrast, model based fault detection must have a mathematical system model to represent the plant, and the types of models could be autoregressive moving average (ARMA) model, state space model, and transfer function. When the information of system model is available, model based approach is generally much more accurate and faster than model free method.

Originated in the early 70's, a large amount of work has been done about the model based fault diagnosis since then. The model based method is proved to efficiently detect faults with a large number of real-time applications:

- Aircrafts: aircraft control system, navigation system, and engines;
- Nuclear power plants;
- Emission control systems;
- Chemical plants;
- Industrial robots;
- Electrical motors.

Considering technological and economic demands on the one hand, efficient fault detection and on-line implementation capability on the other hand, model based method is a powerful tool to solve FDI problems [35].

Why using nonsmooth optimization method?

In the recent years, specialized nonsmooth optimization techniques have been used to solve a variety of difficult problems in controller design [8, 10, 9, 17, 7]. These techniques aim at solving challenging control problems, which are often classified as Non-deterministic Polynomial-time hard (NP-hard). Nonsmooth optimization method avoids using Lyapunov variables, whose number grows quadratically with the plant state dimension. Therefore, due to the execution time, the nonsmooth techniques are much faster than linear matrix inequalities (LMI) based methods, especially for large scale systems. In addition, it can deal with problems where controller structure/architecture is important.

- Plant has large size;
- Specifications are not LMI representable;
- LMI methods are conservative;

- Etc

With the advantages of nonsmooth optimization techniques, it is interesting to apply this technique to a number of challenging problems in fault diagnosis.

1.2. Objectives

Inspired by the motivations in the last section, this thesis is primarily concerned with the application of nonsmooth optimization on some problems in FDI. According to the requirements of fault detection in reality, the desirable features of a fault detection system are given below:

1. Early detection of faults;
2. High sensitivity to faults, under the influences of disturbances or model uncertainties;
3. As few chances as possible of both false alarm and missing alarm;
4. Should not take much time for on-line computation.

Based on these four features of a good fault detection system, three different objectives of this thesis are given as follows.

Design fault detection filter with nonsmooth optimization approach

Model based FDI diagnosis systems estimate the outputs of the plant system, and compare the estimations with the actual measurements of plant system to generate the well known residual signal. The residuals in fault free situations are very close to zero while clearly deviates from zero in the presence of faults. A second objective of fault diagnosis systems is the ability to discriminate between all the possible faults, which is explained as the use of the term isolation. However, because of the unknown inputs (process disturbances, model uncertainties and measurement noises), the residual signal will not be zero even if there is no fault in the process. Both fault detection and isolation should be robust against disturbances and model uncertainties. Since full decoupling of the disturbance on residual is hard to achieve, a decision making solution requires to maximize a performance index allowing the generation of a residual which shows maximum sensitivity in the presence of faults with maximum robustness to disturbances.

- Revising the traditional fault detection filter design in the framework of nonsmooth optimization

  Different criteria are used to evaluate the effects of disturbances and faults on residual. Same as in the problem of control, criteria of H_2 norm and H_∞ norm are good candidates for measuring up the disturbance rejection capability of a fault detection system. For the sensitivity of residual to faults, H_2 norm and H_∞ norm are also used. Specifically, H_- index (is also called as H_- norm), which describes the minimum influence of faults on the residual signal, seems to be an appropriate performance index to describe the sensitivity of residual to faults. With above defined criteria, fault detection problems could be formulated as multiple objective optimization problems by maximizing the effects of faults on residual and minimizing the sensitivity to disturbances with different evaluation specifications, for example, H_2/H_2 problem, H_2/H_∞ problem, H_∞/H_∞ problem, H_-/H_2 problem and H_-/H_∞.

Design fault detection filter in mixed time and frequency domain

To achieve a successful fault detection based on the generated residual, further efforts are needed to extract the information of faults from the residual with disturbances: residual evaluation and threshold setting. In the process, the evaluated residual (with evaluation function) is compared with the threshold, then a decision on the occurrence of a fault could be got if the evaluated residual exceeds the threshold. Threshold selection is a very important task in fault diagnosis. If the threshold is selected too high, the residual with some kinds of faults may not exceed the threshold, which will result into a missed detection. On the other hand, if the threshold is selected too low, a false alarm may appear due to the effects of unknown input. Since the threshold is used to filter the perturbations of disturbances on the decision making, the threshold setting is dependent on the disturbances.

As shown in the desirable features of fault detection system, the fault should be detected as soon as possible, or even predicted before fault occurs. To decrease fault detection delay, the dynamics of residual when fault appears and threshold setting should be considered together. A lower threshold will give a faster reaction of residual for the alarm. However, the low threshold has a risk of false alarm. Additional, the design of low threshold may decrease the ability of fast transients of residual.

- Design observer based fault detection filter in time and frequency domain.

  Some specifications to evaluate the transients of responses in frequency domain could be used to improve the dynamics of residual, such as eigenvalues [27, 115, 119]. However, these kinds of performance index are not available for some cases and they cannot be represented in time domain directly. An alternative method is considering the transients of response in time domain directly. To guarantee the design's abilities of fast fault detection, high fault sensitivity and high disturbances robustness of the design, mixed specifications in time and frequency domain should be considered simultaneously to design the fault detection filter.

Active diagnosis

The task of fault diagnosis is difficult in hybrid systems with combined continuous and discrete behaviors, because the changing mode make diagnosability hard to achieve. Including additional sensors can improve diagnosability, whereas, this method is not always feasible. An alternative strategy is active fault diagnosis, where the diagnosis results could be improved by executing or blocking controllable events.

- Active diagnosis design in time and frequency domain.

  In the active fault diagnosis method, an auxiliary signal should be exerted to excite the system, which should be large enough so that the fault could be observable from the input and output of the system. Simultaneously, the designed auxiliary signal should minimally disturb system operation. It is interesting to propose some new performance indices in time and frequency domain to compensate the disadvantages of traditional specifications in the design of auxiliary signal.

1.3. Contributions

In this section, the major contributions are summarized as follows:

1. Considering the fault detection in the "worst case", which means that minimum influences of faults and maximum influences of disturbances on the residual, an optimization framework of H_-/H_∞ is formulated. The formulated optimization problem could be solved by the nonsmooth optimization method. The criterion H_∞ is used to evaluate the maximum robustness to the disturbances, while the minimum sensitivity to faults is represented with the specification H_-. Different from the classical methods, like LMI and gradient method, the nonsmooth optimization method shows faster calculation with the similar suboptimal results. The criterion of eigenvalues is also added into the optimization to improve the dynamics of residual. With the aid of min-max formulation, a unique observer and residual weighting matrix are designed for a system with multiple models. Since the parameters A, B, C and D of system in observer should be known, the switching signal of the multiple models system should be introduced into the observer to switch the parameters in the observer and the threshold. Then, a robust fault detection filter is formalized to detect faults for a system with multiple models if the exact information of models are unknown. The "robustness" of this filter means that the filter is not only robust to the disturbances, but also robust to the model of system. A unique threshold is proposed to work with the robust filter to detect fault for this kind of system.

2. In order to improve the transients of residual responses from faults, traditional method always utilizes suitable frequency domain performance indices to meet the corresponding constraints in time domain. However, this kind of method is inaccurate and indirect. To overcome the shortages of classical methods, a direct approach to adjust the transients of residual from some specific faults is introduced, where the specific faults could be steps, ramps or other typical interesting signals. In this method, an envelope (a lower bound envelope and an upper bound envelope) of residual should be given in advance, which will constrain the upper and lower bound of residual in time domain. With the aid of an iterative algorithm to minimize fault detection delay, a fast fault detection design could be realized with the mixed criteria H_-/H_∞ for unknown faults and disturbances in frequency domain and the specifications of lower bound envelope for specific faults in time domain. Furthermore, the upper bound envelope is added into the optimization to decrease false alarms when fault disappears.

3. With the aid of a framework of fault detection, a new framework of active fault diagnosis is proposed to design auxiliary signal. For the two models discrimination case (called model detection), a discrimination condition is introduced to distinguish these two models. Different from the ideas in related literature [21], the worst "disturbances" of auxiliary signal on system is explained as the peak amplitude of the responses from auxiliary signal on the control signal and output of system. On the other hand, for the multiple models case (also could be called as model isolation), a framework with a bank of reference models is introduced. The generated multiple residuals can be used to discriminate the multiple models with a typical logic.

1.4. Outline

The outline of this thesis is displayed in the following overview.

Chapter 2: **Review and algorithms** - presents a review of optimization in model-based fault detection. Typically, a great deal of attention is paid on the different performance indexes used in literature. For multiobjective optimization formulation, two different types of multiobjective optimization are introduced: Min-max method and Lexicographic method. Finally, two algorithms are briefly described to solve the proposed multiobjective optimization problem.

Chapter 3: H_-/H_∞ **fault detection filter design via nonsmooth optimization approach** - The design in the worst case for single model case is revised for nonsmooth optimization method, and a comparison between the nonsomooth optimization method and other classical methods is given. Then, two frameworks for multiple models case are introduced. A theoretical example is given to show that the proposed method could give a result as good as other classical methods, while a practical example about vehicle lateral dynamics system is provided to show the effects of the designs with two proposed frameworks.

Chapter 4: **FDI observer design using time and frequency domain specifications** - An integrated design for unknown faults and specific faults is introduced. To minimize the fault detection delay for some specific faults and decrease false alarms when the fault disappears, specifications with lower and upper bound envelope in time domain are proposed. In addition, the ability of fault detection for unknown faults could be evaluated with the mixed criterion H_-/H_∞ in frequency domain. Finally, the usefulness of the proposed methodologies is shown by the vehicle lateral dynamics example.

Chapter 5: **Auxiliary signal design for active fault diagnosis** - A new framework of active fault diagnosis is proposed, and then, a discrimination condition is introduced to design the auxiliary signal. Two models case (Model detection) and multiple models case (Model isolation) are introduced respectively. Finally, a design example is given for both cases to illustrate the theory.

Chapter 6: **Conclusion and perspective** - gives a summary of thesis, and provides some recommendations for possible further research.

2. Review and algorithm

In this chapter, a review about multiobjective optimization in model based fault detection is given. A particular attention has been paid to the used specifications in the parity space method and observer based method. In order to overcome the shortage of classical methods in the multiobjective optimization for fault detection, a min-max formulation and the lexicographic formulation are introduced. Finally, a brief description of two different algorithms is given.

2.1. Review of optimization in model based fault detection

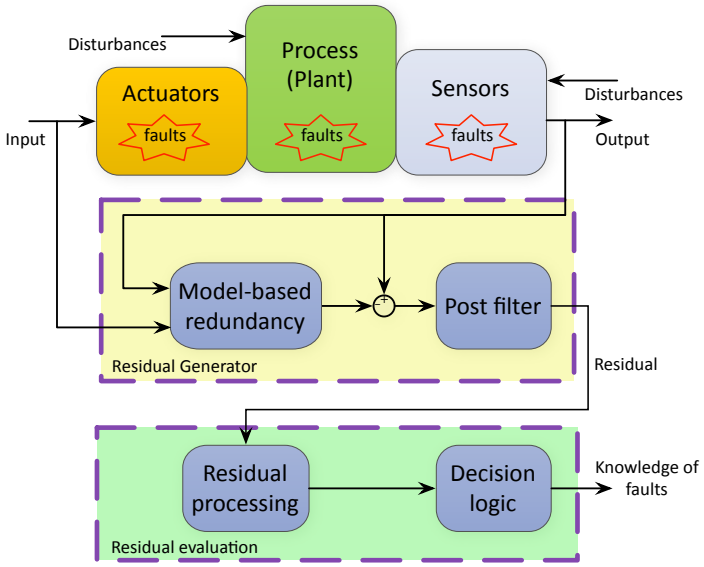

Figure 2.1.: Schematic description of model-based fault detection scheme

A general structure of a model-based fault detection (FD) containing two stages of residual generation and residual evaluation is illustrated in Fig. 2.1. In the residual generation part, a fault indicating residual is generated by the comparison between the system outputs and the output estimations. In ideal case, the residual should only carry the information of fault:

$$\text{if residual} \neq 0 \text{ then fault, otherwise fault-free}$$

In order to guarantee reliable FD, the information of fault in residual generation should be kept as "much" as possible. However, there always exist unknown disturbances and the modeling errors, to which the residual is also sensitive. Therefore, it is necessary to extract the fault information from the residual signals, which is done in the residual evaluation part.

2.1.1. Parity space methods

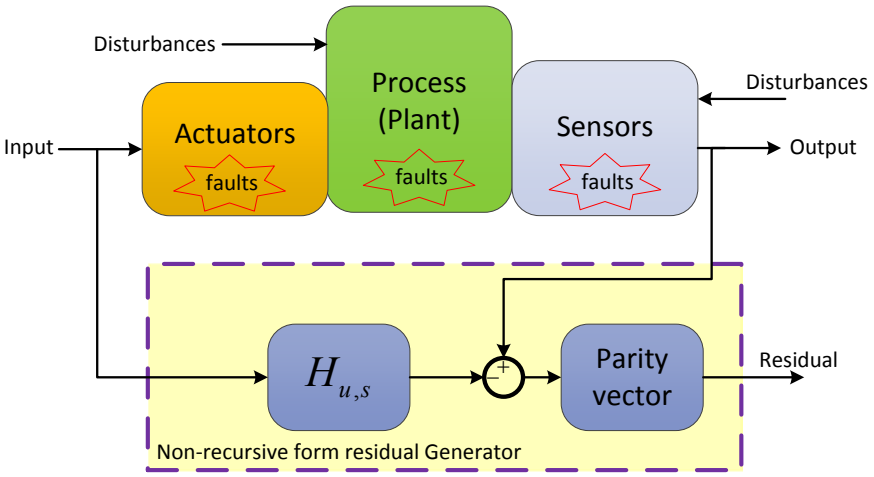

Figure 2.2.: Schematic description of residual generation with parity space methods

The parity space method is one of the most applied approach to generate residual. Initiated in [30], the parity space method could be explained with a state space model of a linear discrete time system:

$$\begin{cases} x(k+1) = Ax(k) + B_u u(k) + B_f f(k) + B_d d(k), \\ y(k) = Cx(k) + D_u u(k) + D_f f(k) + D_d d(k), \end{cases} \tag{2.1}$$

where $u(k) \in \mathbb{R}^r$ is the input vector, $y(k) \in \mathbb{R}^m$ is the output vector, $x(k) \in \mathbb{R}^n$ is the state vector, $f(k) \in \mathbb{R}^g$ represents a fault vector, and $d(k) \in \mathbb{R}^q$ denotes the disturbance.

The so-called parity relation could be constructed by collecting a series of data over an observable window of length $s > 0$:

$$y_s(k) = H_{o,s} x(k-s) + H_{u,s} u_s(s) + H_{f,s} f_s(s) + H_{d,s} d_s(s) \tag{2.2}$$

with

$$y_s(k) = \begin{bmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{bmatrix}, u_s(k) = \begin{bmatrix} u(k-s) \\ u(k-s+1) \\ \vdots \\ u(k) \end{bmatrix}, f_s(k) = \begin{bmatrix} f(k-s) \\ f(k-s+1) \\ \vdots \\ f(k) \end{bmatrix},$$

$$d_s(k) = \begin{bmatrix} d(k-s) \\ d(k-s+1) \\ \vdots \\ d(k) \end{bmatrix}, H_{o,s} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^S \end{bmatrix}, H_{u,s} = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{s-1}B & \cdots & CB & CBD \end{bmatrix},$$

$$H_{f,s} = \begin{bmatrix} D_f & 0 & \cdots & 0 \\ CB_f & D_f & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{s-1}B_f & \cdots & CB_f & CB_fD_f \end{bmatrix}, H_{d,s} = \begin{bmatrix} D_d & 0 & \cdots & 0 \\ CB_d & D_d & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{s-1}B_d & \cdots & CB_d & CB_dD_d \end{bmatrix}$$

Hence, the residual could be

$$\begin{aligned} r(k) &= v_s\left(y_s(k) - H_{u,s}u_s(s)\right) \\ &= v_s\left(H_{d,s}d_s(k) + H_{f,s}f_s(k)\right) \end{aligned} \tag{2.3}$$

where $v_s \in \mathbb{R}^{(s+1)m}$ is the residual generating vector, which also could be a matrix for multi-dimensional residual case:

$$v_s \in P_s, P_s = \{v_s \mid v_s H_{o,s} = 0\} \tag{2.4}$$

In the ideal case of fully decoupling from the unknown disturbances, the residual signal $r(k)$ should be zero for the fault-free case and non-zero for the faulty case:

$$v_s H_{d,s} = 0,\ v_s H_{f,s} \neq 0 \tag{2.5}$$

However, the conditions of (2.5) can only be achieved when the number of independent disturbances is smaller than the number of independent measurements [35]. A trade-off between the robustness of the residual generator to the disturbance and its sensitivity to the faults should be considered via

$$\min_{v_s \in P_s} J_s = \min_{v_s \in P_s} \frac{\|v_s H_{d,s}\|_F^2}{\|v_s H_{f,s}\|_F^2} \tag{2.6}$$

which is one of the most often used indices [48, 37, 118, 124, 123]. This multiobjective optimization problem could be solved by Generalized singular value decomposition (GSVD) [48] or method based on matrix pencil [27]. Comparing with other introduced methods in [27], the main advantage of GSVD and method based on matrix pencil is that the result is reliable and easily handle the case of nearly singular when the matrix $H_{d,s}$ is almost rank deficient. This performance index of the parity space approach will converge to the optimal result of the H_2 approach for observer based method when the window length s increases [124]. [122, 121, 120] developed approaches based on wavelet transform and infinite impulse response (IIR) to achieve both a good performance (2.6) and a low order parity vector (small a window length s). This method is also extended to Linear Parameter varying (LPV) systems recently [107, 106].

Particularly, for the fault sensitivity, a minimum influence of the fault on the residual attracts large attentions, which could be expressed as

$$S_{f,\ \_} = \inf_{f \neq 0} \frac{\|v_s H_{f,s} f\|}{\|f\|} \tag{2.7}$$

which is not a norm. Since the parity space method implements in time domain, $S_{f,\ \_}$ could be

$$S_{f,\ \_} = \underline{\sigma}\left(v_s H_{f,s}\right) \tag{2.8}$$

where $\underline{\sigma}\left(v_s H_{f,s}\right)$ represents the minimum singular value of matrix $v_s H_{f,s}$.

An analytical solution is proposed in [35] for the same formulation considering different criteria

to evaluate the matrices $H_{f,s}$ and $H_{d,s}$:

$$\min_{v_s \in P_s} J_1 = \min_{v_s \in P_s} \frac{\|v_s H_{d,s}\|_1}{\|v_s H_{f,s}\|_1} \tag{2.9}$$

$$\min_{v_s \in P_s} J_2 = \min_{v_s \in P_s} \frac{\|v_s H_{d,s}\|_2}{\|v_s H_{f,s}\|_2} \tag{2.10}$$

$$\min_{v_s \in P_s} J_\infty = \min_{v_s \in P_s} \frac{\|v_s H_{d,s}\|_\infty}{\|v_s H_{f,s}\|_\infty} \tag{2.11}$$

also

$$\min_{v_s \in P_s} J_- = \min_{v_s \in P_s} \frac{\|v_s H_{d,s}\|_-}{\|v_s H_{f,s}\|_2} \tag{2.12}$$

where $\|v_s H_{d,s}\|_\infty$ is known as a worst-case measurement of the effects of disturbances on the residual, while $\|v_s H_{f,s}\|_-$ is introduced as an evaluation of minimum influence of fault on the residual. $\|v_s H_{f,s}\|_\infty$ means a best-case handling of the effects of fault on the residual, which is meaningless in practice. Especially, the ratio criteria (2.12) means the worst-case from the FDI viewpoint.

Since the three conditions (2.4) and (2.5) are difficult to meet for nonlinear systems and systems with uncertainties, the criterion (2.6) is extended as

$$\min_{v_s \in P_s} J_s = \min_{v_s \in P_s} \frac{\left\| v_s \begin{bmatrix} H_{o,s} & H_{d,s} \end{bmatrix} \right\|_F^2}{\|v_s H_{f,s}\|_F^2} \tag{2.13}$$

in [77, 27, 48]. In this case, the residual will also be affected by the states of the system, and the states of the system is considered as the disturbances in the design (2.13).

Most work has been done on maximizing the fault sensitivity and minimizing the disturbance rejections simultaneously with the parity space method. However, the design of a good fault detector needs to consider other aspects with parity space method.

When the fault is a kind of step signal, i.e.

$$f_s(k) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

the parity space method may meet some problems to generate a suitable residual in time domain.

Example 2.1. The model is from [48]:

$$A = \begin{bmatrix} 0.5 & -0.7 & 0.7 & 0 \\ 0 & 0.8 & 0 & 0 \\ -1 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0.4 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}^T, B_f = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, B_d = \begin{bmatrix} 0 \\ 0.6 \\ 0 \\ -1 \end{bmatrix}, D_f = D_d = 0$$

The window length s is chosen as 6. A pulse fault starts from $1s$ to $4s$, and the disturbance is white noise: where the threshold is selected with the methods in [123]

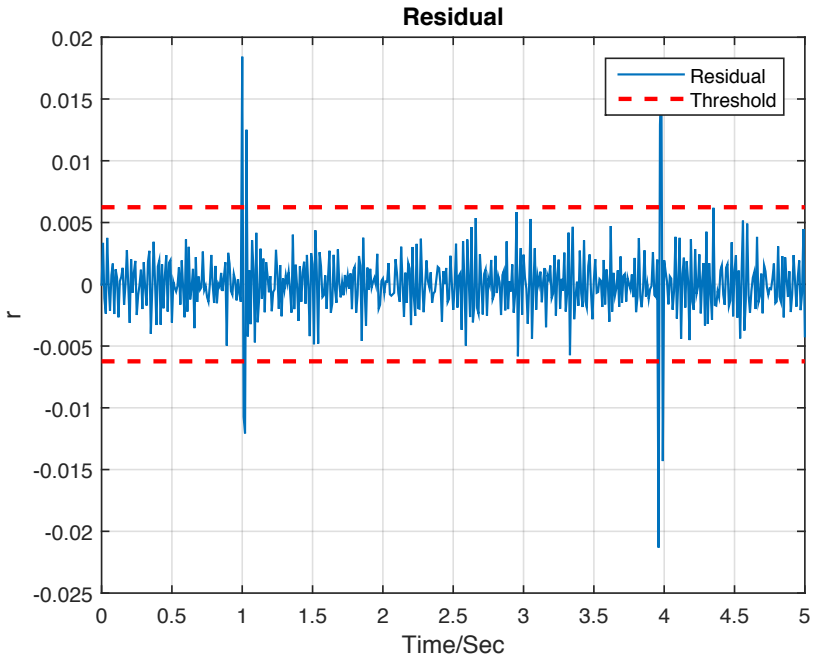

Figure 2.3.: Simulation of parity space method

$$J_{th,1} = \sup_{f=0,d_2} \left| v_s H_{d,s} H_{d,s}^T v_s^T \right| = \left(v_s H_{d,s} H_{d,s}^T v_s^T \right)^{1/2} \max_d |d(k)| \tag{2.14}$$

The simulation in Fig. 2.3 shows that the residual exceeds the threshold during $1s$ to $1.023s$, which means that the fault could be detected in this time interval. However, this residual decreases to be under the threshold fast, as a result, the alarm of fault detection vanishes even the fault persists. In practice, it is difficult to detect this pulse fault with this kind of short feasible fault detection interval. Furthermore, if the disturbance is too large, the pulse fault is hard to be detected with this kind of residual. A better residual with pulse signal fault should still be larger than the threshold after the wild oscillations, which can give more time to detect the pulse fault. However, it is difficult to add this kind of residual transients specification into the design with the classical parity space method. Thus, other framework

with some approximate specifications about the transients of residual should be considered for this typical case.

2.1.2. Observer based methods

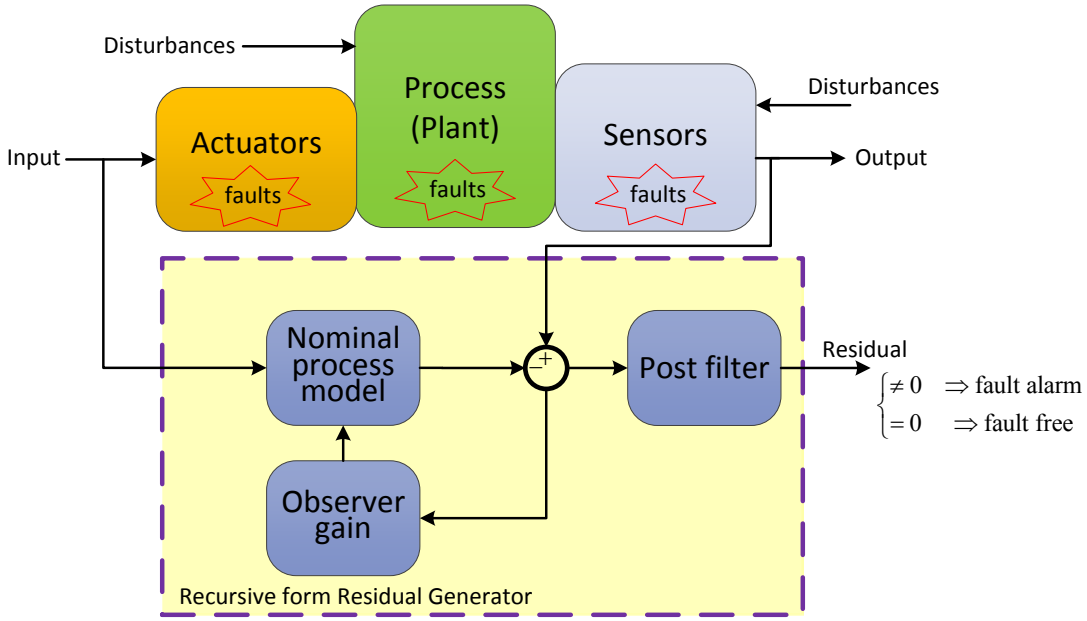

Figure 2.4.: Schematic description of residual generation with observer

Observer based method is one of the most applied model-based scheme for fault detection. The function of the observer is to estimate the output of the system, then, a residual is generated by comparing the outputs of the system and the output estimations. Different from the parity space methods in Fig. 2.2, the implementations of observer based methods use a recursive form while the parity space method represents a non-recursive form: as shown in Fig. 2.4, the outputs of systems are also introduced into the observer to revise the output estimations.

Two kinds of observers are proposed since early 70's: fault detection filter (FDF) and diagnostic observer. The diagnostic observer may lead to a reduced order residual generator for on-line implementation, while the fault detection filter needs the full order estimation.

With the framework of observer, the effects of inputs u are decoupled from the residual signal:

$$r = G_{rf} f + G_{rd} d \tag{2.15}$$

where G_{rf} is the transfer function from fault to residual and G_{rd} is the transfer function from disturbances to residual.

The formulations of residual with observer based method in (2.15) and with parity space in (2.3) are similar: the residual will be affected not only by the disturbances, but also by the faults. Owing to the recursive form of the observer based method, it provides the possibility to adjust the transients of residual, which cannot be realized by parity space method.

With the same reasons as parity space method, the perfect unknown decoupling conditions

$$G_{rd} = 0, G_{rf} \neq 0 \tag{2.16}$$

are difficult to achieve. An alternative formulation is considered

$$\min \|G_{rd}\|, \ \max \|G_{rf}\|$$

where $\|\cdot\|$ represents a norm of the matrix. Various specifications in frequency domain are applied to evaluate the effects of fault, model uncertainty and disturbances on the residual from different views, i.e. in the form of norms, like H_∞ and H_2 [55, 56, 70, 58, 59, 71]. The physical meaning of H_2 norm is that it evaluates the amplification of a transfer matrix which maps the inputs (which are either fixed or have a fixed power spectrum) into the outputs. By contrast, H_∞ measures the biggest amplification of a transfer matrix that maps the input with finite energy into the output.

Definition 2.1. The H_∞ norm of a transfer function $G(s)$ over the frequency range $[\omega_1, \omega_2]$ is defined as

$$\|G_{rd}(s)\|_\infty^{[\omega_1,\omega_2]} = \sup_{\omega\in[\omega_1,\omega_2]} \overline{\sigma}\left[G_{rd}(j\omega)\right] \tag{2.17}$$

where $\overline{\sigma}$ represents the maximum singular value.

Particularly, with the same general definition in (2.7), H_- index (in frequency domain) which describes the smallest fault detectability of a system is received a great deal of attentions. The H_- norm is firstly introduced in [54] by using the minimum singular value of the transfer function from faults to residual at zero frequency, i.e. $\omega = 0$. In this case, only the singular value at zero frequency is considered to design the observer, which will be too conservative. Then, the H_- notation is extended to nonzero singular value over finite frequency and infinite frequency ranges by [35, 93, 28, 27, 36]. In [116, 69], the H_- norm is further extended to H_- index, which contains the possible zero singular values of the transfer function from the fault to the residual over the infinite or finite frequency range.

Definition 2.2. The H_- index of a transfer function $G(s)$ over the frequency range $[\omega_1, \omega_2]$ is defined as

$$\|G_{rf}(s)\|_-^{[\omega_1,\omega_2]} = \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}\left[G_{rf}(j\omega)\right] \tag{2.18}$$

where $\underline{\sigma}$ represents the minimum singular value.

Different from the definition of $S_{f,\_}$ in (2.8), $\|G_{rf}(s)\|_-$ is not a scalar, but depends on the frequency. The importance of frequency range for H_- index is shown in Fig. 2.5: when the system is proper Fig. 2.5 (a), $\|G_{rf}\|_- \neq 0$ for $\omega \in [0, \infty]$. The minimum singular value of $\|G_{rf}(s)\|$ is β_1 when the fault is in high frequency. However, when system is strictly proper, the singular value of $G_{rf}(j\omega)$ is going to be 0 when the frequency ω tends to be ∞. Then, the specification H_- index is $\|G_{rf}\|_- \equiv 0$ with $\omega \in [0, \infty)$. Any design will obtain the same result $\|G_{rf}\|_- \equiv 0$, which means that the specification H_- index is meaningless. In this case, a frequency range $[\omega_1, \omega_2]$ should be added into the criterion, which could contain the information of the interesting faults (as shown in Fig. 2.5 (b)).

When fault and disturbance are modeled as an unknown energy or power bounded signal, H_- and H_∞ are reasonable performance indices to evaluate the sensitivity to the fault and robustness to the disturbance of the residual [71, 70].

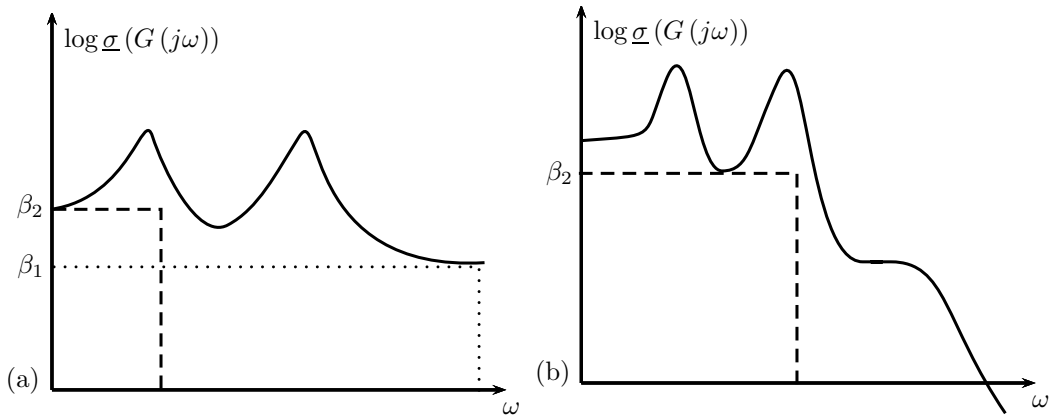

Figure 2.5.: Explanation of H_- index (a. Proper case, b. Strictly proper case)

Calculation of H_- index

First introduce a Lemma:

Lemma 2.1. *([112])Define the frequency range Ω by*

$$\Omega = \left(\left[\omega_l^1, \omega_h^1\right], \ldots, \left[\omega_l^n, \omega_h^n\right]\right)$$

Given transfer functions W with

$$\inf_{\omega\in\Omega} \underline{\sigma}\left[W\left(j\omega\right)\right] = \delta \quad \left\|W - G_{rf}\right\|_\infty = \alpha \tag{2.19}$$

then

$$\underline{\sigma}\left[G_{rf}\left(j\omega\right)\right] \geq \delta - \alpha, \forall\omega \in \Omega \tag{2.20}$$

To calculate the H_- index, in general, there are two ways:

- The indirect way: the fault sensitivity criterion is translated into an optimal H_∞ tracking problem[93, 112, 51, 52] with Lemma 2.1. The key point of this method is that the reference model (W in Lemma 2.1) should be carefully selected.
- The direct way: calculates the lowest singular value directly [24, 18, 91, 119, 116, 69, 35], using like linear matrix inequality (LMI) or the nonsmooth optimization method.

These two different ways have different advantages respectively: if there is some prior information about the faults, the indirect way is a better choice for the design because of the easy selection of the reference model W. Otherwise, the direct way can give us more freedoms to design potentially a "better" residual generator for the unknown faults case.

2.1.2.1. Mixed criterion to maximize fault sensitivity and minimize disturbance robustness

With different specifications to evaluate the effects of fault and disturbance on the residual, the mixed optimization problems could be separated with used specifications as H_2/H_2 case, H_2/H_∞ case, H_∞/H_∞ case and H_-/H_∞ case.

H_2/H_2 case

The problem is formulated as

$$\inf J = \inf \frac{\|G_{rd}\|_2}{\|G_{rf}\|_2} \tag{2.21}$$

which could be solved by the generalized singular value decomposition in frequency domain [40]. LMI and factorization techniques also could be applied to this mixed criterion [70, 35]. The relationship between the parity space method and H_2/H_2 is introduced in [124]. The H_2/H_2 residual generator has a property of bandpass, which means that the filter will be only sensitive to a fault signal near frequency ω_0. In the other word, most kinds of faults are deterministic, thus such filter is not very useful in practice [70, 35]. This characteristic of band-limited also could be found with the parity space method when $s \to \infty$.

H_2/H_∞ case

The problem is formulated as

$$\inf J = \inf \frac{\|G_{rd}\|_\infty}{\|G_{rf}\|_2} \tag{2.22}$$

A unified optimal solution is proposed for this mixed criterion in [71, 70]. The H_2/H_∞ filter is robust to the variations on B_f and D_f matrices. When $D_f \neq 0$, the value of mixed criteria (2.22) will be 0, which means that this fault detection filter is very sensitive to the sensor fault.

H_∞/H_∞ case

The problem is formulated as

$$\inf J = \inf \frac{\|G_{rd}\|_\infty}{\|G_{rf}\|_\infty} \tag{2.23}$$

which was solved by [89, 36, 40, 41, 94]. The interpretation of $\|G_{rf}\|_\infty$ is the maximum effects of fault on the residual, which means a best-case handling of the influence of fault on residual. The optimization of (2.23) will maximize the differences between the biggest effects of fault and disturbance on the residual. In the general case, this criterion is not interesting, due to the fact that a large $\|G_{rf}\|_\infty$ does not imply the residual is well sensitive to the fault since the faults may occur in the other direction where $\|G_{rf}\|_\infty$ is not large. However, if a bandpass filter is added to this optimization problem, this criterion will be interesting. In this case, the minimized cost function (2.23) makes the frequency points ω_{rf} ($G_{rf}(j\omega_{rf}) = \|G_{rf}\|_\infty$) and ω_{rd} ($G_{rd}(j\omega_{rd}) = \|G_{rd}\|_\infty$) be equal, which implies that the mixed specification (2.23) maximizes the ability of fault detection when the fault and disturbance are in the same frequency.

H_-/H_∞ case

The problem is formulated as

$$\inf J = \inf \frac{\|G_{rd}\|_\infty}{\|G_{rf}\|_-} \tag{2.24}$$

which is considered as an optimization in the worst case: improve the ability to detect fault in the case of minimum influence of fault and maximum effects of disturbance on the residual. The physical meaning of minimizing the mixed specification H_-/H_∞ (2.24) can be explained as the maximizing the set of strongly detectable faults [36]: the fault only can be detected when

$$\inf_d \|r\| > J_{th} \tag{2.25}$$

where J_{th} is a threshold with the definition $J_{th} = \sup_{f=0} \|r\|$ and $\|\cdot\|$ is the H_2 norm. Note that

$$\inf_d \|r\| = \inf_d \|G_{rf}f + G_{rd}d\| = \|G_{rf}f\| - \sup_d \|G_{rd}d\| \tag{2.26}$$

As a consequence, a fault can be detected for any disturbances if

$$\|G_{rf}f\| > \sup_d \|G_{rd}d\| + J_{th} = 2J_{th} \tag{2.27}$$

Only the fault f, satisfying (2.27), are strongly detectable in the presence of disturbances. Since

$$\|G_{rf}f\| > \|G_{rf}\|_- \|f\| \tag{2.28}$$

Then, a fault f is strongly detectable if the presence of disturbances if

$$\|G_{rf}\|_- \|f\| > 2J_{th} \tag{2.29}$$

which is equivalent to

$$\|f\| > \frac{2J_{th}}{\|G_{rf}\|_-} \tag{2.30}$$

Thus, the minimization of $\frac{2J_{th}}{\|G_{rf}\|_-}$ implies the maximization of the set of strongly detectable faults. Since the minimization of $\frac{2J_{th}}{\|G_{rf}\|_-}$ is equivalent to (2.24), the mixed specification (2.24) also implies the maximization of the set of strongly detectable faults [36].

A great deal of work has been done on this topic [53, 51, 52, 57, 69, 115, 93, 104, 112, 67, 91]. With the aid of LMI method, [91] calculated the H_- norm and designed the fault detection observer with the criterion of H_-/H_∞, whose condition is sufficient but not necessary. With the defined index, [116, 69] developed an LMI formulation for the multiple objectives of the fault detection observer and used the iterative linear matrix inequality (ILMI) to obtain the solutions. Considering the same mixed H_-/H_∞ criterion, some numerical optimization methods such as genetic algorithm [27] proposed to design the robust fault detection observer. Different from the approximate values with method in [116, 69, 91, 93], [114] gives an accurate characterization of H_- index in the low frequency domain with developed generalized Kalman-Yakubovic-Popov

(KYP) Lemma.

However, most discussed methods cannot guarantee that the obtained results are optimal. One exception case is that by using the co-inner-outer factorization approach, [36, 35] give an optimal solution for mixed H_-/H_∞ multiobjective optimization problem. The solution could be obtained by solving one algebraic Riccati equation (ARE), which leads to an optimal trade-off between fault detection rate and the false alarm rate. Then, [71, 70] proposes an optimal result in an observer form by solving a standard algebraic Riccati equation for continuous and discrete linear systems. In [66], the results are developed to linear continuous time-varying and linear discrete time-varying systems.

2.1.2.2. Standard H_∞ filtering with reference model

Since the specification $\|G_{rf}\|_-$ is difficult to calculated for systems with uncertainties, a new performance index is introduced to make the residual be as "similar" to a reference residual as possible. As shown in Fig. 2.6, P is the plant to be monitored, W is the reference model, and F is the filter to design. The estimation error is defined as $e = r - r_f$, where r is the residual vector.

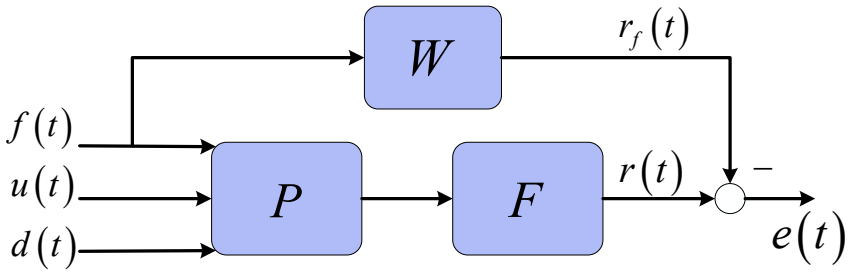

Figure 2.6.: Schematic description of H_∞ filter

Then, the design problem could be translated as

$$\sup_{\substack{w\in L_2\\ w\neq 0}} \frac{\|r - r_f\|_2}{\|w\|_2} = \sup_{\substack{w\in L_2\\ w\neq 0}} \frac{\|e\|_2}{\|w\|_2} \tag{2.31}$$

where $w = \begin{bmatrix} u^T & f^T & d^T \end{bmatrix}^T$. This problem could be formulated as

$$\min_F \|G_{rf}F - W\|_\infty \text{ and } \min_F \|G_{rd}F\|_\infty \tag{2.32}$$

With different selection of the reference model W, the fault isolation could be achieved [23, 60, 34, 68, 74, 111, 125, 85]. Typically, if the reference model is set as identity matrix $W = I$, the objective function (2.32) could be used to estimate or identify the faults, as in [84, 75]. Some methods to choose this reference model are introduced in [125, 42].

2.1.3. Extra criteria for optimization

The previous part mainly discussed the specifications of fault sensitivity and disturbance robustness. But in the field of FDI, these two specifications are not enough for the design. A

relative problem of parity space method is shown in Example Example 2.1. With the introduced properties of a good fault detection system, some other specifications should be met for fault diagnosis in practice, e.g., early fault detection.

Since the fast fault detection depends on the transients of residual, some criteria to evaluate the transients of residual responses should be considered. Due to the characteristic of poles of system, it is reasonable to consider the poles of designed filter as a specification. With aid of developed KYP, [32] considered a multiobjective optimization problem for calculating the real values of the H_- and H_∞ specifications in a finite frequency range with a constraint of filter poles in a specified region. To simplify the optimization, the region of poles is approximated as a disk region, which also decrease the feasible region to design. For linear uncertain dynamic systems, besides the mixed criteria H_-/H_∞ and regional constraints on filter poles, a generalized H_2 performance index is added to evaluate the peak amplitude of the residual caused by the energy bounded disturbances in the fault free case [51]. In [115], the main idea is to use the pole assignment approach to translate the fault detection problem into an unconstrained optimization problem, and then search for a desirable observer gain with the aid of a gradient-based optimization approach for both the infinite and finite cases. But in this method, the target poles of observer should be selected in advance, which also means that the criterion of poles is fixed in the optimization.

In these methods, a constraint of system poles is considered to improve the transients of the residual. However, the effects of fast transients responses on the ability of fault detection are not systematically detailed. The threshold selection also affects the time to detect faults. For the fast fault detection design, both transients of residual from fault and threshold selection should be considered simultaneously. Furthermore, some typical transients of residual responses in time domain should be formulated to decrease false alarms rate and missing alarm, e.g. the introduced shortage of parity space method (too short interval of fault detection). In this case, a constraint of residual in time domain could be added into the design to increase the interval of fault detection.

2.2. Multiobjective optimization

The design of a "good" fault detection system with different requirements is a typical multiobjective optimization problem. Due to the fact that it is impossible to make every specification achieve the best result, a trade-off among the different performance indexes has to be considered. Before introducing the classical methods used in the multiobjective optimization of FDI, the basic definition of multiobjective optimization and Pareto optimality are introduced first.

2.2.1. Definition of multiobjective optimization

Definition 2.3. (Multiobjective optimization problem [72])

$$\begin{aligned} \underset{x}{\text{Minimize}}\ & f(x) = [f_1(x), f_2(x), \ldots, f_k(x)]^T \\ \text{subject to}\ & g_j(x) \leqslant 0,\ j = 1, 2, \ldots, m \\ & h_l(x) = 0,\ j = 1, 2, \ldots, e \end{aligned} \tag{2.33}$$

where the vector $x \in \mathbb{R}^n$ is a vector of decision (also called design variables) variables representing the quantities for which values are chosen in the optimization problem, k is the number of objective function (Soft constraints of the design) $f_i : \mathbb{R}^n \to \mathbb{R}\,(i = 1, \ldots, k;\ k \geqslant 2)$, m is the number of inequality constraints (g_j: hard constraints of the design), and e is the number of equality constraints (h_l). In multi optimization, a compromise should be sought among the different objectives. In such case, there will not be only one optimal solution, but a set of equally feasible solutions, known as the Pareto optimal set.

The feasible set $\mathcal{X}$ (also called the feasible decision space or constraints set) is implicitly defined by the reference inequality and equality constraints. The feasible objective function space $\mathcal{Z}$ (also called the feasible attainable set or cost space) is defined as the set $\{f\,(x) \mid x \in \mathcal{X}\}$.

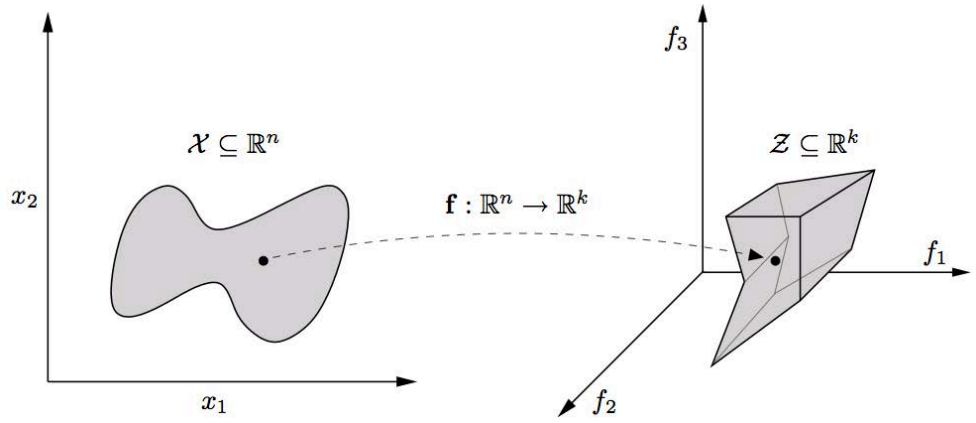

Figure 2.7.: Space definition in multiple objective optimization

Definition 2.4. Pareto Optimal: A point, $x^* \in \mathcal{X}$, is Pareto optimal iff there does not exist another point, $x \in \mathcal{X}$, such that $f_i\,(x) \leqslant f_i\,(x^*)$, and $f_i\,(x) < f_i\,(x^*)$ for at least one function.

Definition 2.5. Weakly Pareto Optimal: A point, $x^* \in \mathcal{X}$, is weakly Pareto optimal iff there does not exist another point, $x \in \mathcal{X}$, such that $f_i\,(x) < f_i\,(x^*)$.

2.2.2. Shortages of classical methods in fault detection for multiobjective optimization

2.2.2.1. Classical methods in fault detection for multiobjective optimization

Although the unified solution given in [35, 36] is optimal for H_i/H_∞ (i is either ∞ or $-$) case by factorization techniques, this method could not support the optimization with extra performance indices. It is the reason that LMI techniques are proposed to solve FDI multiobjective optimization problem. A mixed specifications of H_-/H_∞, H_2/H_2 are proposed to solve by LMI [35, 91, 69, 116, 85, 84, 104, 126].

Considering the same mixed H_-/H_∞ criterion, some numerical optimization methods such as genetic algorithm [27, 26] are proposed to design the robust fault detection observer. Kowalczuk et al developed genetic algorithms to solve multi optimization problem to detect faults in the framework of observer based method [63, 61, 62]. The evolutionary algorithms could be extended to nonlinear systems [86]. Both of them consider the problem into a H_∞ formulation.

Major work about FDI multiobjective optimization has been done by LMI method, the following part tries to analysis the LMI methods, and illustrate some weaknesses of LMI in the view of multiobjective optimization design.

2.2.2.2. Shortages of LMI

Definition 2.6. A linear matrix inequality has the form

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0 \tag{2.34}$$

where $x \in \mathbb{R}^n$ is the variable and the symmetric matrices $F_i = F_i^T \in \mathbb{R}^{n \times n}$, $i = 0, \ldots, m$ are given. The inequality symbol in (2.34) is positive-definite, i.e., $u^T F u > 0$ for all nonzero $u \in \mathbb{R}^n$. The LMI (2.34) is equivalent to a set of n polynomial equalities in x.

The LMI techniques arise in FDI and control theory can be formulated as convex optimization problem. The convexity is a typical important characteristic in the optimization theory, and the foundation for analytical tools used in LMI techniques [12].

The optimization with LMI constraints or semidefinite programming (SDP) needs the mathematical theory for translating control criteria into LMI constraints. However, as introduced in [95], some criteria formulated with LMI may give conservative result. In reality, some of control and FDI engineering specifications are not convex. Therefore, many approaches are proposed to solve the nonconvex problems with LMIs techniques, such as LMI relaxations [45, 64, 50, 65]. However, conservatism will still arise when simplifying the nonconvex specifications. Most of the approaches are based on sufficient conditions but not necessary and sufficient conditions [50]. In addition, some specifications are difficult to translate the criterion into formulation of LMIs. Even the criterion could be translated into the formulation of LMIs, it is difficult to evaluate accurately for some practical problems in control [15] and FDI problems [113, 69] because of the conservatism inherent in LMI techniques. [12] included that it might solve some problems in some proper way, the LMI techniques are not entirely prepared to handle truly multiobjective control problems.

In the view of consumption of calculation , the approach LMI needs the introduction of Lyapunov variables, which grow quadratically with the number of state variables. Consequently, the total number of variables can be quite large and even problems of moderate size can lead to numerical difficulties. It is the reason why LMI techniques are not appropriate for large or even moderate scale systems.

2.2.3. Weighted min-max formulation for nonsmooth optimization

With the developed nonsmooth optimization, it is possible to solve FDI multiobjective optimization problem better than the classical method, like LMI. The nonsmooth optimization solves the multiobjective optimization problem with a min-max formulation, which has better characteristics than the formulation of weighted-sum typically for nonconvex problem [19]. Before introducing the algorithm, the min-max formulation is given, which formulates the objective

functions as the maximum of the weighted criteria:

$$\begin{aligned} \underset{x\in\mathcal{X}}{\text{Minimize}}\ \mathrm{f}_{\infty}(\mathrm{x}) &= \max\left(\lambda_1 f_1(x), \ldots, \lambda_k f_k(x)\right) \\ \text{subject to } g_j(x) &\leqslant 0,\ j = 1, 2, \ldots, m \\ h_l(x) &= 0,\ l = 1, 2, \ldots, e \end{aligned} \tag{2.35}$$

where λ_i are nonnegative weights, which are meant to express the importance of the criteria.

One variation of the weighted-max objective (2.35) is to add a constant offset for each criterion f_i:

$$\begin{aligned} \underset{x\in\mathcal{X}}{\text{Minimize}}\ &\max\left(\lambda_1\left(f_1(x) - \gamma_1\right), \ldots, \lambda_k\left(f_k(x) - \gamma_k\right)\right) \\ \text{subject to } g_j(x) &\leqslant 0,\ j = 1, 2, \ldots, m \\ h_l(x) &= 0,\ l = 1, 2, \ldots, e \end{aligned} \tag{2.36}$$

Among the results of multi criterion optimization, the Pareto optimal ones are wise choices for the design. Because the Pareto optimal results depend on the selection of weights λ_i, it remains to choose a suitable one of these, i.e., "search" the tradeoff surface for the design. The selection of the weights is always done interactively. The weights λ_i is repeatedly adjusted with the weighted max objective (2.35), then, the resulting optimal design could be evaluated. [19] introduces several procedures to choose the weights for the design in the views of application.

The formulation of min-max (2.35) and (2.36) are continuous but not differentiable everywhere, which means that the objective function f_∞ in (2.35) is nonsmooth. For example, Fig. 2.8 shows a typical case with $k = 2$: the objective function f_∞ at the point of x^* is continues but not differentiable, which means that the min-max formulation is typically a nonsmooth optimization problem. In this case, some approaches, which do not depend on the gradient, should be considered. Consequently, the differentiable optimization methods (gradient descent, Newton and quasi-Newton sequential quadratic programming problems under constraints, etc.) are not sufficient.

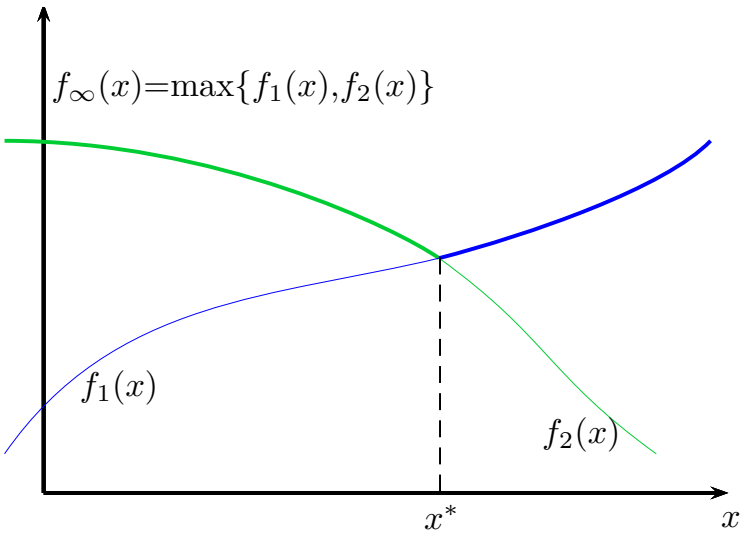

Figure 2.8.: Max function on $\mathbb{R}$: f_∞ is not derivable at x^*

2.3. Algorithm

As introduced in [13], in practice, if the (sub)gradients are available, it is better to apply the (sub)gradients for the efficiency. However, if obtaining (sub)gradient information is difficult (or the analytical formulation of a specification does not exist), an approximate (sub)gradients (i.e. a derivative-free method) could be considered, even the efficiency is not high. In this section, two algorithms are given. Since min-max formulation (2.35) is typically a nonsmooth optimization problem, the first algorithm about nonsmooth optimization is introduced for the case that the subgradients could be obtained. The second algorithm is genetic algorithm (GA) , which is used when the subgradient information is hard to obtain.

2.3.1. Nonsmooth optimization

Different from the classical methods, like LMI, the nonsmooth optimization method avoids the Lyapunov variables. Therefore, the nonsmooth optimization method works faster than LMI, and is suitable for the large size plant. Using the information of subgradients, the nonsmooth optimization method also works faster than the derivative-free methods (e.g. genetic algorithm).

In this part, the key ingredients of nonsmooth optimization are introduced, the reader could refer to [11, 98, 10, 16] for the details. Here, the general procedure is introduced.

For the sake of clarity, the formulation of multiobjective optimization (2.35) could be represented as a more general form:

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad g(x) \leq 0 \end{aligned} \tag{2.37}$$

A progress function is proposed with an idea in [87]:

$$F\left(x^{+}, x\right) = \max\left\{f\left(x^{+}\right) - f(x) - \mu g(x)_{+}; g\left(x^{+}\right) - g(x)_{+}\right\} \tag{2.38}$$

where $\mu > 0$ with some fixed value. x represents the present iterate, x^{+} is a candidate for the next iterate. It is shown in [87] that a critical point $\overline{x}$ of $F(\cdot, \overline{x})$ is also the critical point of the constrained optimization problem (2.37), except the case when $\overline{x}$ is a local minimum of the constraint violation $g(\overline{x}) > 0$.

A first-order approximation $\hat{F}(\cdot, x)$ of $F(\cdot, x)$ around x could be

$$\begin{aligned} \hat{F}(x+h, x) = \max\Bigg\{ &\max_{(\phi_f, \Phi_f) \in \Upsilon_f(x)} \phi_f - f(x) - \mu g(x)_{+} + \Phi_f^T h, \\ &\max_{(\phi_g, \Phi_g) \in \Upsilon_g(x)} \phi_g - g(x)_{+} + \Phi_g^T h \Bigg\} \end{aligned} \tag{2.39}$$

where h is the displacement for the design parameter x, $g(x)_+ \triangleq \max\{0, g(x)\}$ with

$$\phi_f = \phi_f(x,y), \phi_g = \phi_g(x,y)$$
$$\Phi_f \triangleq \nabla_x \phi_f(x,y), \Phi_g \triangleq \nabla_x \phi_g(x,y)$$
$$\Upsilon_f(x) \triangleq \{(\phi_f, \Phi_f) : y \in Y_f\}, \Upsilon_g(x) \triangleq \{(\phi_g, \Phi_g) : y \in Y_g\}$$

Then a tangent program

$$\underset{h \in R^q}{\text{minimize}} \qquad \hat{F}(x+h, x) + \frac{\delta}{2}\|h\|^2 \tag{2.40}$$

with $\delta > 0$ a fixed value could be obtained. As introduced in [98], the formulation (2.40) could be transformed to

$$\begin{aligned} \underset{t,h \in R^q}{\text{Minimize}} \quad & t + \frac{\delta}{2}\|h\|^2 \\ \text{subject to} \quad & \phi_f - f(x) - \mu g(x)_+ + \Phi_f^T h \le t, (\phi_f, \Phi_f) \in \Upsilon_f(x) \\ & \phi_g - g(x)_+ + \Phi_g^T h \le t, (\phi_g, \Phi_g) \in \Upsilon_g(x) \end{aligned} \tag{2.41}$$

which is standard convex quadratic program (CQP) when the eigenvalue multiplicity of all maximum eigenvalue functions equals to 1.

Algorithm 2.1 A nonsmooth algorithm program

Step 1. Initialize. Select initial x^1, a counter $j = 1$, $\delta > 0$, $0 < \beta < 1$, $0 < \gamma < 1$;
Step 2. Stopping test. At counter j, stop if $0 \in \partial_1 \hat{F}(x^j, x^j)$ and return x^j; Otherwise continue;
Step 3. Compute descent direction. Solve tangent programs (2.40) or (2.41)

$$\underset{h \in R^q}{\text{minimize}} \qquad \hat{F}(x+h, x) + \frac{\delta}{2}\|h\|^2$$

The search direction h^j could be obtained.
Step 4. Line search. Find $\xi = \beta^v$, $v \in N$, satisfying the Armijo condition [33]

$$F\left(x^j + \xi h^j, x^j\right) - F\left(x^j, x^j\right) \le \gamma \xi F'\left(\cdot, x^j\right)\left(x^j, h^j\right) < 0$$

Step 5. Update. Put $x^{j+1} = x^j + \xi h^j$, $j = j + 1$ and return to step 2.

2.3.2. Genetic algorithm

Genetic algorithm does not need the information of (sub)gradients of the objective function, which is appropriate for the case when the objective function is not analytical. This algorithm could be a good candidate when the introduced nonsmooth optimization in ([10, 99]) is not available.

Genetic algorithm (GA) solves optimization problems based on the mechanics of natural selection and genetics, which are part of a broad class of search techniques within the area of evolutionary algorithms (EAs). The standard genetic algorithm (SGA), presented in [46], con-

tains three genetic operators: reproduction, crossover, and mutation. Probabilistic rules are proposed to perform the evolution, which manipulate the "genetic code".

Replication (Selection)

The function of replication or selection part is to select the best individuals to participate in the production of offspring. Typical selection operators for this process could be:

- Sampling with replacement or roulette wheel selection.
- Stochastic universal sampling.
- Tournament selection.

Crossover (Recombination)

After replication, the crossover part exchanges genetic information with other to produce a single individual with the genetic features of two or more parents.

Mutation

According to some probabilistic rules, the individual's genotypes are changed in this genetic operation. The mutation is in the random way. In artificial genetic algorithms, an irrecoverable loss of some potentially useful genetic material is protected against by the mutation operator.

The implement of the SGA is shown in Fig. 2.9.

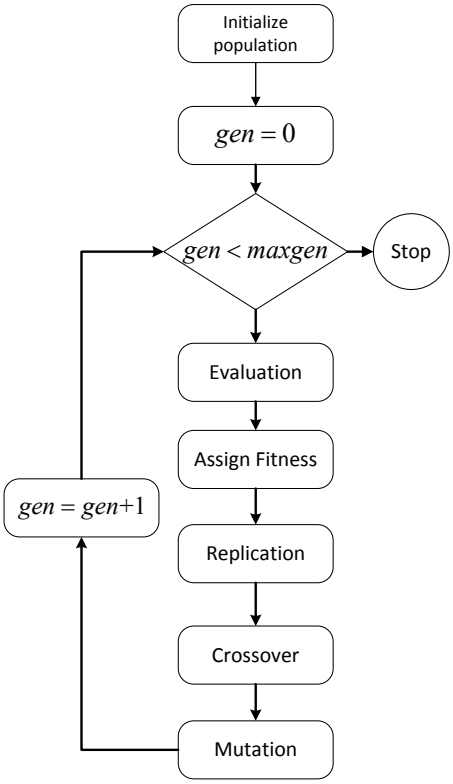

Figure 2.9.: Flowchart of the Standard Genetic Algorithm

Some other extensions of genetic algorithm could be found in [76][49][44][12].

3. H_-/H_∞ fault detection filter design via nonsmooth optimization approach

3.1. Introduction

As introduced in the previous part, in general, there are several following problems for LMI methods:

- Conservative solutions. Since the mixed H_-/H_∞ could be formulated as two LMIs, a same additional matrix is always introduced to meet these two LMIs simultaneously. As illustrated in [91, 112, 114], a common LMI solution result is always obtained, and some conservatism is introduced because of the same additional matrix.
- Low efficiency. Some work has been done to relax the constraint of same additional matrix, such as [116, 69]. Nevertheless, the iterative linear matrix inequality (ILMI) algorithm has to be introduced to search the local solution [116, 69, 47, 112, 104]. A disadvantage of ILMI algorithm is that the search step is not optimal, thus, the convergence rate could not be guaranteed. In addition, some pre-defined variables always have to be introduced into the ILMI algorithm for optimization, which sometimes are difficult to choose for the optimal result.
- Not suitable for large scale systems. The LMI approach uses Lyapunov variables, whose number grows quadratically with the system state size. Therefore, the LMI approach will meet some problem for medium and large scale systems [10].

Due to the advantages of nonsmooth optimization approach, this chapter tries to revise the mixed H_-/H_∞ fault detection filter design problems with proposed nonsmooth optimization approach and design with H_-/H_∞ specifications for multiple models system.

In order to compare classical methods with the proposed nonsmooth optimization approach, a formulation of classical fault detection observer with a mixed performance index H_-/H_∞ is introduced in Section 3.2. The cost function in this part includes both disturbance attenuation with H_∞ norm and fault sensitivity with H_- index, which is a typical nonsmooth optimization problem. What's more, a constraint of fast transients of residual responses from faults can be added into the optimization. Typically, this technique is easily applied to design an observer with a unique observer gain and residual weighting matrix for multiple models. In this design, the observer has to work with an injection of switching signals (information about model parameters of the system) for the multiple models, because the model parameters in the observer (A, B, C and D) are relevant to the parameters of the system. For a typical case that the system stays in one of the multiple models, but the exact model is unknown, the framework of observer is not appropriate for this case (observer needs the exact system parameters). Alternatively, a framework of robust fault detection filter (RFDF) is formulated in Section 3.3 for this typical case, which does not need to know the exact model. The RFDF is designed by means of

a deconvolution approach [23, 24, 60, 100], which could also be solved by proposed nonsmooth optimization approach. The solver with nonsmooth optimization approach is adopted to solve the formulated problems in Section 3.4. Then in Section 3.5, a comparison between classical methods and the nonsmooth optimization approach is given with an academic example for the single model case. The relationship between the formulations of fault detection observer and robust fault detection filter is detailed with the aid of this single model case. For the multiple models case, the compromised designs with the formulations of fault detection observer and robust fault detection filter are illustrated with a vehicle lateral dynamics switched system with 3 subsystems. Finally, a conclusion is given in Section 3.6.

Recently, solvers relying on nonsmooth optimization techniques like Hinfstruct and Systune [10, 8, 11] are well developed. This is the first work about the application of the nonsmooth optimization techniques in FDI problems with this tool.

3.2. Mixed H_-/H_∞ fault detection observer design

As introduced in previous chapter, fault detection observer is one of the mostly applied model-based approach to detect faults. Therefore, it is reasonable to revise the traditional problem of mixed H_-/H_∞ fault detection observer design with the new developed nonsmooth optimization techniques, and compare these new techniques with classical methods. This section tries to show the effectiveness of nonsmooth optimization techniques on the fault detection filter design: it can be used to design H_-/H_∞ fault detection observer with an additional constraint of rapidity of residual response, and works faster than classical methods. Furthermore, a more complex problem about a trade-off design for multiple models could be also easily solved by the nonsmooth optimization approach.

3.2.1. Residual generation

Assuming that we have $N \geqslant 1$ models describing different behaviors of a system, the linear time invariant (LTI) system for multiple models with faults and disturbances is described by

$$\Sigma_0 \begin{cases} \dot{x} = A^i x + B_u^i u + B_f^i f + B_d^i d, \\ y = C^i x + D_u^i u + D_f^i f + D_d^i d, \end{cases} \tag{3.1}$$

where $i \in \{1, 2, \ldots, N\}$ represents that the system is in i^{th} operational mode, $x \in \mathbb{R}^{n_x}$ is the system state vector, $y \in \mathbb{R}^{n_y}$ represents the output measurement vector, $f \in \mathbb{R}^{n_f}$ denotes the fault vector, which can be process faults, sensor faults, or actuator faults. $d \in \mathbb{R}^{n_d}$ is the unknown input vector, including disturbance, modeling error, process and measurement noise or uninterested fault. $u \in \mathbb{R}^{n_u}$ is the control input vector. Matrices A^i, B_u^i, C^i, D_u^i, B_f^i, D_f^i, B_d^i, D_d^i are constant with appropriate dimensions. The single model case could be described by the above model Σ_0 with $N = 1$.

Without loss of generality, the following assumptions are used:

- (A^i, C^i) is detectable;
- f and d are $\mathcal{L}_2$ norm bounded.

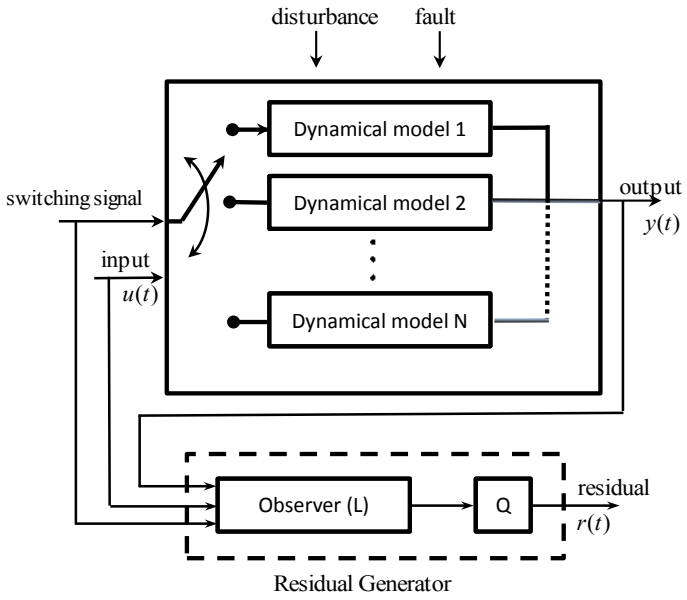

Figure 3.1.: Residual generator with observer

As shown in Fig. 3.1, a full-order observer in the following form [27] is used to generate residual r for the multiple models system:

$$\Sigma_1 \begin{cases} \dot{\hat{x}} = A^i\hat{x} + B_u^i u + L(y - \hat{y}), \\ \hat{y} = C^i\hat{x} + D_u^i u, \\ r = Q\left(y - \hat{y}\right). \end{cases} \tag{3.2}$$

where $i \in \{1, 2, \ldots, N\}$. $\hat{x} \in \mathbb{R}^{n_x}$ and $\hat{y} \in \mathbb{R}^{n_y}$ are the system's state and output estimations, $r \in \mathbb{R}^{n_r}$ is the corresponding residual vector, $L \in \mathbb{R}^{n_x \times m_y}$ is the observer gain to design, and $Q \in \mathbb{R}^{n_r \times n_y}$ is the residual weighting matrix, which could be static or dynamic as $Q(s)$. In this framework, the switching signal is known, which is used to switch the system parameters A^i, B_u^i, C^i and D_u^i in (3.2).

Connecting the observer $\sum_1$ (3.2) with the system $\sum_0$ (3.1) together as shown in Fig. 3.1, and considering the state estimation error as $e = x - \hat{x}$, the residual error dynamic equations can be obtained:

$$\Sigma_2 \begin{cases} \dot{e} = (A^i - LC^i)e + (B_f^i - LD_f^i)f \\ \qquad\quad + (B_d^i - LD_d^i)d, \\ r = QC^i e + QD_f^i f + QD_d^i d \end{cases} \tag{3.3}$$

The corresponding residual responses from faults and disturbances are:

$$\begin{aligned} r =& \mathfrak{L}^{-1}\left(Q\{D_f^i + C^i(sI - A^i + LC^i)^{-1}(B_f^i - LD_f^i)\}f\right) \\ &+ \mathfrak{L}^{-1}\left(Q\{D_d^i + C^i(sI - A^i + LC^i)^{-1}(B_d^i - LD_d^i)\}d\right) + \Psi_i\left(x\left(0\right)\right) \\ =& \mathfrak{L}^{-1}\left(G_{rf}^i(L,Q)f\right) + \mathfrak{L}^{-1}\left(G_{rd}^i(L,Q)d\right) + \Psi_i\left(x\left(0\right)\right) \end{aligned} \tag{3.4}$$

where $\mathfrak{L}^{-1}$ means the inverse Laplace transform, $\Psi_i\left(x\left(0\right)\right)$ is the natural response of i^{th} model with the initial state $x\left(0\right)$ when the switching of system and observer occurs, which depends

on the state of system. The part $\Psi_i(x(0))$ appears on residual when switching arises, and converges to 0 after switching.

In order to detect faults under the perturbations of disturbances, the residual should be sensitive to the faults and robust to the disturbances and uncertainties simultaneously. The following two conditions should be satisfied if perfect decoupling is achieved:

$$G_{rf}^i f \neq 0 \tag{3.5}$$

$$G_{rd}^i d = 0 \tag{3.6}$$

where G_{rf}^i means the transfer function from fault to the residual, G_{rd}^i represents the transfer function from disturbance to the residual. In this case, the residual (3.4) is

$$r = \mathfrak{L}^{-1}\left(G_{rf}^i(L,Q)f\right) + \Psi_i(x(0)) \tag{3.7}$$

Because the switching signal i is known, the time when $\Psi_i(x(0))$ appears is also known. The effects of $\Psi_i(x(0))$ on residual could be decreased by waiting for enough time after the switching.

Since the perfect decoupling conditions (3.5) and (3.6) are hard to meet in reality, a trade-off between the maximization of the faults sensitivity and the minimization of the disturbances sensitivity is always considered instead. Therefore, the multiobjective design of robust fault detection observer (design the observer gain L and the residual weighting matrix Q) should involve the following objectives:

i) The $1, ..., N$ residual error dynamics equations (3.3) with the observer gain L should be stable,

ii) Minimize the effects of disturbances on residual r,

iii) Maximize the effects of faults on residual r,

iv) Faster transients of residual r.

In order to achieve early fault detection, a residual with fast transients will give more potentials to design a fast robust fault detection observer. A specification to evaluate the rapidity of the residual should be included. Therefore, some certain performance indexes should be proposed to measure the above four design objectives exactly.In the worst case, if fault f is modeled as an unknown energy or power bounded signal, $\| G_{rf} \|_-$ is a reasonable specification to measure the minimum fault sensitivity of the residual. As the problem in control design, the norm H_∞ is a widely accepted measurement and $\| G_{rd} \|_\infty$ is a reasonable performance index of disturbance rejection, if disturbance d is also modeled as an unknown energy or power bounded signal.In order to improve the transients of residual, the spectral abscissa of $A^i - LC^i$ ($\alpha(A^i - LC^i)$) could be used to evaluate the rapidity of responses in frequency domain:

$$\alpha\left(A^i - LC^i\right) \triangleq \max_j Re\left(\lambda_j\left(A^i - LC^i\right)\right) \tag{3.8}$$

where the symbols $\lambda_1(A^i - LC^i)$, ..., $\lambda_n(A^i - LC^i)$ are the eigenvalues of $A^i - LC^i$, and $Re(\lambda_j(A^i - LC^i))$ represents the real part of $\lambda_j(A^i - LC^i)$. Fast transients could be achieved with a minimization of following specification

$$\min \alpha\left(A^i - LC^i\right) \tag{3.9}$$

The designed observer is asymptotically stable if the spectral abscissa of matrix $A^i - LC^i$ are negative.

Based on the definitions of $\mathcal{H}_\infty$ norm (2.17) and $\mathcal{H}_-$ index (2.18), the above four aspects could be formulated as follows:

i) $A_0^i = A^i - LC^i$ is asymptotically stable,

ii) $$\max_{L,Q} \parallel G_{rf}^i \parallel_- = \max_{L,Q} \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}(G_{rf}^i)$$

$$= \max_{L,Q} \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}\left(QD_f^i + QC^i(j\omega I - A^i + LC^i)^{-1}(B_f^i - LD_f^i)\right),$$

iii) $$\min_{L,Q} \parallel G_{rd}^i \parallel_\infty = \min_{L,Q} \sup_{\omega'\in[\omega_1',\omega_2']} \bar{\sigma}(G_{rd}^i)$$

$$= \min_{L,Q} \sup_{\omega'\in[\omega_1',\omega_2']} \bar{\sigma}\left(QD_d^i + QC^i(j\omega I - A^i + LC^i)^{-1}(B_d^i - LD_d^i)\right),$$

iv) $$\min_{L} \ \varsigma, \quad \alpha\left(A^i - LC^i\right) < \varsigma, \ \varsigma < 0.$$

[27] proposes that $\max_{L,Q} \left\|G_{rf}^i\right\|_-$ can be calculated by the following formulation:

$$\begin{aligned} &\min_{L,Q} \sup_{\omega\in[\omega_1,\omega_2]} \bar{\sigma}((G_{rf}^i)^{-1}) \\ &= \min_{L,Q} \sup_{\omega\in[\omega_1,\omega_2]} \bar{\sigma}\left(\left(QD_f^i + QC^i(j\omega I - A^i + LC^i)^{-1}(B_f^i - LD_f^i)\right)^{-1}\right) \end{aligned} \tag{3.10}$$

with the condition that matrix G_{rf}^i is invertible because of

$$\left\|G_{rf}^i\right\|_- \left\|(G_{rf}^i)^{-1}\right\|_\infty = 1 \tag{3.11}$$

When the system is singular or non square, which means that the inverse of G_{rf}^i does not exist, the formulation in (3.10) is no longer appropriate to calculate $\max_{L,Q} \left\|G_{rf}^i\right\|_-$. Then, two different cases should be considered:

- When the system is singular, or there is only actuator faults in the model ($D_f = 0$ in (3.1)), the value of $\left\|G_{rf}^i\right\|_-$ is always zero for $\omega \in [0, \infty)$. Then, a finite frequency range should be added into $\left\|G_{rf}^i\right\|_-$. Lemma 2.1 could be used in this case with a appropriate selection of the reference model.

- When the transfer function matrix G_{rf}^i from fault to the residual is not square, an alternative approach is to calculate the Moore–Penrose pseudo inverse $\left(G_{rf}^i\right)^+$:
  Lemma 3.1. *([92, 104])For a given system in transfer function as G, if its Moore-Penrose pseudo inverse system exists denoted by G^+, then the calculation of* $\max \ \|G\|_-$ *is equivalent to calculate* $\min \ \|G^+\|_\infty$*, which also could be represented as* $\|G\|_- > \beta$ *if and only if* $\|G^+\|_\infty < \frac{1}{\beta}$.

3.2.2. Formulation for nonsmooth optimization

In [35, 91, 116], there are numerous formulations for the problem of H_-/H_∞ fault detection observer design, and the formulation to be solved by LMI method is:

$$\begin{aligned} \left\|G_{rf}^i(L,Q)\right\|_- &> \beta_i \\ \left\|G_{rd}^i(L,Q)\right\|_\infty &< \gamma_i \\ A^i - LC^i \text{ is asymptotically stable} & \end{aligned} \tag{3.12}$$

where the β_i and γ_i are the parameters to maximize and minimize. For the LMI approach, another important work is to use iterative method to search the minimum γ_i and the maximum β_i [112, 116, 69]. As introduced in the previous part, in view of introducing the sufficient but not necessary conditions, non-optimal search length and quadratically growing Laypunov variables, results of ILMI are always conservative and this iterative method converges slowly. By contrast, from the optimization point of view, a ratio formulation of (3.12) is considered

$$\min_{L,Q} \frac{\|G_{rd}^i(L,Q)\|_\infty}{\left\|G_{rf}^i(L,Q)\right\|_-}, \tag{3.13}$$

$$A - LC \text{ is asymptotically stable} \tag{3.14}$$

The optimization with the objective function (3.13) is a kind of semi-infinite programming (SDP) problem, which is nonsmooth. Since the gradients of this objective function (3.13) does not exist, it is difficult to minimize this nonsmooth objective function by classical gradient-based methods.

Because it is impossible to achieve optimal results with the criteria of (3.13) for each model simultaneously, Pareto optimal results should be introduced in this case:

$$\min_{L,Q} \max_{i=1,\ldots,N} \left(\lambda_i \frac{\|G_{rd}^i(L,Q)\|_\infty}{\left\|G_{rf}^i(L,Q)\right\|_-} \right), \tag{3.15}$$

$$A^i - LC^i \text{ is asymptotically stable} \tag{3.16}$$

where λ_i are appropriate non-negative weights.

The problem (3.15) is formulated as minimax optimization, which is also a typically nonsmooth optimization design problem. With a view of LMI formulation, $2N$ inequalities have to be considered [2, 3]. Too many constraints of inequalities will decrease the calculation efficiency with LMI method. On the other hand, the minimax formulation problem (3.15) can be easily solved by nonsmooth optimization method even with many models.

Remark 3.1. For the multiple model case, the objective function (3.15) also could be formulated

as a weighted-sum case:

$$\min_{L,Q} \sum_{i=1}^{N} \left(\lambda_i \frac{\|G_{rd}^i(L,Q)\|_\infty}{\left\|G_{rf}^i(L,Q)\right\|_-} \right), \tag{3.17}$$
$$A^i - LC^i \text{ is asymptotically stable}$$

where λ^i are appropriate non-negative weights. However, if the formulated problems are not convex, some Pareto optimal results could not be obtained with any selection of weights for different models with the weighted-sum formulation [19]. The minimax formulation in (3.15) has better characteristics than the weighted-sum formulation (3.17).

Remark 3.2. Since the transients of responses with spectral abscissa (3.9) near 0 are unsuitable for FDI, the constraint of stability for multiple models (3.16) could be replaced by the spectral abscissa of $A^i - LC^i$ (3.9) to improve the transients of residual responses. However, the corresponding design region of observer gain L will be more strict than the previous one. A smaller spectral abscissa of $A^i - LC^i$ will have a smaller design region to select the observer gain L.

Normally, the regions for the observer gain L to stabilize the observer for different models are different, therefore, the unique observer gain may not exist in the case of no intersection of the feasible region for different models. This section discusses the case where the intersection of the feasible region for different models exists, which means that there exists a unique observer gain L, which stabilizes all the models at the same time. The objective of the proposed optimization with (3.15) and (3.16) could be interpreted by Fig. 3.2. As introduced in [36], a smaller $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ means a larger set of strongly detectable faults. Assuming that the points A, B and C in Fig. 3.2 are the optimal designs with the H_-/H_∞ mixed criterion for Model 1, 2 and 3 respectively, which provide largest sets of strongly detectable faults for the corresponding model. However, the design A may not give large sets of strongly detectable faults or even no sets of strongly detectable faults (if A is out of the feasible region of corresponding models) for the other models, like Model 2 and Model 3. The same conditions may happen for the design B and C. The objective of the optimization in (3.15) and (3.16) is to compromise the sets of strongly detectable faults for the three models: decrease the redundant sets of strongly detectable faults of some models to compensate the small sets of strongly detectable faults for other models. As shown in Fig. 3.2, the integrated design D will give compromised sets for all three models separately.

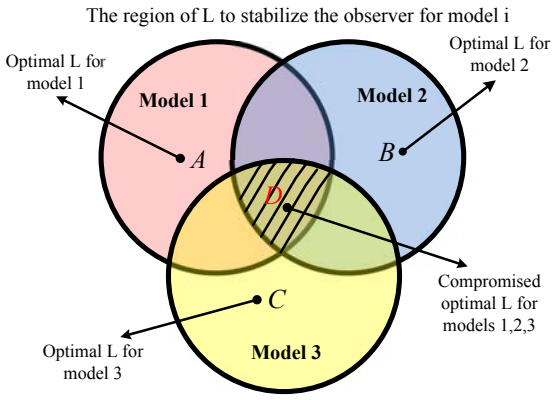

Figure 3.2.: Interpretation of the formulation (3.15) and (3.16)

Obviously, different Pareto results will be given by the proposed objective function (3.15) with different selection of the weights λ_i. One practical background of the weights λ_i is that they are meant to express the importance of different models: for example, a system contains three models. In 80% process duration, the system is in the first model, and the other two models will take up 10% duration of the process respectively. Then, the importance of the different models could be represented by the time duration of the different models. In this case, the weights could be set as $\lambda_1 = 0.8$, $\lambda_2 = 0.1$ and $\lambda_3 = 0.1$. From a theoretical point of view, to search the Pareto optimal surface for an appropriate design, some other general design procedures to choose the initial designs and design iterations for the weights λ_i are recommended in [19]. The objective of this part does not focus on the selection of the weights λ_i, but tries to show the effectiveness of the nonsmooth optimization approach on the design for multiple models case with some selected weights λ_i. Typically, as introduced in [19], one method for selecting the weights is to scale every $\|G^i_{rd}(L,Q)\|_\infty / \left\|G^i_{rf}(L,Q)\right\|_-$ by a typical or nominal value, for example, if the weights λ_i are selected as the reciprocal of the best value of $\|G^i_{rd}(L,Q)\|_\infty / \left\|G^i_{rf}(L,Q)\right\|_-$, the objective function in (3.15) could be interpreted as a balanced design, which considers the influences of all the models simultaneously with the performance indexes $\|G^i_{rd}(L,Q)\|_\infty / \left\|G^i_{rf}(L,Q)\right\|_-$. The following part will consider this setting for the weights λ_i in (3.15) to design the unique observer gain for multiple models. The details about selection of the weights λ_i and the integrated design are shown in Algorithm 3.1.

Algorithm 3.1 Procedure to design for multiple models case

Step 1. Design the observer gain L_i and residual weighting matrix Q_i individually ($i \in \{1,\ldots,N\}$) with objective function (3.13) and the constraints of stabilization for all the models:

$$\min \quad \left(\frac{\|G_{rd}^i(L_i,Q_i)\|_\infty}{\left\|G_{rf}^i(L_i,Q_i)\right\|_-} \right) \tag{3.18}$$
$$A^j - L_iC^j \text{ is asymptotically stable, } j \in \{1,\ldots,N\}$$

Step 2. Check the spectral abscissa of $A^j - L_iC^j$ ($i,\ j \in \{1,\ldots,N\}$). If the minimum value of $\alpha\left(A^j - L_iC^j\right)$ for $i,\ j \in \{1,\ldots,N\}$ is too small or not appropriate enough with some L_i, note $p = 1$. Then, a constraint of specification (3.9) should be added into the optimization (3.18) to replace the condition of asymptotic stability to obtain a series of L_i and Q_i:

$$\min \quad \left(\frac{\|G_{rd}^i(L_i,Q_i)\|_\infty}{\left\|G_{rf}^i(L_i,Q_i)\right\|_-} \right) \tag{3.19}$$
$$\alpha\left(A^j - L_iC^j\right) < \varsigma,\ j \in \{1,\ldots,N\}$$

where the negative parameter ς is chosen with the requirement of residual dynamics. Otherwise, $p = 0$.

Step 3. Select the weights λ_i to be the reciprocal of the best nominal value of $\|G_{rd}^i(L_i,Q_i)\|_\infty / \left\|G_{rf}^i(L_i,Q_i)\right\|_-$ for different models with the series of L_i and Q_i;

Step 4. If $p = 0$, optimize the criterion (3.15) under the constraints of (3.16) by using nonsmooth optimization method. Otherwise, optimize with the transient specification (3.9):

$$\min_{L,Q} \max_{i=1,\ldots,N} \left(\lambda_i \frac{\|G_{rd}^i(L,Q)\|_\infty}{\left\|G_{rf}^i(L,Q)\right\|_-} \right) \tag{3.20}$$
$$\alpha\left(A^i - LC^i\right) < \varsigma,\ i \in \{1,\ldots,N\} \tag{3.21}$$

Remark 3.3. Since the design region for the multiple models case is the intersection of stability region for multiple models, the mixed criteria H_-/H_∞ may be cut by the bound of the intersection region for some models. As a result, the optimal design of H_-/H_∞ in Step 1 for some model is in the bound of the stability region, which means that the corresponding spectral abscissa will be near zero. Therefore, in this case, the Step 2 in Algorithm 3.1 has to be introduced. However, the constraint of transient specification (3.9) in Step 2 of Algorithm 3.1, which is used to replace the asymptotically stability, will decrease the design region to choose the unique observer gain L. A too small ξ will introduce conservativeness into the design. The design in Step 1 could give an estimation to choose the parameter ς:

$$\min_{i,\ j\in\{1,\ldots,N\}} \alpha\left(A^j - L_iC^j\right) < \varsigma$$

3.2.3. Residual evaluation and threshold design

After residual generation, the next step is to select the evaluation function and the threshold to detect faults based on the generated residual signal. In the steady state of residual after switching, if the perfect decoupling conditions (3.5) and (3.6) could be achieved, the thresholds will be zero in this case. Otherwise, thresholds should be considered to filter the effects of the unknown inputs, like disturbances. Since it is difficult to achieve perfect decoupling in the face of limitations of measurement information and unstructured uncertainties, robust residual evaluation, which can achieve small false alarm rate with a guaranteed sensitivity to faults, should be considered. One way of robust residual evaluation is adaptive threshold, which will be used in this chapter. The problem of a fixed threshold is that a too low threshold will increase the rate of false alarms while conservativeness will be introduced if the threshold is set too high. Therefore, it is better to design a threshold, which depends upon the uncertainties of the system and adapts to the system input.

Among a great deal of residual evaluation functions, the time window Root Mean Square (RMS) is a convenient residual measure in practice [24, 23, 40, 60]:

$$\|r\|_{\text{rms}} := \left(\frac{1}{T} \int_{t-T}^{t} r^T(\tau)\, r(\tau)\, d\tau \right)^{\frac{1}{2}} \tag{3.22}$$

where T denotes the finite time window. Since it is impractical to evaluate the residual signal over the whole time range, the windowed RMS evaluation method is proposed in practice to detect fault as early as possible.

The RMS evaluation function is used here for the following reasons [40, 4, 35]:

1. It is widely used in practice.
2. Through the average calculation of the instantaneous value of the residual over the time window T, the effects of disturbances are reduced.
3. The residual signal over the moving time window is smooth with the RMS evaluation.

Remark 3.4. In general, an evaluated residual with a big time window T will be more smooth and harder to be affected by the disturbances, but introduces a big delay to detect faults. By contrast, a small time window T results in a small delay to detect faults, but gives an evaluated residual with more oscillations because of the disturbances. In the view of optimal design, the evaluation part should be designed with the criteria used to generate residual simultaneously. In some typical cases, the selection of time window T will also affect the ability of fault detection. For example, if the frequency ω_1 (which makes $|G_{rd}(j\omega_1)| = \|G_{rd}\|_\infty = \gamma_d$) is the same as the frequency ω_0 (which makes $|G_{rf}(j\omega_0)| = \|G_{rf}\|_- = \beta_f$), the mixed criteria (3.49) has to work with a bandpass filter $Q(s)$, which maximizes the difference between $\|G_{rf}\|_-$ and $\|G_{rd}\|_\infty$ at the same frequency point $\omega_1 = \omega_0$. The residual r will be a kind of sinusoidal signal because of the filtering by the bandpass filter $Q(s)$. In this case, the finite time window T could be set as $T = \frac{2\pi}{\omega_0}$, as a consequence, the evaluated residual $\|r\|_{\text{rms}}$ will be constant if the amplitude of the sinusoidal signal is constant. In the view of FDI, this constant evaluated residual has better characteristics of fault detection for sinusoidal fault signal. Since this chapter focuses on the general case with the mixed specifications H_-/H_∞ (ω_1 may not be equivalent to ω_0), the selection of finite time window T is out of this thesis.

One objective of threshold selection is to reduce or avoid false alarms. Therefore, the evaluated residual $\|r\|_{\text{rms}}$ should be under the threshold value J_{th}^i when the system is in i^{th} operational model:

$$J_{th}^i = \sup_{f=0} \|r\|_{\text{rms}} \tag{3.23}$$

In the fault free case, the residual (3.4) changes to be

$$r = \mathfrak{L}^{-1}\left(G_{rd}^i d\right) + \Psi_i\left(x\left(0\right)\right) \tag{3.24}$$

Therefore, the threshold in (3.23) depends on the part $\Psi_i\left(x\left(0\right)\right)$. In order to detect fault during the switching, the initial state $x\left(0\right)$ should be estimated. If there is no information about initial state $x\left(0\right)$, the threshold with (3.23) and (3.24) cannot be exactly selected. An alternative way is to detect fault during steady state after switching. In the steady state after switching, the residual is

$$r = \mathfrak{L}^{-1}\left(G_{rd}^i d\right) \tag{3.25}$$

According to the Parseval Theorem [19, 127] and RMS norm relationship [19]

$$\begin{aligned} \|r\|_{\text{rms},\ f=0} &= \left\|\mathfrak{L}^{-1}\left(G_{rd}^i d\right)\right\|_{\text{rms}} \leqslant \left\|G_{rd}^i\right\|_\infty \|d\|_{\text{rms}} \\ &= \gamma_d^i \|d\|_{\text{rms}} \leqslant \gamma_d^i \cdot \max\left(\|d\|_{\text{rms}}\right) \end{aligned} \tag{3.26}$$

where γ_d^i is the value of $\|G_{rd}^i(L,Q)\|_\infty$ with the optimized L and Q from the objective function (3.15). $\max\left(\|d\|_{rms}\right)$ is a convenient upper bound to the rms-norm with the worst disturbances acting on the plant, which could be calculated off-line [24, 23, 60].

Considering to (3.23) and (3.26), the thresholds for multiple models could be:

$$J_{th}^i = \gamma_d^i \cdot \max\left(\|d\|_{\text{rms}}\right) \tag{3.27}$$

Then, by comparing the evaluated residual $\|r\|_{\text{rms}}$ with the proposed threshold J_{th}^i, the fault could be detected if fault makes the evaluated residual $\|r\|_{\text{rms}}$ exceed the threshold J_{th}^i. Following logic decision unit is considered to detect faults:

$$\begin{cases} \|r\|_{\text{rms}} > J_{th}^i & \text{Fault is detected (Alarm)} \\ \|r\|_{\text{rms}} \leq J_{th}^i & \text{Fault free (No Alarm)} \end{cases} \tag{3.28}$$

Remark 3.5. The threshold in (3.27) only estimates the possible maximum effects from the disturbances on the residual. If the frequency of the disturbances (which achieves $\max\left(\|d\|_{\text{rms}}\right)$) is different from the frequency ω_1 (which makes $G_{rd}^i\left(j\omega_1\right) = \|G_{rd}^i\|_\infty = \gamma_d^i$), the parameters γ_d^i and $\max\left(\|d\|_{rms}\right)$ cannot be achieved simultaneously by any disturbance. As a consequence, the threshold J_{th} will be higher than the peak value of the evaluated residual $\|r\|_{\text{rms}}$.

Remark 3.6. As shown in (3.4), during the switching of the system and observer, the initial state $x\left(0\right)$ of switching is not zero. (3.1) and (3.2) use common state vector x and $\hat{x}$ with changing A^i, B_u^i, C^i and D_u^i, which will not jump after the switching. Because the threshold J_{th}^i are designed in the steady state of the residual, the natural response $\Psi\left(x\left(0\right)\right)$ caused by

the initial state $x(0)$ when switching arises will affect the decision of fault detection during the switching. Because threshold J_{th}^i is also dependent on the switching signal i, there will be some false alarms in some typical cases:

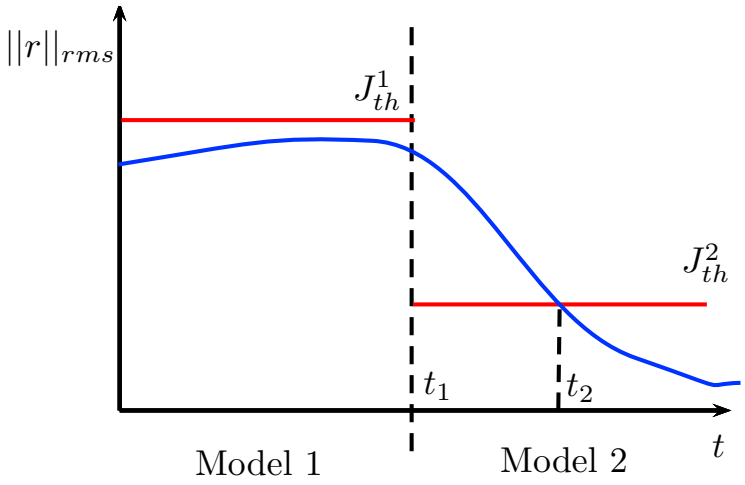

Figure 3.3.: False alarms during the switching

Fig. 3.3 shows a switching at t_1 from Model 1 to Model 2. J_{th}^1 and J_{th}^2 ($J_{th}^1 > J_{th}^2$) are the thresholds for Model 1 and Model 2 respectively. The natural response $\Psi_2(x(0))$ at t_1 will make the evaluated residual $\|r\|_{\text{rms}}$ be higher than the threshold J_{th}^2 even the system is in Model 2. According to the design logic in (3.28), there will be a false alarm from t_1 to t_2. In order to avoid this kind of false alarms, on the one hand, the specification (3.9) could be optimized to improve the rapidity of convergency, on the other hand, the fault detection could be waited for enough long time until there is no alarm after the switching. Therefore, in practice, both aspects should be considered together to decrease false alarms rate with framework of fault detection in Fig. 3.1.

Remark 3.7. As introduced in chapter 2, the minimization of mixed specification H_-/H_∞ in (2.24) can be explained as maximizing the set of strongly detectable faults. With a fixed energy bound disturbances, the effects of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ will be shown on the lower bound of strongly detectable faults. A design with smaller value of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ has a smaller lower bound of the strongly detectable faults, which is validated with the following example:

Example 3.1. Considering a numerical example:

$$A = \begin{bmatrix} -0.9231 & 0.5422 \\ -0.9442 & -0.6764 \end{bmatrix}, \; C = \begin{bmatrix} 0.5432 & 0.4595 \end{bmatrix}$$
$$B_f = \begin{bmatrix} 0.4141 \\ -0.3287 \end{bmatrix}, \; B_d = \begin{bmatrix} 0.2093 \\ 0.1224 \end{bmatrix}$$
$$D_f = 0.7525, \; B_d = 0.0834$$

We have two solutions: $L_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ with $\phi_0 = \frac{\|G_{rd}(L_0)\|_\infty}{\|G_{rf}(L_0)\|_-} = \frac{0.1973}{0.5204} = 0.3791$ and $L_1 = \begin{bmatrix} 0.6557 \\ -0.2478 \end{bmatrix}$ with $\phi_1 = \frac{\|G_{rd}(L_1)\|_\infty}{\|G_{rf}(L_1)\|_-} = \frac{0.1870}{0.6465} = 0.2892$, where the frequency ω meeting $\|G_{rf}\|_- =$

$|G_{rf}(j\omega)|$ is at the low frequency for both solutions $\omega = 0.001 rad/s$. The bode diagrams of two observer gains are shown in Fig. 3.4.

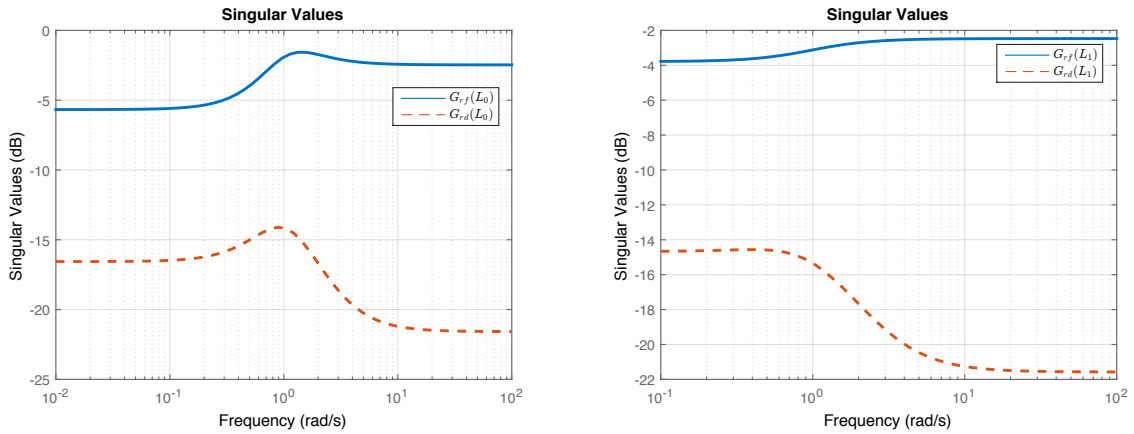

Figure 3.4.: Singular values of G_{rd} and G_{rf} for L_0 and L_1

If the fault f is a sinusoidal signal, the lower bound of strongly detectable faults could be represented by the smallest amplitude of the sinusoidal signal when the sinusoidal signal is strongly detectable. Assuming $\max \|d\|_{\text{rms}} = 1$, then the threshold J_{th} for the design L_0 is 0.1973 while the threshold J_{th} for the design L_1 is 0.1870. The strongly detectable faults could be represented as $\|\mathfrak{L}^{-1}(G_{rf}f)\|_{\text{rms}} > 2J_{th}$. Then, we can find the smallest amplitude (AMP) of the sinusoidal signal $sin(0.001t)$ in the worst case, which meets $\|\mathfrak{L}^{-1}(G_{rf}f)\|_{\text{rms}} > 2J_{th}$, for the design with L_0 is $AMP_0 = 0.75826$ and the design with L_1 is $AMP_1 = 0.5785$. Notice that

$$\phi_1/\phi_0 = AMP_1/AMP_0$$

we have

$$AMP_1 = AMP_0 \frac{\phi_1}{\phi_0}$$

which means that the lowest amplitude of sinusoidal signal is proportional to the ratio of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$.

3.3. Robust fault detection filter design

With the formulation of fault detection observer for the multiple models case, the parameters (except the unique observer gain L and residual weighting matrix Q) in the fault detection observer and the thresholds have to change with the switching of models (parameter i). However, from a practical point of view, in some cases, you know that the system (3.1) is in one of N model, but you don't know the system is in which model exactly. In this case, it is necessary to design a fault detection "observer" and some threshold, which could work without model information i. This part introduces a new framework to generate residual, and then designs a unique threshold to achieve fault detection. Both residual generation and threshold are independent upon the model information i.

3.3.1. Residual generation

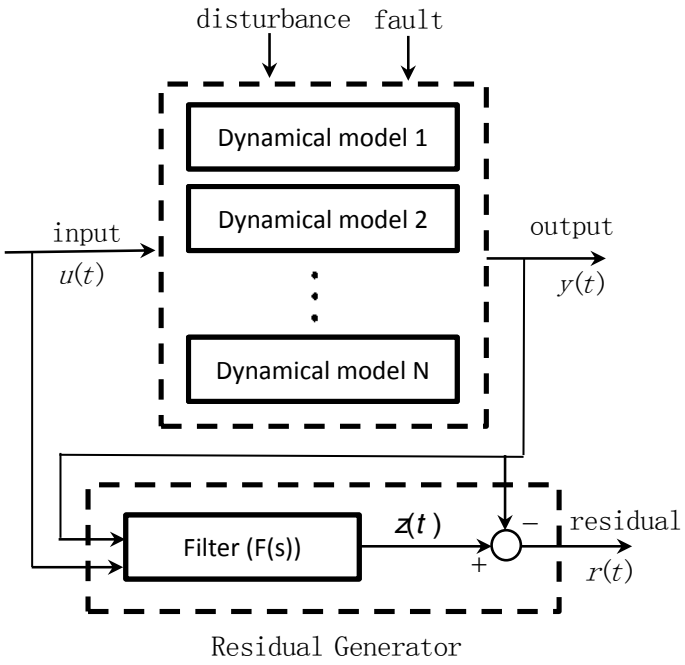

Figure 3.5.: Robust residual generator

With the same plant models in (3.1), a framework of robust residual generator is illustrated in Fig. 3.5. Different from the previous case, there is no switching in the system. The system is in one of the N dynamical models, which is unknown. Comparing with the residual generator in Fig. 3.1, the inputs of the filter F in Fig. 3.5 are inputs u and outputs y without the model information i. Therefore, in this case, the new framework with filter F is robust to the model information i.

The form of robust fault detection filter is introduced in [23, 24, 60, 100]:

$$F : \begin{cases} \dot{x}_F = A_F x_F + B_F \begin{bmatrix} y \\ u \end{bmatrix} \\ z = C_F x_F + D_F \begin{bmatrix} y \\ u \end{bmatrix} \end{cases} \tag{3.29}$$

where $x_F \in \mathbb{R}^{n_F}$, $A_F \in \mathbb{R}^{n_F \times n_F}$, $B_F \in \mathbb{R}^{n_F \times (n_y+n_u)}$, $C_F \in \mathbb{R}^{n_y \times n_F}$, $D_F \in \mathbb{R}^{n_y \times (n_y+n_u)}$. In the full order case [24, 23], the filter dimension (the dimension of A_F) satisfies $n_F = n_x + n_u$. In the view of transfer function for (3.29), the filter will be

$$F = C_F(sI - A_F)^{-1}B_F + D_F \tag{3.30}$$

The residual r is defined as

$$r = z - y \tag{3.31}$$

The dynamic equations of residual r are:

$$\begin{cases} \dot{x}_r = A_r^i x_r + B_{r,u}^i u + B_{r,f}^i f + B_{r,d}^i d \\ r = C_r^i x_r + D_{r,u}^i u + D_{r,f}^i f + D_{r,d}^i d \end{cases} \tag{3.32}$$

$$\begin{aligned}
x_r &= \begin{bmatrix} x \\ x_F \end{bmatrix} \\
A_r^i &= \begin{bmatrix} A^i & 0 \\ B_F \begin{bmatrix} C^i \\ 0 \end{bmatrix} & A_F \end{bmatrix}, \\
C_r^i &= \begin{bmatrix} D_F \begin{bmatrix} C^i \\ 0 \end{bmatrix} - C^i & C_F \end{bmatrix} \\
B_{r,u}^i &= \begin{bmatrix} B_u^i \\ B_F \begin{bmatrix} D_u^i \\ I \end{bmatrix} \end{bmatrix}, D_{r,u}^i = \left[D_F \begin{bmatrix} D_u^i \\ I \end{bmatrix} - D_u^i \right] \\
B_{r,f}^i &= \begin{bmatrix} B_f^i \\ B_F \begin{bmatrix} D_f^i \\ 0 \end{bmatrix} \end{bmatrix}, D_{r,f}^i = \left[D_F \begin{bmatrix} D_f^i \\ 0 \end{bmatrix} - D_f^i \right] \\
B_{r,d}^i &= \begin{bmatrix} B_d^i \\ B_F \begin{bmatrix} D_d^i \\ 0 \end{bmatrix} \end{bmatrix}, D_{r,d}^i = \left[D_F \begin{bmatrix} D_d^i \\ 0 \end{bmatrix} - D_d^i \right]
\end{aligned}$$

In the view of transfer matrices, residual r is given by (assuming that the initial state is zero)

$$\begin{aligned}
r &= z - y \qquad (3.33) \\
&= C_F x_F + D_F \begin{bmatrix} y \\ u \end{bmatrix} - \left\{ C^i x + D_u^i u + D_f^i f + D_d^i d \right\} \\
&= \mathfrak{L}^{-1} \Bigg(C_F (sI - A_F)^{-1} B_F \begin{bmatrix} y \\ u \end{bmatrix} + D_F \begin{bmatrix} y \\ u \end{bmatrix} - \left\{ C^i (sI - A^i)^{-1} [B_u^i u + B_f^i f + B_d^i d] \right. \\
&\quad \left. + D_u^i u + D_f^i f + D_d^i d \right\} \Bigg) \\
&= \mathfrak{L}^{-1} \left([C_F (sI - A_F)^{-1} B_F + D_F] \begin{bmatrix} y \\ u \end{bmatrix} - \left\{ G_{yu}^i u + G_{yf}^i f + G_{yd}^i d \right\} \right) \\
&= \mathfrak{L}^{-1} \left(F \begin{bmatrix} G_{yu}^i u + G_{yf}^i f + G_{yd}^i d \\ u \end{bmatrix} - \left\{ G_{yu}^i u + G_{yf}^i f + G_{yd}^i d \right\} \right) \\
&= \mathfrak{L}^{-1} \left(\left\{ F \begin{bmatrix} G_{yf}^i \\ 0 \end{bmatrix} - G_{yf}^i \right\} f + \left\{ F \begin{bmatrix} G_{yu}^i \\ I \end{bmatrix} - G_{yu}^i \right\} u + \left\{ F \begin{bmatrix} G_{yd}^i \\ 0 \end{bmatrix} - G_{yd}^i \right\} d \right) \\
&= \mathfrak{L}^{-1} \left(G_{rf}^i f + G_{ru}^i u + G_{rd}^i d \right) \qquad (3.34)
\end{aligned}$$

The transfer matrices, G^i_{rf}, G^i_{ru} and G^i_{rd}, are defined as follows:

$$G^i_{rf} = F\begin{bmatrix} G^i_{yf} \\ 0 \end{bmatrix} - G^i_{yf} \tag{3.35}$$

$$G^i_{ru} = F\begin{bmatrix} G^i_{yu} \\ I \end{bmatrix} - G^i_{yu} \tag{3.36}$$

$$G^i_{rd} = F\begin{bmatrix} G^i_{yd} \\ 0 \end{bmatrix} - G^i_{yd} \tag{3.37}$$

where G^i_{rf} means the transfer function from fault to the residual, G^i_{rd} represents the transfer function from disturbance to the residual and G^i_{ru} is the transfer function from control input to the residual.

Obviously, equation (3.34) shows that the dynamics of residual vector rely on faults signals, disturbances and the control inputs, which also depend on the model information $i \in \{1, 2, \ldots, N\}$. Since the model information i is unknown, all the $i \in \{1, 2, \ldots, N\}$ model should be considered into the design. With the aim of residual generation for fault detection, the robust fault detection filter (RFDF) design problem could be described as designing a filter $F(A_F, B_F, C_F, D_F)$ to make the residual r be as sensitive as possible to the faults f and simultaneous as robust as possible to the disturbances d and control input u for all models. The functions of different parts in (3.34) could be explained as: the part G^i_{rf} means the ability of fault sensitivity, the part G^i_{ru} evaluates the performance of control input tracking on the residual, while the part G^i_{rd} represents the capability of disturbances rejection.

Remark 3.8. From the aspect of residual generation, different from the classical methods (like observer based method and parity space method), robust fault detection filter does not decouple the control input u from residual r, but removes the model information i from the filter F (3.29), which gives the possibility to detect faults for the system in (3.1) even without the model information i. When $N = 1$, the classical methods could decouple the effects of the control input u from the residual r by subtracting $G^i_{ru}u$ from the residual r (3.34) in the parity space method or by the structure of the observer in the observer based method. When $N > 1$, the effects of $G^i_{ru}u$ also could be decoupled from the residual by the similar way (observer based method or parity space method) if the model information i is known. However, without the model information i, the part of $G^i_{ru}u$ in the residual is hard to be decoupled accurately if $G^i_{ru} \neq G^j_{ru}$ with $i \neq j$ even when the control inputs u are known. Fortunately, if the control inputs are known, the influences of the control input on results of fault detection could be decreased by an appropriate selection of threshold.

In order to detect faults, the residual should be sensitive to the faults and at same time remain robust to the disturbances, uncertainties and control inputs. The following three conditions should be satisfied if perfect decoupling is achieved:

$$G^i_{rf} f \neq 0 \tag{3.38}$$

$$G^i_{rd} d = 0 \tag{3.39}$$

$$G^i_{ru} u = 0 \tag{3.40}$$

Remark 3.9. Comparing with the perfect decoupling conditions of observer based formulation,

the formulation of robust fault detection needs another decoupling condition for control input u. When $N = 1$, the condition (3.40) is not necessary if the control inputs u are known, because any deviations away from the known dynamics of G_{ru} could be used to detect fault with the conditions (3.38) and (3.39). However, when $N > 1$, the dynamics of G^i_{ru} will be difficult to determine if the model information is unknown. As a result, the undefined dynamics of $G^i_{ru}$with the conditions (3.38) and (3.39) will decrease the fault detection rate and cause false alarm. Hence, it is better to add the constraint (3.40) into the conditions of perfect decoupling to decrease the influences of the undefined dynamics of G^i_{ru}.

The constraints (3.38), (3.39) and (3.40) may be satisfied for one model case ($N = 1$) for a filter F. However, when $N > 1$, it is more difficult to find a unique filter F to satisfy the above constraints for different model at the same time. Therefore, the FDI module has to deal with the problem without perfect disturbance decoupling.

3.3.2. Unique threshold design for RFDF

On account of the fact that it is difficult to decouple the disturbance d and control input u from the residual r, it is necessary to definite a threshold to avoid false alarm. The objective of RFDF is to detect fault for N models without model information i, hence, the designed threshold should be also robust to the model information i. In other words, there is only one threshold in the design. Owing to this objective, the following part describes the procedures to design the unique threshold for the system with N models.

In the similar way of 3.2.3, the time window RMS (3.22) is considered to evaluate residuals here. Different from the case in (3.24) under fault free conditions, the residual is described as

$$r = G^i_{rd} d + G^i_{ru} u \tag{3.41}$$

which means that the residual in the fault free conditions are affected not only by disturbances d, but also by the control inputs u of systems and model information i.

According to the Parseval Theorem [19], the unique threshold for the system with $N > 1$ models could be:

$$\begin{aligned}
J_{th} &= \sup_{f=0} \|r\|_{\text{rms}} \\
&= \left\| \mathfrak{L}^{-1} \left(G^i_{ru} u + G^i_{rd} d \right) \right\|_{\text{rms}} \\
&\leqslant \left\| G^i_{ru} \right\|_\infty \|u\|_{\text{rms}} + \left\| G^i_{rd} \right\|_\infty \|d\|_{\text{rms}} \\
&\leqslant \max \left(\left\| G^i_{ru} \right\|_\infty \right) \|u\|_{\text{rms}} + \max \left(\left\| G^i_{rd} \right\|_\infty \right) \|d\|_{\text{rms}} \\
&\leqslant \max \left(\left\| G^i_{ru} \right\|_\infty \right) \|u\|_{\text{rms}} + \max \left(\left\| G^i_{rd} \right\|_\infty \right) \max \left(\|d\|_{\text{rms}} \right) \\
&= \gamma_u^{max} \|u\|_{\text{rms}} + \gamma_d^{max} \max \left(\|d\|_{\text{rms}} \right)
\end{aligned} \tag{3.42}$$

where $\gamma_u^{max} = \max\limits_{i=1,\ldots,N} \{\|G^i_{ru}\|_\infty\}$ and $\gamma_d^{max} = \max\limits_{i=1,\ldots,N} \{\|G^i_{rd}\|_\infty\}$. $\max \left(\|d\|_{\text{rms}} \right)$ is a convenient upper bound to the rms-norm with the worst disturbances acting on the plant, which could be considered to be calculated off-line [24, 23, 60]. $\|u\|_{\text{rms}}$ is calculated on-line.

3.3.3. Robust residual generation with unique filter

After the definition of the unique threshold, the following part tries to generate residual with a unique filter F (3.29), which is also robust to the model information i.

Because of the difficulty to achieve perfect disturbance and control input decoupling, a satisfactory performance of residual should consider the trade-off between the maximization of the effects of faults on the residual and the maximization of the robustness to the disturbances and control inputs. In fact, the following objectives of the robust residual are conflicting each other:

$$\max_F \left\|G_{rf}^i\right\| \tag{3.43}$$

$$\min_F \left\|G_{rd}^i\right\| \tag{3.44}$$

$$\min_F \left\|G_{ru}^i\right\| \tag{3.45}$$

The residual generation problem is to design a filter F in the form of (3.29) with the objectives in (3.43), (3.44) and (3.45), which is a multiobjective optimization problem.

With the application of H_∞ norm and H_- index in the previous part,it leads to consider the H_∞ norm for the minimization of $\|G_{rd}^i\|$ and $\|G_{ru}^i\|$, and H_- index for the maximization of $\left\|G_{rf}^i\right\|$. Then, the following conflicting specifications should be considered:

$$\max_F \left\|G_{rf}^i\right\|_- \tag{3.46}$$

$$\min_F \left\|G_{rd}^i\right\|_\infty \tag{3.47}$$

$$\min_F \left\|G_{ru}^i\right\|_\infty \tag{3.48}$$

Definition 3.1. When $N = 1$, given positive reals a, b and c, the robust fault detection filter (RFDF) design problem using multiobjective H_∞/H_- optimization is to find a filter realization F such that:

$$\begin{gathered}\min_F \ (a\gamma_d + b\gamma_u)/c\gamma_f \\ \left\|F\begin{bmatrix} G_{yd} \\ 0 \end{bmatrix} - G_{yd}\right\|_\infty \leqslant \gamma_d, \ \gamma_d > 0 \\ \left\|F\begin{bmatrix} G_{yu} \\ I \end{bmatrix} - G_{yu}\right\|_\infty \leqslant \gamma_u, \ \gamma_u > 0 \\ \left\|F\begin{bmatrix} G_{yf} \\ 0 \end{bmatrix} - G_{yf}\right\|_- \geqslant \gamma_f, \ \gamma_f > 0\end{gathered} \tag{3.49}$$

where a, b, and c are used to weight the relative importance of the conflicting requirements in (3.46), (3.47) and (3.48).

Definition 3.2. When $N > 1$, given positive reals a^i, b^i, c^i and λ_i, the RFDF design problem

using multiobjective H_∞/H_- optimization is to find a filter realization F such that:

$$\min_{F} \max_{i=1,\ldots,N} \left(\lambda_i \frac{(a^i \gamma_d^{max} + b^i \gamma_u^{max})}{c^i \gamma_f^i} \right) \tag{3.50}$$
$$\left\| G_{rd}^i \right\|_\infty = \left\| F \begin{bmatrix} G_{yd}^i \\ 0 \end{bmatrix} - G_{yd}^i \right\|_\infty \leqslant \gamma_d^i \leqslant \gamma_d^{max},\ \gamma_d^i, \gamma_d^{max} > 0$$
$$\left\| G_{ru}^i \right\|_\infty = \left\| F \begin{bmatrix} G_{yu}^i \\ I \end{bmatrix} - G_{yu}^i \right\|_\infty \leqslant \gamma_u^i \leqslant \gamma_u^{max},\ \gamma_u^i, \gamma_u^{max} > 0$$
$$\left\| G_{rf}^i \right\|_- = \left\| F \begin{bmatrix} G_{yf}^i \\ 0 \end{bmatrix} - G_{yf}^i \right\|_- \geqslant \gamma_f^i,\ \gamma_f^i > 0$$

where $\gamma_d^{max} = \max_{i=1,\ldots,N} \{\gamma_d^i\}$, $\gamma_u^{max} = \max_{i=1,\ldots,N} \{\gamma_u^i\}$. a^i, b^i, and c^i are used to weight the relative importance of the conflicting requirements (disturbances rejection $\|G_{rd}^i\|_\infty$, input tracking $\|G_{ru}^i\|_\infty$ and fault sensitivity $\left\|G_{rf}^i\right\|_-$) for the i^{th} model. Different from the previous design, the weights λ_i means the probability of the i^{th} model, which the system (3.1) is in. If there is no prior information about this probability, an alternative way is selecting the weights λ_i as the reciprocal of the best value of $\|G_{rd}^i(L,Q)\|_\infty / \left\|G_{rf}^i(L,Q)\right\|_-$, as introduced in the previous part. The objective of the optimization in (3.50) with kind of weights is to reduce the redundant value of $\|G_{rd}^i(L,Q)\|_\infty / \left\|G_{rf}^i(L,Q)\right\|_-$ for some models to compensate the ability of fault detection in the worst for other models. The following part will take this setting for the weights λ_i to show the effectiveness of design.

Remark 3.10. As introduced in [23, 24], control input u is treated as a disturbance in the above formulation. In some typical cases, the fault signal f is added into the control input u ($G_{rf} = G_{ru}$) and their frequency windows may be overlapped. Then, the optimization of objective function (3.49) when $N = 1$ is translated to:

$$\frac{a\|G_{rd}\| + b\|G_{ru}\|}{c\|G_{rf}\|} = \frac{a\|G_{rd}\| + b\|G_{rf}\|}{c\|G_{rf}\|} = \frac{a\|G_{rd}\|}{c\|G_{rf}\|} + b/c \tag{3.51}$$

Consequently, the objective function of the optimization (3.49) is equivalent to the maximization of the fault sensitivity (3.43) and maximization of the disturbance robustness (3.44). In such a case, the selection of the weights b will not affect the optimization. However, when the effects of fault sensitivity and input robustness are measured by H_- index and H_∞ norm as (3.46) and (3.48) respectively:

$$\frac{a\|G_{rd}\|_\infty + b\|G_{ru}\|_\infty}{c\|G_{rf}\|_-} = \frac{a\|G_{rd}\|_\infty + b\|G_{rf}\|_\infty}{c\|G_{rf}\|_-} \tag{3.52}$$

According to the definitions of $\|\cdot\|_\infty$ (A.5) and $\|\cdot\|_-$ (2.18), the frequency range is typically important for the design. If the frequency windows of fault signal f and control input u are overlapped in some part, some part of the frequency range Φ_1 for $\|\cdot\|_\infty$ and the frequency range Φ_2 for $\|\cdot\|_-$ are also overlapped. Consequently, the maximization of the fault sensitivity (3.43) and maximization of the input decoupling (3.45) are not possible to realize simultaneously. It is the reason why [23, 24] proposes that the decoupling of input is not considered by imposing $b = 0$ for this typical case.

As introduced in [23, 24], the full order filter F can give satisfying abilities to detect the faults. Since the reduced order filter has the advantages of faster data processing and a reduction in

detection filter complexity, it is reasonable to design a reduced order filter with an equivalent robustness for the disturbances and sensitivity for the faults to the full order filter [60]. Procedures to design the reduced order RFDF F_{mix} when $N > 1$ are introduced in Algorithm 3.2.

Algorithm 3.2 Procedure to design the reduced order RFDF F_{mix} when $N > 1$

Step 1. Select the interesting frequency ranges Φ for $\|G_{rf}\|_-$, $\|G_{rd}\|_\infty$ and $\|G_{ru}\|_\infty$ (which could be different) and the RFDF dimension (the dimension of A_F) satisfies $n_F = n_x + n_u$, ξ_1 is a small value;

Step 2. Design the robust fault detection filter F separately with the objective function (3.49) to obtain a series of F_i with $i \in \{1, \ldots, N\}$. Note that

$$\vartheta_{ij} = \frac{(a^i \gamma_d^{max} + b^i \gamma_u^{max})}{c^i \gamma_f^i}$$

represents that ϑ_{ij} is obtained with the model (i) and F_j, $i,\ j \in \{1, \ldots, N\}$.

Step 3. Decrease the RFDF dimension $n_F = n_F - 1$, and repeat the Step 2 to obtain a new series of performance index ϑ_{ij}^1;

Step 4. If $\left|\vartheta_{ij} - \vartheta_{ij}^1\right| \leqslant \xi_1$ with $i = j \in \{1, \ldots, N\}$, repeat the Step 3 to obtain a new series of performance index ϑ_{ij}^2. Otherwise, $\vartheta_{ij}^3 = \vartheta_{ij}$ and go to Step 6 ;

Step 5. If $\left|\vartheta_{ij} - \vartheta_{ij}^2\right| \leqslant \xi_1$ with $i = j \in \{1, \ldots, N\}$, go to Step 4. Otherwise, $\vartheta_{ij}^3 = \vartheta_{ij}^1$ and go to Step 6;

Step 6. Select the weights λ_i to be the reciprocal of ϑ_{ij}^3 with $i = j \in \{1, \ldots, N\}$;

Step 7. Design the unique robust fault detection filter F_{mix} with the objective function for (3.50) with the selected weights λ_i.

The point we should note here is that the above threshold J_{th} is also independent on the model information i. In this case, all the processes of the fault detection with the filter F and the threshold J_{th} are robust to the model information i.

3.4. Nonsmooth optimization

Recently, H_∞ synthesis nonsmooth problem with structural constraints [7, 1, 43] could be solved by the solvers based on nonsmooth optimization techniques: Hinfstruct and Systune ([10, 8, 11]). According to the formulations of the criteria introduced in the previous section, both the mixed H_-/H_∞ fault detection observer and robust fault detection filter could be designed with these solvers.

The formulations of the objective functions and constraints for the optimization should be transformed to nonsmooth casts of the following form, which could be solved by the Systune solver:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g(x) \leq c,\ c \in R \end{aligned} \tag{3.53}$$

where both f and g are max-functions

$$f(x) := \max_{i=1,\ldots,N_f} f_i(x), g(x) := \max_{j=1,\ldots,N_g} g_j(x), \tag{3.54}$$

and x gathers all the parameters to design.

In the design of fault detection observer, $f(x)$ could be the ratio objective function (3.13) for the single model case with $N_f = 1$, while for the multiple models case, (3.15) with $N_f = N$ is the function $f(x)$ to minimize. The constraint of asymptotic stability (3.14) could be represented by the hard constraint part $g(x)$ with (3.9)< 0. When the constraint of fast response transients is added into the optimization, the hard constraint $g(x)$ is replaced by the (3.9)$< \xi$.

For the case of robust fault detection filter design, the objective function (3.49) is the soft minimization part $f(x)$ when $N = 1$, while the formulation of maximization for multiple $f_i(x)$ has to be considered for the complex objective function (3.50) when $N > 1$.

3.5. Results

3.5.1. Single model case ($N = 1$)

The single model case is used to show the effectiveness of the proposed nonsmooth optimization techniques and design formulation. For the fault detection observer design with single model case, a comparison between the classical methods and nonsmooth optimization approach is given to. Particularly, the advantages of the proposed nonsmooth optimization approach are outlined. The robust fault detection filter design with single model case is used to illustrate the essences of the proposed filter formulation. A comparison of fault detection capability the between the fault detection observer and robust fault detection filter is also given.

3.5.1.1. Model to consider

A Multiple Input Multiple Output (MIMO) model in the form of (3.1) from [116] as follows is considered:

$$A = \begin{bmatrix} -0.1210 & -0.5624 & -0.2651 & -0.2500 \\ 4.0000 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 0.2500 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} -1.4140 & -0.4373 & -0.1768 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}^T \quad D_f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$B_d = \begin{bmatrix} 0.02 & -0.02 & 0 \\ 0.02 & 0.1 & 0 \\ 0.02 & -0.02 & 0 \\ 0.02 & 0.1 & 0 \end{bmatrix} \quad D_d = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This MIMO model is used to show the effects of fault detection observer design and robust filter design for the single model case.

3.5.1.2. Fault detection observer design

The H_-/H_∞ optimal observer is designed by two classical methods and the proposed nonsmooth optimization method to validate the effectiveness of the proposed method: (a) the inner-outer factorization method of [36], (b) the ILMI method of [116] and (c) the proposed nonsmooth optimization method with (3.13) and (3.14) .

Solving the algebraic Riccati equation (ARE) in [36]

$$\widetilde{A}^T Y + Y\widetilde{A} - YC^T P^{-1} CY + B_d \left(I - D_d^T P^{-1} D_d\right) B_d^T = 0$$

where $P = D_d D_d^T$ and $\widetilde{A} = A - B_d D_d^T P^{-1} C$, an observer gain L and a residual weighting matrix Q could be obtained:

$$L_{Ding} = \left(B_d D_d^T + YC^T\right) P^{-1} = \begin{bmatrix} -0.0201 & 0.0981 & -0.0202 & 0.1016 \\ -0.0001 & -0.0034 & 0 & 0.0089 \end{bmatrix}^T$$

$$Q_{Ding} = P^{-1/2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

With the ILMI algorithm in [116], an observer gain L and a residual weighting matrix Q are obtained:

$$L_{Wang} = \begin{bmatrix} -0.0209 & 0.0916 & -0.0161 & 0.1036 \\ -0.0045 & 0.0227 & -0.0406 & 0.0520 \end{bmatrix}^T$$

$$Q_{Wang} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In this chapter, the residual weighting matrix Q is set as a static matrix to compare the performance of the result by proposed nonsmooth optimization method with these two classical methods. Whereas, the residual weighting matrix Q also could be a dynamic matrix as $Q(s)$ in [27, 119], which will increase the freedoms to design. Through the optimization by nonsmooth optimization method, we can get

$$L_{nonsmooth} = \begin{bmatrix} 0.0513 & 0.2324 & 0.1541 & -0.0586 \\ -0.1532 & 0.2302 & 0.7581 & 0.0988 \end{bmatrix}^T$$

$$Q_{nonsmooth} = \begin{bmatrix} -1.4814 & -0.7006 \\ -1.2189 & 1.0062 \end{bmatrix}$$

Fig. 3.6 shows the singular values plots for the three methods. The minimum singular value of G_{rf} for different methods is at the same frequency point.

Tab. 3.1 lists the effects of different methods with the mixed specifications $\|G_{rd}\|_\infty / \|G_{rf}\|_-$, from which, the design by Ding's method is optimal, while the designs by Wang and Nonsmooth

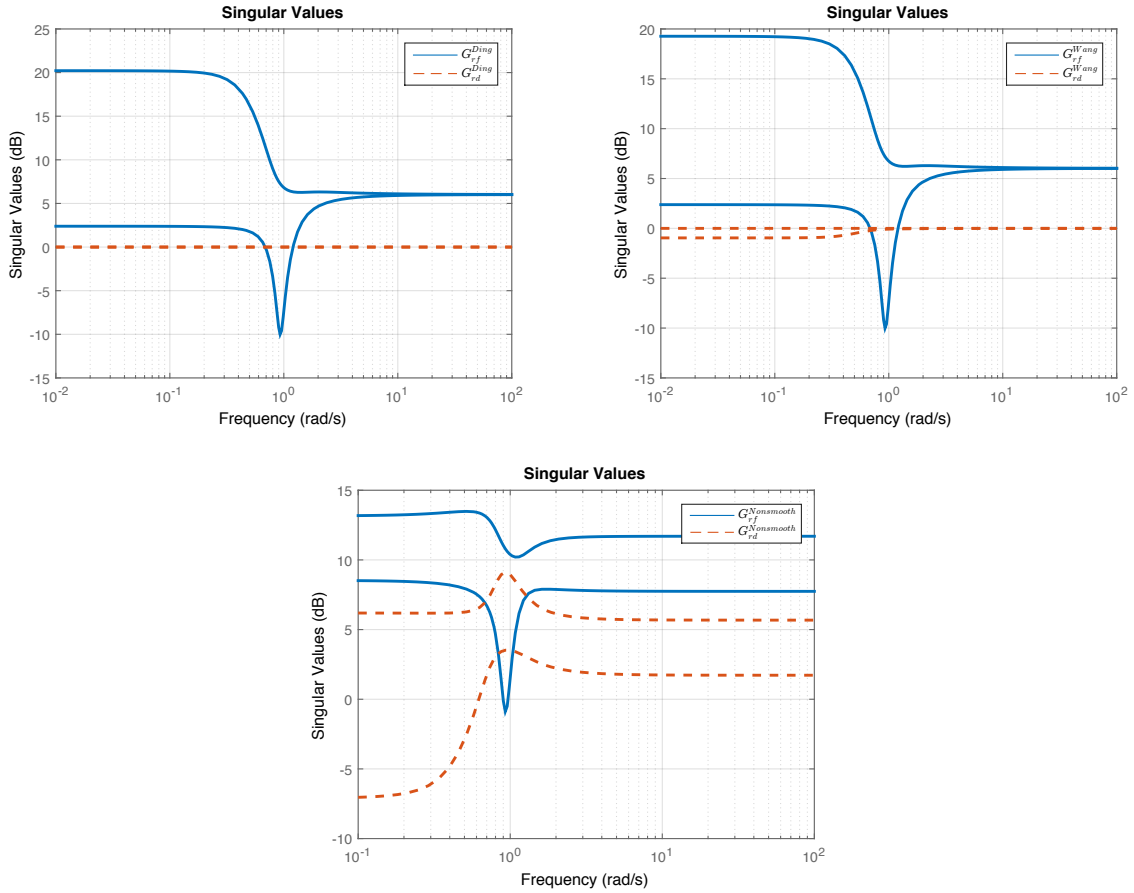

Figure 3.6.: Singular values of G_{rf} and G_{rd} for different designs

are suboptimal solutions, which show a comparable abilities of fault sensitivity and disturbances rejection in the view of ratio specification (3.13).

Table 3.1.: Comparison of results for different methods

Method	$\|G_{rf}\|_-$	$\|G_{rd}\|_\infty$	$\|G_{rd}\|_\infty / \|G_{rf}\|_-$
Ding	0.3154	1.0000	3.1706
Wang	0.3153	1.0000	3.1716
Nonsmooth	0.9015	2.8585	3.1708

As introduced in [116], the ILMI costs more computational power than Ding's method. A comparison of computation time for the three methods is shown in Tab. 3.2. The proposed nonsmooth optimization works much faster than Wang's method, even for small scale model (dimension of system is 4).

In order to decrease the fault detection delay, the performance index (3.9) should be considered. A fast robust fault detection observer could be designed by proposed nonsmooth optimization

Table 3.2.: Comparison of computation time for different methods

Method	Time
Ding	0.163s
Wang	91.7s
Nonsmooth	1.3s

for the example given above:

$$L_{fast} = \begin{bmatrix} -0.2896 & 0.9076 & 0.6822 & -1.1426 \\ 0.1885 & -0.6323 & -0.5896 & 0.9882 \end{bmatrix}^T$$

$$Q_{fast} = \begin{bmatrix} 0.4913 & 0.4282 \\ -0.2321 & 1.5970 \end{bmatrix}$$

With L_{fast} and Q_{fast}, the $\|G_{rf}\|_-$ is 0.5521, and the $\|G_{rd}\|_\infty$ is 1.7513. The corresponding ratio of $\|G_{rd}\|_\infty$ to $\|G_{rf}\|_-$ is 3.1721, which means that the above solution under the constraint of spectral abscissa (3.9) has almost the same properties of robustness to the disturbances and sensitivity to the faults in the worst case.

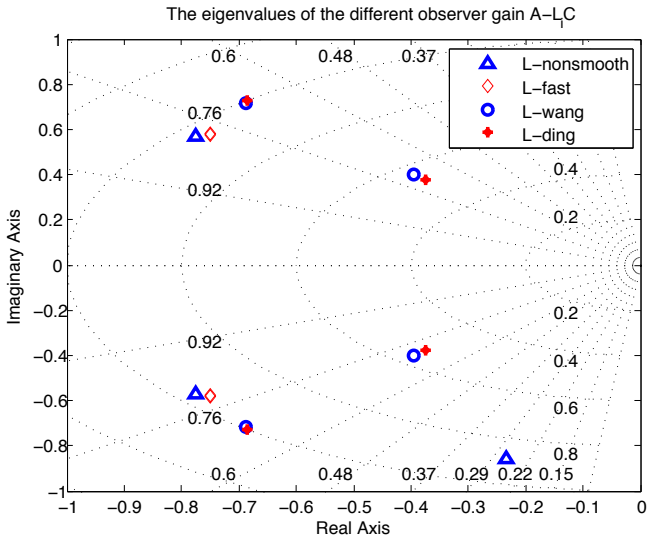

Figure 3.7.: The diagram of the eigenvalues of the different observer gains L

Fig. 3.7 shows the locations of all the eigenvalues of $A - LC$ with the solutions of above introduced methods and the designed L_{fast}. The eigenvalues of $A - L_{fast}C$ are $-0.7503 \pm 0.5810i$ and $-0.7503 \pm 0.5735i$, whose smallest real part are furthest from the imaginary axis comparing with the other observer gains L.

The calculated result shows that, Ding's method has the advantages of fast and easy calculation of the ratio objective function (3.13), however, it is difficult to consider the additional specifications into the design. On the other hand, the ILMI could solve multiobjective design with the ratio criterion (3.13) and transient specification (3.9), but with low efficiency. Comparing with ILMI and Ding's method, the proposed nonsmooth optimization approach not only works

faster for high efficiency calculation, but also can solve the H_-/H_∞ fault detection problem with other additional constraints like eigenvalues.

3.5.1.3. Robust filter design:

This part tries to find the relationships of fault detection abilities (fault sensitivity and disturbance rejection) between fault detection observer and the robust fault detection filter with this single model. In the previous MIMO model, the parameters of inputs $u(t)$ are not given, but this will not affect the following analysis. Decoupling the effects of control inputs u, the fault detection observer only considers the trade-off between fault sensitivity and disturbance rejection. In order to compare the effects of robust fault detection and fault detection observer, only the fault sensitivity and disturbance rejection are considered first. Thus, the weights could be set as $a = 1,\ b = 0,\ c = 1$.

As proposed in [23, 24, 60], the full order ($n_F = 4(\text{plant order}) + 3(\text{reference model order}) = 7$) dynamic filter F (3.29) is designed with the criterion (3.49) by the nonsmooth optimization method:

$$A_F = \begin{bmatrix} -22.3005 & -1.1005 & 69.2127 & 1.0250 & -0.8055 & 15.3692 & -1.8447 \\ 11.2239 & -9.7874 & -37.8748 & 5.2057 & -7.2661 & -0.1119 & -1.4335 \\ -13.5252 & 264.3752 & -10.7090 & -0.0816 & 2.4196 & 18.7329 & 7.0969 \\ 2.4498 & 10.0744 & 3.9286 & -9.2500 & -3.8128 & -3.1133 & -6.8618 \\ 3.0651 & 1.8160 & 5.7046 & -2.2092 & -3.9871 & 10.7254 & -1.7312 \\ 0.8960 & 2.1921 & -3.8481 & -3.8844 & -6.3914 & -9.9065 & -1.7009 \\ 1.8522 & -3.0990 & -5.9203 & 3.3595 & -7.7163 & 4.7342 & -1.9170 \end{bmatrix}$$

$$B_F = \begin{bmatrix} -3.013 & -4.525 & 3.939 & 0.2416 & 0.0178 \\ -4.095 & -3.802 & 11.139 & 1.8829 & 8.1373 \\ -5.791 & -5.904 & 0.634 & 3.9912 & 7.3635 \\ 4.593 & 2.769 & 2.127 & 1.2368 & 0.8320 \\ 5.611 & 6.859 & 0.578 & 53.6836 & 7.9155 \\ 4.542 & -4.934 & 1.697 & 0.2463 & 1.5265 \\ -2.133 & -5.904 & 2.754 & 0.1832 & 1.9534 \end{bmatrix}$$

$$C_F = \begin{bmatrix} -2.8842 & -0.4516 & -2.3262 & 6.0438 & -2.4983 & -0.5313 & 5.1765 \\ -2.1210 & 10.4952 & 1.1555 & 1.9319 & 2.5549 & -0.9477 & 4.5274 \end{bmatrix}$$

$$D_F = \begin{bmatrix} -4.322 & -2.876 & 19.908 & 40.988 & -2.403 \\ 1.405 & -0.895 & -1.132 & -1.662 & -1.303 \end{bmatrix}$$

The effect of this filter for above model is:

$$\left\| F \begin{bmatrix} G_{yf} \\ 0 \end{bmatrix} - G_{yf} \right\|_- / \left\| F \begin{bmatrix} G_{yd} \\ 0 \end{bmatrix} - G_{yd} \right\|_\infty = 0.3156$$

As shown in the previous part, the effects of Luenberger observer (with the observer gain L

and residual weighting matrix Q) in (3.2):

$$\|G_{rf}\|_- / \|G_{rd}\|_\infty = 0.3153$$

Comparing with the effects of the classical Luenberger observer with the mixed specifications $\|G_{rf}\|_- / \|G_{rd}\|_\infty$, the proposed structure of the robust fault detection filter gives equivalent abilities of fault sensitivity and disturbances rejection in the view of ratio. Since the weight b is set as 0, the effects of the control inputs on the residual are not considered in the optimization. It is straightforward to get an idea that the introduction of control inputs rejection ($b \neq 0$) will decrease the ability of fault sensitivity and disturbances rejection (the ratio specification $\|G_{rf}\|_- / \|G_{rd}\|_\infty$) with the objective function (3.49). In other words, the structure of robust fault detection filter sacrifices the ability of fault detection to remove the information of system (e.g. system parameters A^i, B_u^i, C^i and D_u^i) from filter F.

Then, reducing the order of fault detection filter (to 2) with Algorithm 3.2, we can get:

$$A_F^* = \begin{bmatrix} -5.420 & -0.8797 \\ 0.4031 & -0.8820 \end{bmatrix},$$

$$B_F^* = \begin{bmatrix} 0.0503 & -0.2247 & 0 & 0 & 0 \\ 0.1430 & -0.0662 & 0 & 0 & 0 \end{bmatrix}$$

$$C_F^* = \begin{bmatrix} 0.7210 & 0.6362 \\ 0.3157 & 0.2537 \end{bmatrix}$$

$$D_F^* = \begin{bmatrix} 1.1095 & 0.1965 & -0.3949 & -0.3949 & -0.3949 \\ 0.0892 & 1.0140 & -0.1155 & -0.1155 & -0.1155 \end{bmatrix}$$

The value of the optimized criterion is:

$$\left\| F \begin{bmatrix} G_{yf} \\ 0 \end{bmatrix} - G_{yf} \right\|_- / \left\| F \begin{bmatrix} G_{yd} \\ 0 \end{bmatrix} - G_{yd} \right\|_\infty = 0.3154$$

Comparing with the full order RFDF and the classical Luenberger observer, the reduced robust fault detection filter has the same ability to be robust to the disturbances and sensitive to the fault for the worst case in the view of the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$.

3.5.2. Multiple models case

3.5.2.1. Models to consider

A vehicle lateral dynamics system [3, 105] is applied to design a unique robust fault detection observer when $N > 1$. This system is also called one-track model. Sketched in Fig. 3.8, the one-track model is derived with the assumption that the vehicle could only move in x axis, y

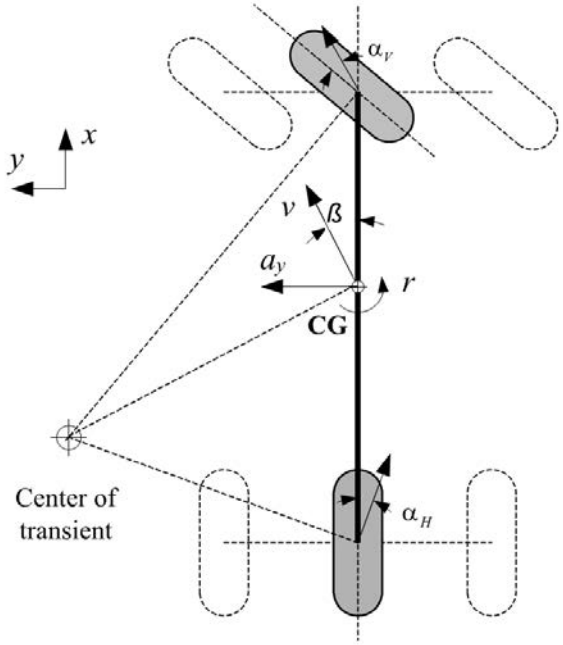

Figure 3.8.: Kinematics of one-track vehicle lateral dynamics model

axis and yaw around z axis. The formulation of model is described by the state space form:

$$\begin{bmatrix} \dot{\beta} \\ \ddot{\psi} \end{bmatrix} = \begin{bmatrix} \frac{Y_1}{mv} & \frac{Y_2}{mv^2} \\ \frac{Y_3}{I_z} & \frac{Y_4}{I_z v} \end{bmatrix} \begin{bmatrix} \beta \\ \dot{\psi} \end{bmatrix} + \begin{bmatrix} \frac{C_{\alpha V}}{mv} \\ \frac{l_V C_{\alpha V}}{I_z} \end{bmatrix} \delta - \begin{bmatrix} \frac{g}{v} \\ 0 \end{bmatrix} d \tag{3.55}$$

$$a_y = \begin{bmatrix} \frac{Y_1}{m} & \frac{Y_3}{mv} \end{bmatrix} \begin{bmatrix} \beta \\ \dot{\psi} \end{bmatrix} + \frac{C_{\alpha V}}{m} * \delta - g * d \tag{3.56}$$

with $Y_1 = -(C_{\alpha V} + C_{\alpha H})$, $Y_2 = l_H C_{\alpha H} - l_V C_{\alpha V} - mv^2$, $Y_3 = l_H C_{\alpha H} - l_V C_{\alpha V}$ and $Y_4 = -(l_V^2 C_{\alpha V} + l_H^2 C_{\alpha H})$, where β is the side slip angle, ψ denotes the yaw rate, a_y is the lateral acceleration, δ is the relative steering wheel angle, d represents road bank angle (considered as disturbance), and v represents the speed of the vehicle. In Tab. 3.3, typical sensor data are listed.

In this application, shown as in state space presentation (3.55) and (3.56), the working speed v affects the operation modes of the vehicle, which is also the switching signal to switch the parameters A^i, B_u^i, C^i and D_u^i in the observer except the unchanged observer gain L. In the following simulation, three subsystems are selected at the speed $v = 7m/s$, $14m/s$ and $20m/s$. In this example, a fault in steering angle δ measurement of the system is considered as the

Variable	Value	Unit	Explanation
g	9.80665	m/s^2	gravity acceleration constant
m	1463	kg	total mass
l_V	1.108	m	distance from CG to front axle
l_H	1.42	m	distance from CG to rare axle
l	$l_V + l_H$	m	distance between front and rare axle
I_z	1805.7	kgm^2	moment of inertia about the z-axis
$c'_{\alpha V}$	103950	N/rad	front axle tire cornering stiffness
$c_{\alpha H}$	108130	N/rad	rare axle tire cornering stiffness
v_{ref}		m/s	longitudinal velocity
β		rad	vehicle side slip angle
r		rad/s	vehicle yaw rate
a_y		m/s^2	vehicle lateral acceleration
δ_L^*		rad	vehicle steering angle

Table 3.3.: Data of the vehicle lateral dynamic system

additive fault, hence we set $B_f = B_u = \left[\begin{array}{cc} \frac{C_{\alpha V}}{mv} & \frac{l_V C_{\alpha V}}{I_z} \end{array}\right]^T$ and $D_f = D_u = \frac{C_{\alpha V}}{m}$ [2, 3, 35]:

$$\left[\begin{array}{c|c|c|c} A^1 & B_u^1 & B_f^1 & B_d^1 \\ \hline C^1 & D_u^1 & D_f^1 & D_d^1 \end{array}\right] = \left[\begin{array}{ccccc} -20.7 & -0.46 & 10.1 & 10.1 & -1.4 \\ 21.2 & -27.3 & 63.7 & 63.7 & 0 \\ \hline -145 & 3.74 & 71 & 71 & -9.8 \end{array}\right]$$

$$\left[\begin{array}{c|c|c|c} A^2 & B_u^2 & B_f^2 & B_d^2 \\ \hline C^2 & D_u^2 & D_f^2 & D_d^2 \end{array}\right] = \left[\begin{array}{ccccc} -9.66 & -0.88 & 4.73 & 4.73 & -0.7 \\ 21.2 & -12.7 & 63.7 & 63.7 & 0 \\ \hline -145 & 1.74 & 71 & 71 & -9.8 \end{array}\right]$$

$$\left[\begin{array}{c|c|c|c} A^3 & B_u^3 & B_f^3 & B_d^3 \\ \hline C^3 & D_u^3 & D_f^3 & D_d^3 \end{array}\right] = \left[\begin{array}{ccccc} -7.24 & -0.93 & 3.55 & 3.55 & -0.5 \\ 21.2 & -9.57 & 63.7 & 63.7 & 0 \\ \hline -145 & 1.31 & 71 & 71 & -9.8 \end{array}\right]$$

3.5.2.2. Robust fault detection observer design

In this application, the dimension of the residual is 1, as a result, there is no effect of the residual weighting matrix Q on the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$. Then, the robust fault detection observer problem is to only design an observer gain L by proposed nonsmooth optimization approach. Note that the proposed nonsmooth optimization approach could also be applied to design observer gain L and residual weighting matrix Q together. In the design, faults over a finite frequency range $[0, 0.1rad/s]$ are investigated instead of the whole frequency range, while the disturbances are on the whole frequency range $[0, \infty)$. Therefore, a high pass filter is selected as the weighting matrix W for the fault sensitivity $\|G_{rf}\|_-$.

Fig. 3.9 shows the feasible regions of the observer gain L to stabilize the observer for the above three subsystems, and the relationship among these stable regions could be obtained (Ψ represents the observer gain stable regions for the corresponding model):

$$\Psi(\text{Subsystem 3}) \subseteq \Psi(\text{Subsystem 2}) \subseteq \Psi(\text{Subsystem 1})$$

where Ψ represents the area of observer gain L to stabilize the observer for the corresponding

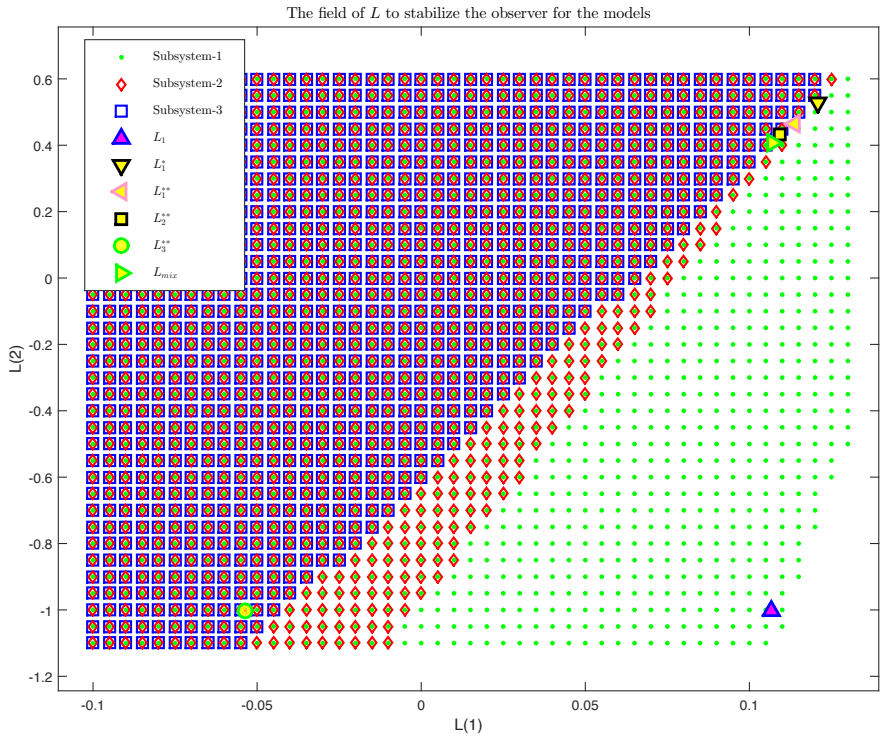

Figure 3.9.: The region of observer gain L to stabilize the three subsystems

model.

The prerequisite to design a unique observer gain L for these above three subsystems is satisfied due to the existence of the intersection of the stable region for the three subsystems. In other words, the designed unique observer gain L will stabilizes the observer for the three subsystems simultaneously. Algorithm 3.1 introduces the steps that how to design the unique observer gain L with the considering of these three subsystems simultaneously.

In order to understand the effects of the simultaneous stability constraints on the fault detection ability criteria $\|G_{rd}^i\|_\infty / \left\|G_{rf}^i\right\|_-$, first, design observers L_i for three subsystems separately with criterion (3.13):

$$\begin{aligned} L_1 &= [0.1065, -1.0039]^T; \quad \|G_{rd}^1\|_\infty / \left\|G_{rf}^1\right\|_- = 0.0220 \\ L_2 &= [0.1093, 0.4311]^T; \quad \|G_{rd}^2\|_\infty / \left\|G_{rf}^2\right\|_- = 0.0114 \\ L_3 &= [-0.0535, -1.001]^T; \quad \|G_{rd}^3\|_\infty / \left\|G_{rf}^3\right\|_- = 0.0181 \end{aligned}$$

where L_1 is typically designed only for Subsystem-1, L_2 is designed only for Subsystem-2 and L_3 is only designed for Subsystem-3.

According to the stable region of Subsystem-1, Subsystem-2 and Subsystem-3 in Fig. 3.9, the obtained observer gain for Subsystem-1 L_1 is in the stable region of Subsystem-1 (green part) but not in the feasible region of Subsystem-2 or Subsystem-3, which means that the designed L_1 cannot stabilize the observer for Subsystem-2 or Subsystem-3.

Then, the stability constraint for all the models should be added into the optimization. With the

formulation in (3.18), three stable observers L_i are designed with the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ and the stability constraint for all the models:

$$\begin{aligned} L_1^* &= [0.1207, 0.5298]^T; \quad \|G_{rd}^1\|_\infty / \left\|G_{rf}^1\right\|_- = 0.1648 \\ L_2^* &= [0.1093, 0.4311]^T; \quad \|G_{rd}^2\|_\infty / \left\|G_{rf}^2\right\|_- = 0.0114 \\ L_3^* &= [-0.0535, -1.001]^T; \quad \|G_{rd}^3\|_\infty / \left\|G_{rf}^3\right\|_- = 0.0181 \end{aligned}$$

where L_1^* is typically designed only for Subsystem-1, L_2^* is designed only for Subsystem-2 and L_3^* is only designed for Subsystem-3.

Shown in Fig. 3.9, the obtained observer gain for Subsystem-1, L_1^* is at the boundary of the intersection for three subsystems, which is also at the boundary of the stability region for Subsystem-2 and Subsystem-3. Comparing the designed observer gains L_1 and L_1^*, the value of mixed criteria $\|G_{rd}^1\|_\infty / \left\|G_{rf}^1\right\|_-$ decreases with the adding of stability region for all the subsystems. In other words, the value of mixed criteria $\|G_{rd}^1\|_\infty / \left\|G_{rf}^1\right\|_-$ will be smaller with the introduction of stricter stability constraint. Since the spectral abscissa of $A^i - L_1^* C^i$ for Subsystem-2 and Subsystem-3 are near 0, the observer gain L_1^* is not good for the design. With the optimization in (3.19), a constraint of spectral abscissa for three subsystems is added into the optimization (ξ is chosen as -0.45):

$$\begin{aligned} L_1^{**} &= [0.1139, 0.4641]^T; \left\|G_{rd}^1\right\|_\infty / \left\|G_{rf}^1\right\|_- = 0.1796 \\ L_2^{**} &= [0.1093, 0.4311]^T; \left\|G_{rd}^2\right\|_\infty / \left\|G_{rf}^2\right\|_- = 0.0114 \\ L_3^{**} &= [-0.0535, -1.001]^T; \left\|G_{rd}^3\right\|_\infty / \left\|G_{rf}^3\right\|_- = 0.0181 \end{aligned}$$

where the obtained L_2^{**} is exactly the same as L_2^*, and $L_3^{**} = L_3^*$. As introduced in the Algorithm 3.1, the weights λ_i are selected as: $\lambda_1 = 1/0.1796$, $\lambda_2 = 1/0.0114$, $\lambda_3 = 1/0.0181$. Then, using the formulation of the objective function (3.20) and the constraint (3.21) with $\xi = -0.45$ to design a compromised observer gain L_{mix}:

$$L_{mix} = [0.1070, 0.4075]^T$$

To show the relationship between the obtained observer gains, the locations of the obtained observer gains L_1^*, L_2^*, L_3^*, L_1^{**} and L_{mix} are illustrated in Fig. 3.9. The comparison on the criterion of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ among the different observer gains are listed in the Tab. 3.4.

Table 3.4.: Comparison for different observers

	Subsystem-1	Subsystem-2	Subsystem-3
L_1^{**}	0.1796	0.0117	0.0239
L_2^{**}	0.1911	0.0114	0.0234
L_3^{**}	0.5645	0.0856	0.0181
L_{mix}	0.1960	0.0115	0.0196

Tab. 3.4 shows that the values of the mixed criterion $\|G_{rd}^i\|_\infty / \left\|G_{rf}^i\right\|_-$ with the observer gain L_1^{**} and Subsystem-1, L_2^{**} and Subsystem-2 and L_3^{**} and Subsystem-3 are smallest, which means

that the observer gain L_i^{**} is the optimal design for the Subsystem-i in the worst case, but they are not the good design for other Subsystem-j ($j \neq i$). Evaluating the three subsystems with selected weights λ_i, the integrated design, L_{mix}, gives the Pareto optimal solution, whose effects are the compromise among the three subsystems with the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$. The compromised observer gain L_{mix} also guarantees fast transient responses ($\alpha\left(A^i - LC^i\right) < -0.45$) for all three subsystems at the same time.

As introduced in [36], a smaller $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ means more faults can be strongly detectable with either fixed disturbances. It is the reason that the mixed performance index $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ is considered to evaluate the system's ability of fault sensitivity and disturbance robustness. In this application, with a fixed energy bounded disturbances, faults with large energy in the frequency range $[0,\ 0.1rad/s]$ could be detected with all the observer gains. Therefore, it is interesting to compare the the lower bounds of the strongly detectable faults in the frequency range $[0,\ 0.1rad/s]$ for the different observer gains with a fixed energy bounded disturbances. Given a pulse fault signal with fixed time interval, the energy of the fault could be represented by the amplitude of pulse fault signal. A larger amplitude of pulse fault signal when the fault could be detected means a smaller set of strongly detectable faults with the fixed bounded energy disturbances. Then, with the fixed energy bounded disturbances, the lower bounds of the strongly detectable faults could be represented by the smallest amplitude of the pulse fault signal when the fault could be detected. If a pulse fault signal with an amplitude cannot be detected, the lower bound of the strongly detectable faults is higher than this amplitude. Based on this idea, the smallest amplitude of pulse fault signal with different observer gains for different subsystems could be compared to illustrate the effects of the integrated design L_{mix}. Taking the pulse faults as

$$f = \begin{cases} \text{Amp1} = 0.0435, & 5 \leq t \leq 15 \\ \text{Amp2} = 0.0094, & 25 \leq t \leq 35 \\ \text{Amp3} = 0.0120, & 45 \leq t \leq 55 \\ 0, & \text{elsewhere} \end{cases}$$

and the energy bounded disturbances are uniform random disturbances during $0 \leq t \leq 60s$. The value of $\max\left(\|d\|_{\text{rms}}\right)$ in the threshold (3.27) could be calculated with the uniform random signal before the simulation in this case.

In the simulation, there a fault in Subsystem-1, Subsystem-2 and Subsystem-3 separately before the switching, and the switching time of the Subsystems is selected at 20s and 40s. The thresholds and the time responses of the residual signals with the proposed faults and disturbances are shown in Fig. 3.10.

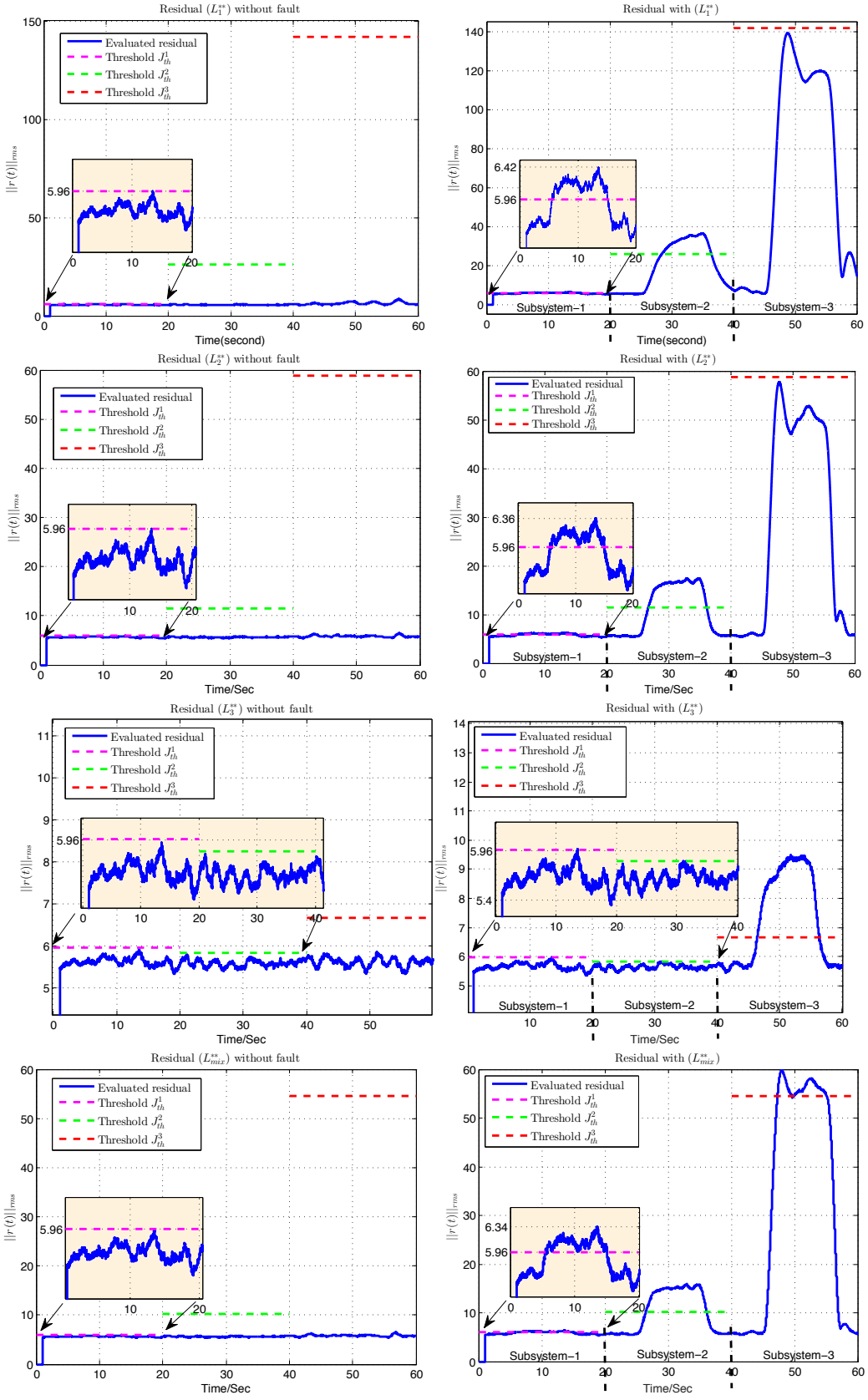

Figure 3.10.: Switching residual (L_1^{**}, L_2^{**}, L_3^{**} and L_{mix}) without fault (left) and with pulse faults (right)

The thresholds are chosen with the definition in (3.27), which are the estimations of possible maximum effects from the disturbances on the residual. For the case of L_1^{**} with Subsystem-2, the maximum singular value of G_{rd} is at low frequency, and the disturbances in the simulation are uniform random disturbances, whose high frequency part always has big amplitudes while the amplitudes of low frequency part always are small. Analyzed in Remark 3.4, when these two frequencies are different, the thresholds for L_1^{**} with Subsystem-2 is higher than the biggest value of the evaluated residual only with disturbances ($20s \leqslant t \leqslant 40s$ for without faults case) in Fig. 3.10. By contrast, for the cases of L_3^{**} with Subsystem-2 and Subsystem-3, the frequency ω_1 for $G_{rd}(j\omega_1) = \|G_{rd}\|_\infty$is in frequency range of the disturbances (which achieve $\max(\|d\|_{\text{rms}})$), thus, the threshold J_{th} is just higher than the peak of the evaluated residual when without faults.

Comparing the effects of the robust fault detection observer gains L_1^{**}, L_2^{**}, L_3^{**} and L_{mix}, we can find that the evaluated residual with L_1^{**} can detect fault for Subsystem-1 and Subsystem-2 with the corresponding thresholds, however, the fault cannot be detected for Subsystem-3. In the view of strongly detectable faults, the lower bound of the strongly detectable faults for L_1^{**} with Subsystem-1 and Subsystem-2 are smaller than Amp1 and Amp2 respectively, but with Subsystem-3 is bigger than Amp3. A similar fault detection result is shown for the case with observer gain L_2^{**}. For Subsystem-1, observer gain L_1^{**} shows more space of the amplitude for the pulse fault signal to decrease when the fault could be detected than observer gain L_2^{**}, while inverse results for Subsystem-2. Both faults for Subsystem-1 and Subsystem-2 cannot be detected by observer gain L_3^{**}. In other words, the smallest amplitude of the strongly detectable faults for L_3^{**} with Subsystem-1 and L_3^{**} with Subsystem-2 are larger than Amp1 and Amp2 respectively. The simulation shows that the designed observer gain L_i^{**} with Subsystem-i ($i : 1,\ 2,\ 3$) has the largest set of strongly detectable faults with fixed energy bounded disturbances. Whereas, at least one lower bound of strongly detectable faults for either observer gain of L_1^{**}, L_2^{**}, L_3^{**} with three subsystems is larger than the amplitudes (Amp1, Amp2 and Amp3) of the provided three pulse faults for Subsystem-1, Subsystem-2 and Subsystem-3. Compromising the redundancy of the strongly detectable faults set for these three observer gains, the observer gain L_{mix} can detect all three faults. This also shows that all the low bounds of strongly detectable faults for three subsystems with L_{mix} are smaller than the amplitudes of the proposed pulse faults for different subsystems, but compared with the optimal design (L_i^{**} for Subsystem-i ($i : 1,\ 2,\ 3$)), the compromised design L_{mix} has fewer space of the amplitude for the pulse fault signal to decrease when the fault could be detected. The meaning of the compromised design L_{mix} is to utilize the large fault margins for some subsystems to compensate the small fault margins for other subsystems. The simulation validates the principle of this design: comparing with the observer gains L_1^{**}, L_2^{**}, L_3^{**} for three subsystems, L_{mix} has compromised set of strongly detectable faults.

3.5.2.3. Robust fault detection filter design

The robust fault detection filter is designed in the case that the system is in one of the N models, which is unknown. There is no model switching in this case.

Firstly, design the RFDF optimal filter (full order filter $n_F = 2$(plant order)+1(reference model order) = 3) for the three subsystems separately with the criterion (3.49) using nonsmooth optimization

method, we can get:

$$F_1 : \left[\begin{array}{ccc|cc} -4.896 & -5.585 & 0.1527 & -0.666 & 6.7746 \\ 2.0631 & -4.699 & 2.1166 & 2.1626 & 1.6277 \\ 0.1858 & -5.705 & -4.3389 & -0.309 & -4.732 \\ \hline 5.561 & 0.635 & -5.2426 & 0.8808 & 6.1849 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^1\right\|_- = 2.8379$$

$$F_2 : \left[\begin{array}{ccc|cc} -3.948 & -0.1674 & -0.258 & -0.043 & 0.9507 \\ 571.09 & -5.8794 & -1.016 & 6.198 & 16.337 \\ -1.001 & -0.6295 & -7.828 & -0.379 & -1.4770 \\ \hline -1.645 & 0.0955 & 0.975 & 1.117 & -6.0674 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^2\right\|_- = 0.8489$$

$$F_3 : \left[\begin{array}{ccc|cc} -10.28 & -9.9863 & 4.252 & -0.405 & -3.559 \\ 14.182 & -3.1785 & -6.398 & -0.808 & 12.174 \\ -3.107 & -5.0183 & -8.118 & -1.103 & 0.7175 \\ \hline 3.461 & 0.3134 & -0.793 & 0.877 & 7.7687 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^3\right\|_- = 0.6585$$

Then, using the complex criteria (2.33) to do the optimization for the three subsystems, the weights are selected as discussed before: $\lambda_1 = 1/2.8379, \lambda_2 = 1/0.8489, \lambda_3 = 1/0.6585$. The optimal unique filter is:

$$F_{mix} = \left[\begin{array}{ccc|cc} -7.50 & -2.413 & 7.301 & 0.5961 & -42.724 \\ 46.79 & -5.043 & 3.878 & -3.383 & 20.421 \\ 5.859 & 2.563 & -604 & -124.7 & 0.325 \\ \hline 1.438 & 0.032 & -0.008 & 0.8005 & 14.204 \end{array}\right]$$

Table 3.5.: Comparison for different filters F

	Subsystem-1	Subsystem-2	Subsystem-3
F_1	2.8379	2.3851	2.5239
F_2	3.3392	0.8489	0.8994
F_3	3.8200	0.9711	0.6585
F_{mix}	3.4558	0.8786	0.8016

Tab. 3.5 shows that the mixed performance indexes $\left(\max_{i=1,\ ,2,\ 3}(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty)\right) / \left\|G_{rf}^i\right\|_-$ of the filter F_1 for Subsystem-1, F_2 for Subsystem-2 and F_3 for Subsystem-3 are smallest, which means that they are the optimal design for the separate subsystem in the worst case, however, these designs are not most suitable for the other subsystems. Considering three subsystems together, the design of the mixed case, F_{mix}, gives the Pareto optimal solution for the three different subsystems, whose effects are the trade-off among the three subsystems with the value of $\left(\max_{i=1,\ ,2,\ 3}(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty)\right) / \left\|G_{rf}^i\right\|_-$.

According to the design procedures in Algorithm 3.2, it is easy to check that $n_F = 1$ will decrease the ability of fault detection. Therefore, design the reduced order RFDF filters (order =2) for the subsystems separately, we can get:

$$F_1^*: \left[\begin{array}{cc|cc} -5.2401 & 0.6672 & 0.0231 & -1.2898 \\ -1.6470 & -5.5341 & 0.0158 & -0.6203 \\ \hline 2.5572 & 1.9186 & 1.0061 & -0.7325 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^1\right\|_- = 2.8377$$

$$F_2^*: \left[\begin{array}{cc|cc} -10.8478 & -6.6473 & -0.0881 & -1.6673 \\ 14.8578 & 0.2586 & 0.0532 & 3.0665 \\ \hline -0.3313 & 0.9508 & 1.0263 & -1.5504 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^2\right\|_- = 0.8489$$

$$F_3^*: \left[\begin{array}{cc|cc} -59.1196 & -38.8492 & -0.4797 & 0.9983 \\ 78.0004 & 48.7694 & 0.5749 & 2.8033 \\ \hline 0.2967 & 0.4156 & 1.0193 & -1.1740 \end{array}\right]$$

$$\left(\max_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^3\right\|_- = 0.6586$$

Then, select the weights as: $\lambda_1 = 1/0.1392, \lambda_2 = 1/0.1689, \lambda_3 = 1/0.1940$ and optimize the complex criteria (3.50) for the three subsystems by proposed nonsmooth optimization method. We get the compromised unique filter design:

$$F_{mix}^* = \left[\begin{array}{cc|cc} 18.3159 & -41.5642 & -1.2354 & -54.7168 \\ 13.4355 & -25.5919 & -1.4087 & -11.0455 \\ \hline 0.0242 & -0.0129 & 0.9839 & 0.8856 \end{array}\right]$$

The effects of the optimized filter F_{mix}^* on the three subsystems are:

$$\frac{\max_{i=1,\ ,2,\ 3}\left(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty\right)}{\left\|G_{rf}^1\right\|_-} = \frac{0.8394}{5.6342} = 3.4480$$

$$\frac{\max_{i=1,\ ,2,\ 3}\left(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty\right)}{\left\|G_{rf}^3\right\|_-} = \frac{1.0978}{6.0811} = 0.9564$$

$$\frac{\max_{i=1,\ ,2,\ 3}\left(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty\right)}{\left\|G_{rf}^3\right\|_-} = \frac{1.2301}{5.9520} = 0.7998$$

From Tab. 3.5 and Tab. 3.6, comparing with the full order filter, the reduced order filter shows the equivalent ability of robustness to the disturbances and sensitivity to the faults for the three subsystems in the worst case.

As introduced in [36], a smaller $\left(\max_{i=1,\ ,2,\ 3}\left(\|G_{rd}^i\|_\infty + \|G_{ru}^i\|_\infty\right)\right) / \left\|G_{rf}^i\right\|_-$ means more faults can be strongly detectable with some fixed disturbances and control inputs. In this application, since some big faults could be detected with fixed disturbances and control inputs by either

Table 3.6.: Comparison for different reduced order filters

	Subsystem-1	Subsystem-2	Subsystem-3
F_1^*	2.8377	2.6061	2.6777
F_2^*	3.3393	0.8489	0.8990
F_3^*	3.8052	0.9674	0.6586
F_{mix}^*	3.4480	0.9564	0.7998

reduced order filter, it is interesting to compare the lower bounds of the strongly detectable faults with different filters, but same fixed disturbances and control inputs.

To illustrate the effects of the designed robust fault detection filter F_{mix}^* with the index in Tab. 3.6 (a design with smaller value will be easier to detect faults), all the designed reduced order filters F_1^*, F_2^*, F_3^* and F_{mix}^* are simulated with the discussed 3 subsystems by taking the faults, disturbances and control inputs as in Tab. 3.7.

	Faults	Disturbances	Control inputs
Subsystem-1	$f_1 = -0.3 + 0.3sin\,(15t)$	$sin\,(100t)$	4
Subsystem-2	$f_2 = 0.4 + 1.4sin\,(10t)$	$12sin\,(100t)$	1
Subsystem-3	$f_3 = -0.1 + 0.5sin\,(15t)$	$sin\,(15t)$	1

Table 3.7.: Settings of fault, disturbance and input for simulation

In the simulation, the proposed four reduced order filters will be switched with an order of F_1^*, F_2^*, F_3^*, F_{mix}^* for each subsystem, and the switching time is selected at 20s, 40s and 60s. The fault f_i arises for Subsystem-i at $5s \leqslant t \leqslant 15s$, $25s \leqslant t \leqslant 35s$, $45s \leqslant t \leqslant 55s$, and $65s \leqslant t \leqslant 75s$. The disturbances and control inputs in Tab. 3.7 are implementing during $0s \leqslant t \leqslant 80s$ with the corresponding subsystem. The time responses of the residual signals and the corresponding thresholds with faults, disturbances and control inputs are shown in Fig. 3.11.

According to (3.42), the thresholds are selected as $\gamma_u^{max}\left\|u\right\|_{rms} + \gamma_d^{max}\max\left(\left\|d\right\|_{\text{rms}}\right)$, which are independent on the model information i, but dependent on the used filter (F_1^*, F_2^*, F_3^* or F_{mix}^*). Therefore, with estimated worst disturbance $\max\left(\left\|d\right\|_{\text{rms}}\right)$, known input $\left\|u\right\|_{\text{rms}}$ and selected reduced order filter, the corresponding threshold is also determined. As a consequence, all the processes of fault detection are achieved without the model information i. On the one hand, the parameters γ_u^{max} and γ_u^{max} are the maximum values of γ_u^i and γ_d^i. On the other hand, the frequency ω_1 of the disturbances d and the frequency ω_2 of the control inputs u are different from the frequency points which achieve $\sigma\left(G_{rd}\left(j\omega_3\right)\right) = \left\|G_{rd}\right\|_\infty$ and $\sigma\left(G_{ru}\left(j\omega_4\right)\right) = \left\|G_{ru}\right\|_\infty$ respectively, the threshold J_{th} will be higher than the peak amplitude of the evaluated residual $\left\|r\right\|_{\text{rms}}$ when without faults, such as cases of Subsystem-2 with either of four filters.

The simulation results in Fig. 3.11 shows that for the Subsystem-1, f_1 could be detected by F_1^*, F_2^* and F_{mix}^* except F_3^* with the selected disturbances and control inputs. The differences among the effects of fault detection with F_1^*, F_2^* and F_{mix}^* are the space of the amplitude for f_1 to decrease when fault could be detected. Obviously, filter F_1^* for Subsystem-1 has the most space, while filters F^*2 for this subsystem have more space than the filter F_{mix}^* for this subsystem, which explains the effects of the values of the criteria $\left(\max\limits_{i=1,\ ,2,\ 3}\left(\left\|G_{rd}^i\right\|_\infty + \left\|G_{ru}^i\right\|_\infty\right)\right) / \left\|G_{rf}^i\right\|_-$

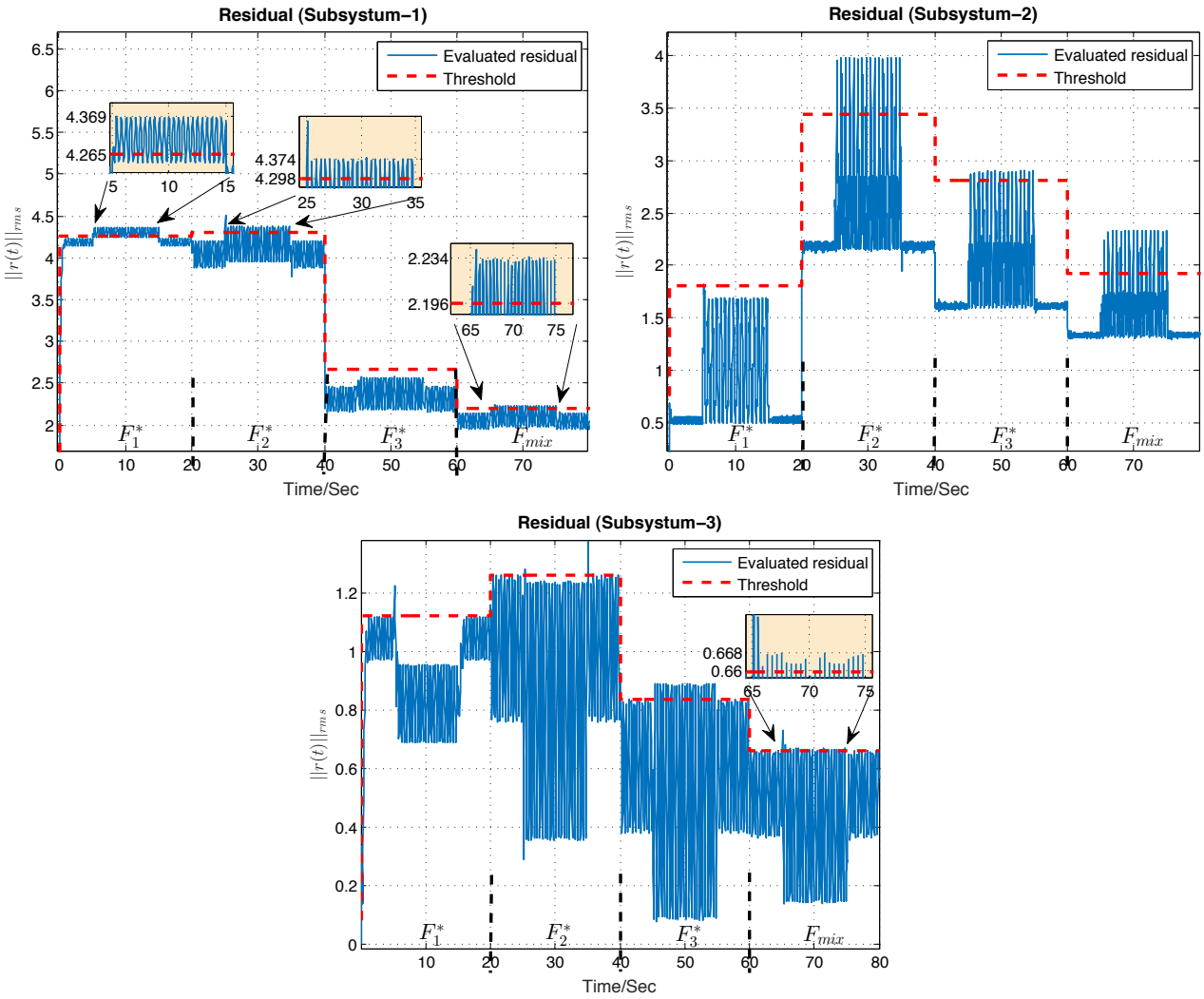

Figure 3.11.: Residual (reduced order filters) with typical faults and disturbances

for Subsystem-1 with these four filters in Tab. 3.7: a smaller value of the criteria will have more space of the amplitude for f_1 to decrease when fault could be detected, which has a better ability of fault detection. In the Subsystem-2, the filter F_1^* fails the detection for the fault f_2. The filter F_2^* has most space while filter F_3^* has fewer space than the filter F_{mix}^*. For the Subsystem-3, filter F_3^* has more space than filter F_{mix}^* while the fault f_3 could not be detected by the filters F_1^* and F_2^*. The simulation also shows that the filter F_i^* with Subsystem-i ($i: \ 1, \ 2, \ 3$) always has the largest space of the amplitude for fault signal to decrease with proposed fixed disturbances and control inputs. However, at least one of the proposed faults cannot be detected for the corresponding subsystem with filter F_i^*. Although the compromised design F_{mix}^* does not have the best ability of fault detection for either subsystem with the corresponding fault, all the proposed faults can be detected for the corresponding subsystem by the filter F_{mix}^*. The meaning of the compromised design F_{mix}^* is to use the best ability of fault detection for some subsystems to compensate the worst ability of fault detection for other subsystems. The simulation in Fig. 3.11 validates the principle of the design F_{mix}^*.

3.6. Conclusion

This chapter has introduced the application of nonsmooth optimization techniques on the FDI problems. Typically, the worst case is considered, where $\|\cdot\|_-$ index is used to measure the fault sensitivity, and $\|\cdot\|_\infty$ is used to describe robustness to disturbances. Different from the method of LMI, the proposed nonsmooth optimization approach with Systune in matlab optimizes the performance index H_-/H_∞ directly and the algorithm has a good convergence to find an optimal solution. To improve the fast transients of the residual when fault appears, a constraint of eigenvalues could be added. In the first framework of the design for multiple models case, a full order observer is applied to generate the residual. The problem of generating residual is transformed to be an optimization problem with a constraint of simultaneously stabilizing the observers for different models. A unique observer gain L and residual weighting matrix Q for multiple models with a Pareto optimal value of $\|\cdot\|_\infty / \|\cdot\|_-$ for different models could be found by proposed nonsmooth optimization approach. However, the model switching signal should be introduced to switch the parameters A, B, C and D in the observer when the model switches. For the typical case that the system stays in one of the multiple models, and the corresponding model is unknown, a new framework with deconvolution approach is proposed to design RFDF, which should be constant and robust to the models information i. The design is formalized as a constrained optimization problem. An unchanged compromised threshold is proposed to compare with the residual to detect fault. Both RFDF filter and threshold are constant, and do not depend on the operational mode of the system. Finally, an academic example and a vehicle lateral dynamics switched system with 3 subsystems are illustrated to prove the effectiveness of the proposed design method.

4. FDI observer design using time and frequency domain specifications

4.1. Introduction

As discussed in the previous section, early fault detection is an important factor of a FDI system. The eigenvalues could be chosen to improve the rapidity of residual responses, which provides potentials to decrease fault detection delay. In [115], pole assignment approach is utilized to transform the fault detection problem into an unconstrained optimization problem, which could be solved by a gradient based optimization method. With the aid of LMI method, constraints on pole location are also added into the mixed H_-/H_∞ design in [2, 32, 51]. In Chapter 3, eigenvalues (poles) were used to improve the transients of residual. The locations of system poles have typical effects on the transients of system (G represents a transfer function):

- Minimum decay rate: $\alpha\left(G\left(j\omega\right)\right) < -\varsigma$
- Minimum damping ratio: $Re\left(G\left(j\omega\right)\right) < -\xi\left|G\left(j\omega\right)\right|$
- Pole magnitude (also called as natural frequency): $\left|G\left(j\omega\right)\right| < \omega_{max}$

Increasing the minimum decay rate value ς results in faster transients. Increasing the minimum damping ratio value ξ results in better damped transients, while decreasing the pole magnitude value ω_{max} prevents fast dynamics. For a low order system, poles could be used to improve the transients of responses. However, for a high order system, this criterion may not make any sense.

For a single input single output (SISO) system, a stable high order transfer function could be decomposed into a sum of some different stable low order transfer functions:

$$G\left(s\right) = \sum_{i=1}^{n} \frac{a_i\left(s\right)}{b_i\left(s\right)} = \eta + \sum_{i=1}^{m} \frac{K_i\omega_i}{\tau_i^1 s + \omega_i} + \sum_{i=m+1}^{n} \frac{\left(\tau_i^2 s + 1\right) K_i\omega_i^2}{\left(\tau_i^3 s^2 + 2\xi_i\omega_i s + \omega_i^2\right)} \tag{4.1}$$

where the parameters τ_i^1, τ_i^2, τ_i^3, ξ_i and ω_i are the parameters of g_i, η is a constant value, and the parameter K_i is the steady state gain of g_i. The real part of the eigenvalues of G could be represented as λ_i ($\lambda_i < 0$ when G is stable). We assume that

$$\max_{\mathrm{i}=1,\ldots,\mathrm{n}} \left(\lambda_{\mathrm{i}}\right) = \lambda_{\mathrm{j}},\ \mathrm{j} \in \{1, \ldots, \mathrm{n}\}$$

In general, the eigenvalues, which have the maximum real part λ_j, determine the major transient behaviors of responses. However, if the steady state gain K_i of g_i ($i \neq j$) is much bigger than the steady state gain K_j of g_j, the dynamics of the responses will mainly depend on g_i. In other words, the eigenvalues, which have the maximum real part λ_j, cannot be used to evaluate the transients of responses. Therefore, in order to adjust the transients of responses with

the specification of poles, both poles and the corresponding steady state gain K_i should be considered simultaneously. As a result, only optimizing the locations of poles is not enough to improve the transients of responses.

Example 4.1. Considering a 4 order system with two transfer functions:

$$T_1 = \frac{1}{s^4 + 4.244s^3 + 92.25s^2 + 186.1s + 98.41}$$
$$T_2 = \frac{1}{s^4 + 3.074s^3 + 87.31s^2 + 175.2s + 313.5}$$

The poles of T_1 are $-1.0616 \pm 9.2458i$ and $-1.0604 \pm 0.1084i$, while the poles of T_2 are $-0.4840 \pm 9.0104i$ and $-1.0530 \pm 1.6558i$. As shown in Fig. 4.1, the spectral abscissa of T_1 is -1.0604, by contrast, the spectral abscissa of T_2 is -0.4840. According to the effects of spectral abscissa (3.8), the responses of T_1 should have faster transients than the responses of T_2.

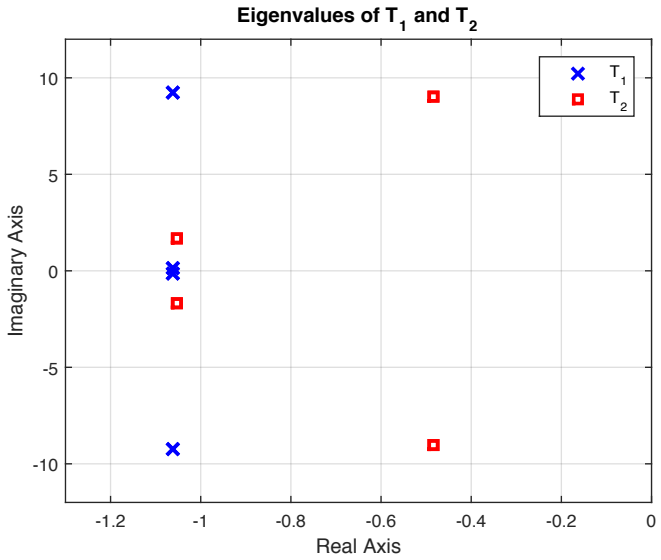

Figure 4.1.: Poles of T_1 and T_2

However, as shown by the simulation of the step responses of T_1 and T_2 in Fig. 4.2(a), the dynamics (rise time, settling time) of responses with T_2 are better than the responses with T_1, which is contrary to the previous analysis. This phenomenon is caused by neglecting the effects of the steady state gain K_i on the transients of responses in (4.1).

Applying the partial fraction decomposition to T_2:

$$T_2 = T_2^1 + T_2^2$$
$$T_2^1 = \frac{-0.0002s - 0.01271}{s^2 + 0.968s + 81.42}, T_2^2 = \frac{0.0002s + 0.01279}{s^2 + 2.106s + 3.85}$$

where the poles of T_2^1 are $-0.4840 \pm 9.0104i$ and the poles of T_2^2 are $-1.0530 \pm 1.6558i$. Since the steady state gain of T_2^1 is much smaller than the steady state gain of T_2^2, the responses of the T_2 mainly depend on T_2^2. As a result, the transients of the responses T_2 should be evaluated by the spectral abscissa of T_2^2, even the eigenvalues of T_2^1 is much nearer the imaginary axis. Simulation in Fig. 4.2(b) validates the analysis. Therefore, for this example, the effects of spectral abscissa is unavailable to evaluate the transients of responses in time domain.

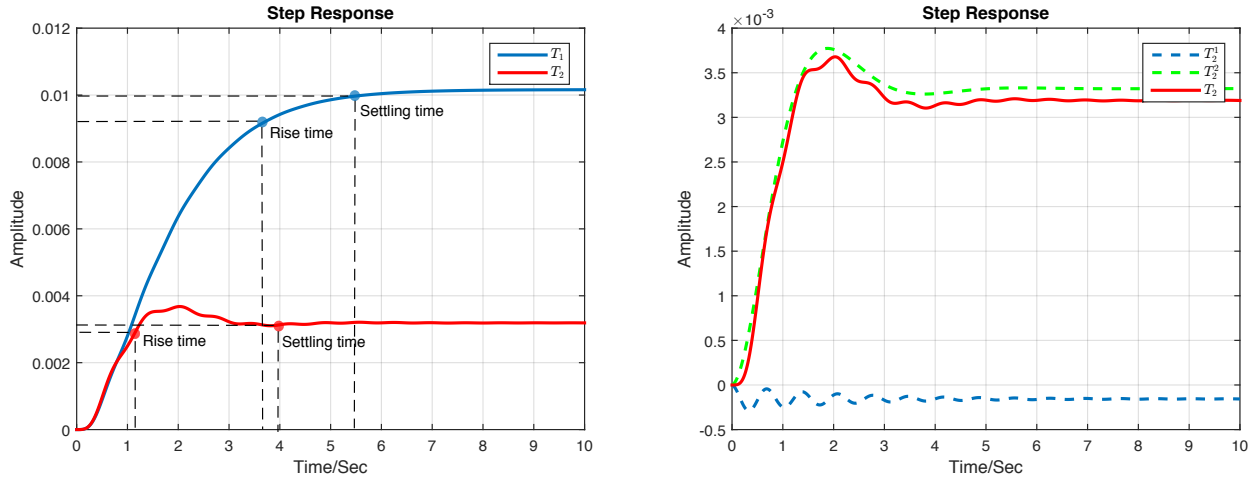

Figure 4.2.: Step responses of T_1, T_2 (a) and T_2^1, T_2^2 and T_2 (b)

To evaluate the transients of response, an alternative way is to design residual in time domain directly. In this chapter, we investigate the design of the residual in time and frequency domain simultaneously. Section 4.2 presents the framework of residual generation with a full-order observer firstly. Then, with the objectives of a "good" fault detection system, the mixed criterion H_-/H_∞ is used to evaluate the ability of fault detection in the worst case for unknown faults, while a specification in time domain is used to improve the transients of the residual for some specific faults. For the specific faults case, an analysis between the different design objectives and the specifications in time and frequency domain is given. In Section 4.3, an integrated design with worst case for unknown faults and fast fault detection for specific faults is introduced. With the aid of an iterative algorithm to minimize the fault detection delay, a design with least fault detection delay for step fault is achieved with an tolerable degradation of the ability of fault detection in the worst case. Furthermore, a constraint of decreasing false alarms when step fault disappears is added into the design, which could be realized with criteria in time and frequency domain. The effectiveness of design is validated with a numerical example in Section 4.4. Finally, Section 4.5 gives a conclusion.

4.2. Problem formulation

4.2.1. Residual generation

The linear time invariant (LTI) system with faults and disturbances is described by

$$\Sigma_0 \begin{cases} \dot{x} = Ax + B_u u + B_f f + B_d d, \\ y = Cx + D_u u + D_f f + D_d d, \end{cases} \tag{4.2}$$

where $x \in \mathbb{R}^{n_x}$ is the system state vector, $y \in \mathbb{R}^{n_y}$ represents the output measurement vector, $f \in \mathbb{R}^{n_f}$ denotes the fault vector, which can be the sensor faults, process faults, or actuator faults. $d \in \mathbb{R}^{n_d}$ is the unknown input vector, including disturbance, modeling error, process and measurement noise or uninterested fault. $u \in \mathbb{R}^{n_u}$ is the control input vector. Matrices

A, B_u, C, D_u, B_f, D_f, B_d, D_d are constant with appropriate dimensions. Without loss of generality, the following assumptions are used:

- (A, C) is detectable.
- f and d are $\mathcal{L}_2$ norm bounded.

For residual generation, a full-order observer in the following form [27] is used:

$$\Sigma_1 \begin{cases} \dot{\hat{x}} = A\hat{x} + B_u u + L(y - \hat{y}), \\ \hat{y} = C\hat{x} + D_u u, \\ r = Q[y - \hat{y}]. \end{cases} \tag{4.3}$$

where $\hat{x} \in \mathbb{R}^{n_x}$ and $\hat{y} \in \mathbb{R}^{n_y}$ are the system's state and output estimations, $r \in \mathbb{R}^{n_r}$ is the residual vector, $L \in \mathbb{R}^{n_x \times n_y}$ is the observer gain to design, and $Q \in \mathbb{R}^{n_r \times n_y}$ is the residual weighting matrix, which could be static or dynamic as a $Q(s)$.

Connecting the observer $\sum_1$ in (4.3) with the system $\sum_0$ in (4.2) together, and considering the state estimation error $e = x - \hat{x}$, we get the residual error dynamic equations:

$$\Sigma_2 \begin{cases} \dot{e} = (A - LC)e + (B_f - LD_f)f + (B_d - LD_d)d, \\ r = QCe + QD_f f + QD_d d. \end{cases} \tag{4.4}$$

The corresponding residual response from faults and disturbances in frequency domain is:

$$\begin{aligned} r =& Q\{D_f + C(sI - A + LC)^{-1}(B_f - LD_f)\}f \\ &+ Q\{D_d + C(sI - A + LC)^{-1}(B_d - LD_d)\}d \\ =& G_{rf}(L, Q)f + G_{rd}(L, Q)d \end{aligned} \tag{4.5}$$

Obviously, the dynamics of residuals depend not only on the effects of disturbances, but also on the effects of faults. Therefore, the multiobjective design of fault detection observer (design the observer gain L and the residual weighting matrix Q) contains the following objectives for unknown faults and disturbances:

i) The residual error dynamics equations (4.4) with the observer gain L should be stable,

ii) Maximize the effects of faults on the residual,

iii) Minimize the effects of disturbances on the residual,

iv) Detect the fault as fast as possible.

The i), ii) and iii) objectives were analyzed in the previous chapter. This chapter tries to propose some performance indexes to decrease fault detection delay (early fault detection) for objective iv).

Obviously, the rapidity to detect fault is an important criterion to design the observer. The time to detect fault depends not only on the transients of residual, but also on the selection of threshold. A lower threshold will give a faster fault detection at a risk of introducing false alarms. By contrast, a higher threshold will miss some faults and increase the fault detection delay. Thus, the objective iv) should be achieved with the condition of zero false alarms. An adaptive threshold was introduced in (3.26), which depends upon the nature of the system

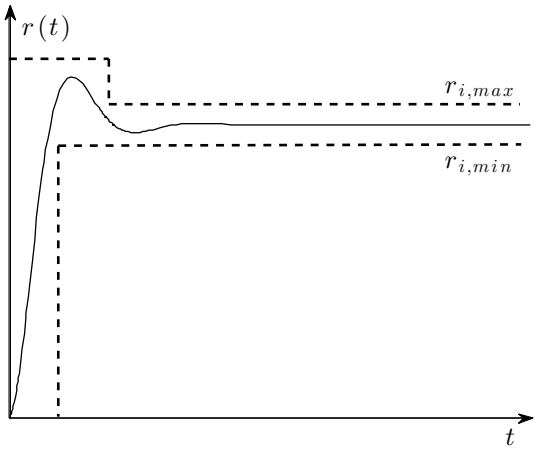

Figure 4.3.: Shape-constraints of the residual r

uncertainties and disturbances. The settings of threshold in this chapter are same as in (3.26):

$$\begin{aligned} J_{th} &= \|r\|_{\text{rms},\ f=0} = \|G_{rd}d\|_{\text{rms}} \leqslant \|G_{rd}\|_{\infty} \|d\|_{\text{rms}} \\ &= \gamma_d \|d\|_{\text{rms}} \leqslant \gamma_d \cdot \max\left(\|d\|_{\text{rms}}\right) \end{aligned} \tag{4.6}$$

where $\max\left(\|d\|_{\text{rms}}\right)$ is the upper bound with the worst disturbances acting on the plant, which can be calculated off-line. Therefore, in view of design, the threshold only depends on $\|G_{rd}\|_{\infty}$. An assumption with $\max\left(\|d\|_{\text{rms}}\right) = 1$ does not introduce any conservativeness, then, the threshold is $J_{th} = \|G_{rd}\|_{\infty}$. In the discussion of the following part, the threshold means the value of $\|G_{rd}\|_{\infty}$

In the view of the dynamics of the generated residual, as discussed before, the specification of pole is not enough to improve the fast transients of residual. In traditional frequency design method for time invariant systems, the criteria in time domain such as overshoot, rise or settling time cannot be addressed directly. An alternative way is to add the time domain constraints on the residuals r for some specific reference inputs such as impulses, steps or other inputs, and then the observer should be designed to let the residuals r follow up the given behaviors. Therefore, the time responses of the residual r from a specific fault signal f with the observer gain L and residual weighting matrix Q must satisfy the envelope constraints

$$\begin{aligned} r_{i,\ min} \leqslant r_i\left(L,\ Q,\ t\right) \leqslant r_{i,\ max}, \\ i \in I := \{1, \ldots, n_r\} \end{aligned} \tag{4.7}$$

The transient behaviors of the residual could be designed with the requirements of fast fault detection. In this chapter, a step fault signal is considered to show the effectiveness of the design, then the constraints in time domain (4.7) could be illustrated with Fig. 4.3.

Owing to the criterion in time domain (4.7), the objective iv) is translated to achieve fast fault detection for the selected fixed fault, which is considered as the specific case. In the case of unknown faults ($\mathcal{L}_2$ bounded), as introduced in chapter 3, the worst case (the case of minimum influence of fault and maximum effects of disturbance on the residual) is more interesting to represent the specifications i), ii) and iii), which is considered as the general case here. This chapter tries to combine the designs of specific case and general case together, which will show good sensitivity to unknown faults and robustness to unknown disturbances with the ability of

fast fault detection for some fixed faults.

Remark 4.1. Due to the fact that fault is detected only when $\|r\|_{\text{rms}} > J_{th}$ and the time to detect faults is affected by the transients of evaluated residual $\|r\|_{\text{rms}}$, it is straightforward to adjust the evaluated residual $\|r\|_{\text{rms}}$ with the envelopes $\|r_{i,\ min}\|_{\text{rms}}$ and $\|r_{i,\ max}\|_{\text{rms}}$ directly. As the optimization in time domain needs the calculation of response every time when the design parameters change, the computation complexity is increased due to the calculation of the evaluated residual $\|r\|_{\text{rms}}$. In fact, when the fault is a step fault signal, the transients of the evaluated residual $\|r\|_{\text{rms}}$ could be approximated by the transients of the residual r directly: a residual r with good transients means the corresponding evaluated residual $\|r\|_{\text{rms}}$ also has good transients. With this setting, the calculation of the RMS norm $\|\cdot\|_{\text{rms}}$ could be avoided. The envelopes with $r_{i,\ min}$ and $r_{i,\ max}$ could be used to adjust the transients of residual directly in the implements of simulation.

Remark 4.2. The influence of the finite time window T on the evaluated residual is detailed in chapter 3. The finite time window T for the residual evaluation in (3.22) will also affect the time to detect fault: as shown in Fig. 4.4, a smaller finite time window T will have faster transients while a larger T will cause worse transients. However, the small finite time window T will cause more effects of disturbances on the evaluated residual, which disturbs the fault detection. This chapter tries to design fast fault detection observer with a fixed time window T.

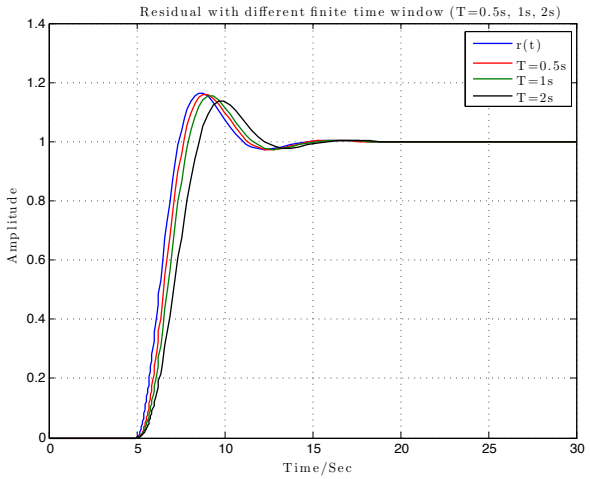

Figure 4.4.: The effects of finite time window T (3.22) on the transients of residual r

4.2.2. Quantitative analysis for the criteria in time and frequency domain

To evaluate the effects of fast fault detection, the time t_{detect} when the observer begins to detect fault could be defined as

$$\{t_{detect} \mid \|r\left(t_{detect}\right)\|_{\text{rms}} \geqslant J_{th}, \|r\left(t_{detect} - \xi\right)\|_{\text{rms}} < J_{th}\} \tag{4.8}$$

where ξ is a tiny positive value. As shown in the red box ϱ_1 of Fig. 4.5 and Fig. 4.6, the fault appears at t_{appear}, which is detected at t_{detec}. Therefore, the delay to detect fault can be defined as $t^1_{delay} = t_{detec} - t_{appear}$. In the view of residual generation, a residual with faster transients

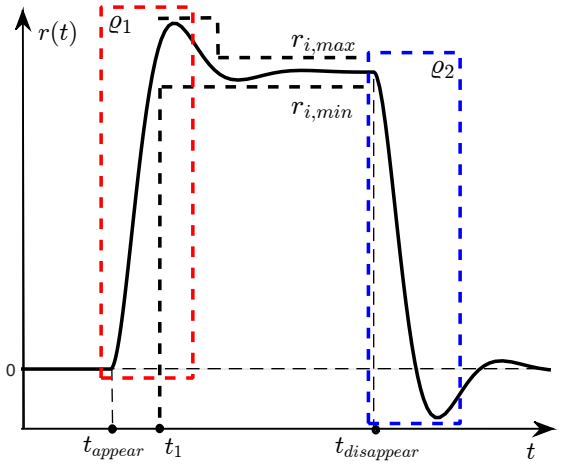

Figure 4.5.: Residual r with pulse fault (Bounded by $r_{i,\ max}$ and $r_{i,\ min}$)

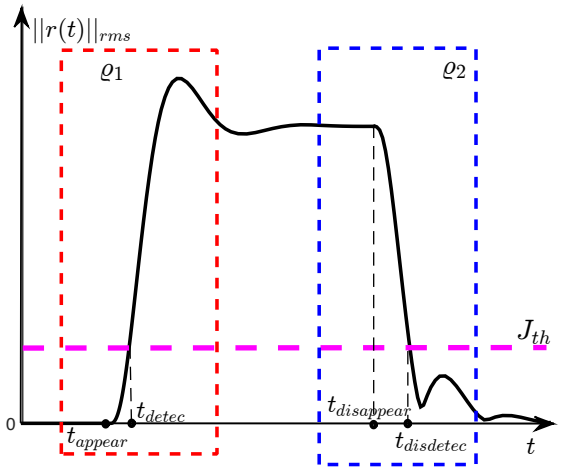

Figure 4.6.: Evaluated residual $\|r\|_{\text{rms}}$with pulse fault

can decrease the delay t^1_{delay} to detect fault. Then, as shown in the red box ϱ_1 of Fig. 4.5, the rapidity of the residual r could be adjusted with the aid of the lower bound envelope $r_{i,\ min}$ in (4.7) (a smaller t_1 will give a residual with faster transients). Thus, the fault detection delay t^1_{delay} can be reduced by selecting a suitable lower bound envelope $r_{i,\ min}$.

When the fault disappears, it is also interesting for the observer to detect that there is no fault. The time $t_{disdetec}$, when the observer detects that fault disappears, is defined as:

$$\{t_{disdetec} \mid \|r\left(t_{disdetec}\right)\|_{\text{rms}} \leqslant J_{th}, \|r\left(t_{disdetec} - \xi\right)\|_{\text{rms}} > J_{th}\} \tag{4.9}$$

As shown in the blue boxes ϱ_2 of Fig. 4.5 and Fig. 4.6, the delay to find that the fault disappears can be defined as $t^2_{delay} = t_{disdetec} - t_{disappear}$. On one hand, the delay t^2_{delay} also depends on the transients of residual. On the other hand, through the evaluation with the time window RMS in (3.22), the negative residual in the blue box ϱ_2 of Fig. 4.5 will be positive, as the evaluated residual in the blue box ϱ_2 of Fig. 4.6. Without suitable design, the oscillations of evaluated

residual in the blue box ϱ_2 of Fig. 4.6 may exceed the threshold J_{th}, which will cause false alarm even the fault disappears, and also increase the delay t^2_{delay}.

Remark 4.3. In order to achieve fast fault detection without false alarm when fault disappears, both intervals t^1_{delay} and t^2_{delay} should be considered. As a consequence, minimizing t^1_{delay} and t^2_{delay} becomes an important criterion to design the observer. In fact, for pulse fault signal, the dynamics of residual when fault appears (in the red box ϱ_1 of Fig. 4.5) and disappears (in the blue box ϱ_2 of Fig. 4.5) are upside down. As a result, the transients of the evaluated residual before the oscillations (dynamics in the blue box ϱ_2 of Fig. 4.6) are similar to the transients of the evaluated residual (dynamics of residual in the red box ϱ_1 of Fig. 4.6), which mainly depend on the transients of residual when fault arises (dynamics in the red boxϱ_1 of Fig. 4.5). Thus, the trends of t^1_{delay} and t^2_{delay} are similar when the oscillations in the blue box ϱ_2 of Fig. 4.6 are lower than the threshold J_{th}. In this case, these two criteria t^1_{delay} and t^2_{delay} can be considered as one criterion.

One critical problem is how to set the upper and lower bound envelopes ($r_{i,\,max}$ and $r_{i,\,min}$) to generate a satisfying residual for fast fault detection. From a practical point of view, the residual should be designed to detect the fault faster without any false alarms when the fault disappears. Therefore, under these objectives, it is straight-forward to propose specifications like rise time, peak time, settling time, overshoot, damping, threshold. The connections among the different criteria and different objectives are illustrated in Fig. 4.7 with following notes.

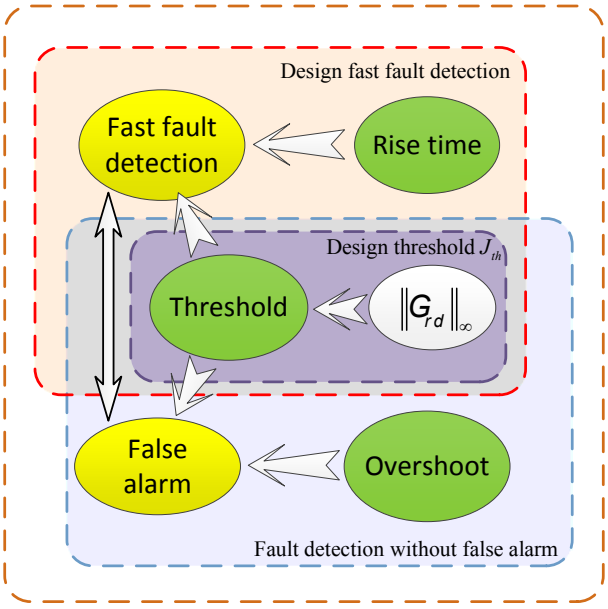

Figure 4.7.: Relationships among different criteria and different design objective

- The factors to affect fast fault detection

  The main point of the fast fault detection design is to minimize the fault detection delay t^1_{delay}. As shown in the red box ϱ_1 of Fig. 4.6, in the view of threshold design, a fast fault detection can be achieved by decreasing the threshold J_{th}. Then, with the definition of threshold in (3.22), the fault detection delay t^1_{delay} could be reduced by minimizing $\|G_{rd}\|_\infty$. On the other hand, a residual with short rise time when fault appears also gives potentials to detect fault faster. With the aid of the time domain specifications in (4.7),

the lower bound envelope $r_{i,\ min}$ could be used to adjust the transients of residual with a small rise time.

- The factors to cause false alarm when fault disappears

  As introduced in the previous part, the false alarm when fault disappears is mainly caused by the evaluation for the negative part of residual in the blue box ϱ_2 of Fig. 4.5. Although a high threshold can filter the oscillations when the fault disappears, the fault detection delay t^1_{delay} is also increased. A better solution is improving the transients of residual to decrease this kind of false alarm. In the view of residual generation, for a step fault signal, the oscillations when fault disappears are mainly caused by the overshoot of residual: a residual with big overshoot always causes big oscillations when fault disappears. Therefore, in order to decrease the false alarm when fault disappears, the overshoot of residual should be restricted, which could be realized by choosing a suitable upper bound envelope $r_{i,\ max}$.

- The relationship between fast fault detection and false alarm when fault disappears

  A fast fault detection needs a residual with a small rise time, which may bring a high overshoot for the responses of residual. However, this kind of high overshoot will increase the false alarm when fault disappears. Therefore, in the integrated design with both objectives, a small rise time with small overshoot will give a good result, which could be achieved by introducing the upper and lower bound envelopes ($r_{i,\ max}$ and $r_{i,\ min}$) simultaneously.

Remark 4.4. As explained in Remark 4.3, if the oscillations are limited, the transients of evaluated residual when fault disappears in Fig. 4.6 are similar to the transients of the residual in Fig. 4.5, whose rapidity is similar to the rapidity of the residual when fault arises. Then, the used lower bound envelope $r_{i,\ min}$ also guarantees a fast transients of the residual when fault disappears to decrease the delay t^2_{delay}.

4.2.3. Settings of the envelopes for two different cases

There are a bit differences of the settings of the envelopes $r_{i,\ min}$ and $r_{i,\ max}$ between the system with $D_f = 0$ and system with $D_f \neq 0$.

- $D_f = 0$

When $D_f = 0$ in (4.2), only the actuator faults are considered. The corresponding transfer function from the fault to the residual G_{rf} will be strictly proper. Considering that if the fault is kind of step signal, the initial value of the corresponding residual responses from the fault f will be zero. In this case, as introduced before, an upper bound envelope $r_{i,\ max}$ is used to eliminate the overshoot of the residual, while the lower bound envelope $r_{i,\ min}$ is employed to improve the fast transients of residual. When the transients of the residual from the fault f is chosen, the fault detection time only depends on disturbances. A fault with smaller disturbances will be detected faster, but the fault detection delay t^1_{delay} will never be 0.

- $D_f \neq 0$

Different from the case of $D_f = 0$, when $D_f \neq 0$ in (4.2), the transfer function G_{rf} is biproper, which will result in a nonzero initial responses of the residual r when the fault f arises. The corresponding value is dependent on the parameters D_f of system. On account of the fact that

the nonzero initial value of response can achieve a zero fault detection delay with some small energy bounded disturbances, the nonzero initial value of response should be utilized as much as possible, which could be realized by setting a constraint of lower bound envelope $r_{i,\ min}$ at $t = 0$. Without suitable design, the residual with a nonzero initial value may go down very fast after achieving the peak. As a result, the time interval of evaluated residual higher than the threshold may be too small to detect fault, which could be avoided by setting a flat $r_{i,\ min}$ $(t_1 \leqslant t \leqslant t_2)$ to obtain enough interval $(t_2 - t_1)$ to detect faults.

4.3. Integrated fast fault detection observer design for general and specific cases

This section introduces the integrated design for general and specific cases, where the design for general case is measured by mixed specifications H_-/H_∞ in frequency domain. For the specific case, two design problems to achieve an ideal fault detection for specific fault are introduced in the previous section: fast fault detection when fault arises and low false alarm rate when fault disappears. The latter design is much stricter than the former design since the upper bound envelope $r_{i,\ max}$ is added into the former design to restrict the overshoot of residual. In the following section, these two different designs are introduced respectively.

4.3.1. Design for fast fault detection

In this case, with the lower bound envelope $r_{i,\ min}$ in time domain, only the constraint of small rise time is considered to achieve fast fault detection for step fault. For the general case, the worst case with $\|G_{rf}\|_-$ and $\|G_{rd}\|_\infty$ is considered for the design.

According to the previous analysis, it is straightforward to obtain the following formulation for integrated design:

i) $A - LC$ is asymptotically stable;

ii) $$\max_{L,Q} \| G_{rf} \|_- = \max_{L,Q} \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}(G_{rf})$$
$$= \max_{L,Q} \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}\left(QD_f + QC(sI - A + LC)^{-1}(B_f - LD_f)\right)$$

iii) $$\min_{L,Q} \| G_{rd} \|_\infty = \min_{L,Q} \sup_{\omega'\in[\omega_1',\omega_2']} \bar{\sigma}(G_{rd})$$
$$= \min_{L,Q} \sup_{\omega'\in[\omega_1',\omega_2']} \overline{\sigma}\left(QD_d + QC(sI - A + LC)^{-1}(B_d - LD_d)\right))^{-1}$$

iv) $r_{i,\ min}(t) \leqslant r_i(L,\ Q,\ t) \quad \underline{t} \leq t \leq \bar{t}$ where $\underline{t}$ and $\bar{t}$ are the starting and ending time points of the simulation in time domain.

Note again that the i), ii) and iii) objectives are proposed for the design of unknown faults and disturbances (general case). The constraint iv) is used to improve the transients of residual for fast fault detection. In the view of fast fault detection for specific fault, the specification iii) should be minimized to decrease the fault detection delay t^1_{delay}. In the design, the specification in time domain is formulated as a constraint, because the lower bound envelope $r_{i,\ min}$ is always fixed in the optimization. Formulating the specifications ii) and iii) as a ratio, the optimization problem could be obtained as:

$$A - LC \text{ is asymptotically stable} \tag{4.10}$$

$$\underset{L,\ Q}{\text{minimize}} \frac{\|G_{rd}(L,\ Q)\|_\infty}{\|G_{rf}(L,\ Q)\|_-} \tag{4.11}$$

$$r_{i,\ min}(t) \leqslant r_i(L,\ Q,\ t) \quad \underline{t} \leq t \leq \bar{t} \tag{4.12}$$

For the fast fault detection, the main objective is to minimize the fault detection delay t^1_{delay} by the design. Because the formulation in (4.7) does not minimize the fault detection delay t^1_{delay} in the direct way, an algorithm to optimize the fault detection delay t^1_{delay} is proposed for the step fault with the formulation in (4.7) first. Then, the minimization of fault detection delay t^1_{delay} is added into the optimization (4.7) to achieve an integrated design for general case and specific case simultaneously.

4.3.1.1. An iterative method to minimize the fault detection delay t^1_{delay} with lower bound envelope $r_{i,\ min}$

As introduced before, the fault detection delay t^1_{delay} is affected by the threshold selection ($\|G_{rd}\|_\infty$) and the transients of residual. Fixing the residual with same transients, fast fault detection can be achieved by minimizing $\|G_{rd}\|_\infty$. By contrast, with a same threshold (same $\|G_{rd}\|_\infty$), a residual with better transients could detect fault faster. However, a smaller $\|G_{rd}\|_\infty$ always results in a residual with unsatisfying transients, while better transients of residual always increase $\|G_{rd}\|_\infty$. A better design of either specification always causes a worse characteristic of the other specification, as a consequence, both cases cannot decrease the fault detection delay t^1_{delay} effectively. Therefore, a design of fast fault detection should consider the minimization of $\|G_{rd}\|_\infty$ and the improvement of the transients of residual simultaneously.

Due to the fact that the minimization of $\|G_{rd}\|_\infty$ and improvement of the transients of residual affect the minimization of fault detection delay t^1_{delay} simultaneously, it is reasonable to fix one of the two factors and then optimize the other factor. Based on this idea, if we fix the specification $\|G_{rd}\|_\infty$ as $\|G_{rd}\|_\infty = 1$ (with the aid of the residual weighting matrix Q, this equation does not introduce any conservativeness into the design), the threshold could be $J_{th} = 1$ with the proposed assumption.

As shown in Fig. 4.8, the threshold is fixed as 1, same as the amplitude of the lower bound envelope $r_{i,\ min}$. Note that $r_{i,\ min}\left(t,\ t^1_{ref}\right)$ means that the lower bound envelope $r_{i,\ min}$ starts at t^1_{ref}, and $t^1_{ref} > t^2_{ref} > t^3_{ref}$. First, with the lower bound envelope $r_{i,\ min}\left(t,\ t^1_{ref}\right)$, we find a solution that the corresponding residual r_1 satisfies the constraint, which means that this design can detect fault at t^1_{ref} with proposed threshold. Then, the lower bound envelope $r_{i,\ min}\left(t,\ t^1_{ref}\right)$ is replaced by $r_{i,\ min}\left(t,\ t^2_{ref}\right)$. If a solution could be found that the corresponding residual r_2 meets the new constraint, the new solution is a better design: the design could detect fault at t^2_{ref}, earlier than the previous design. Then reduce the starting time of the low bound envelope to be t^3_{ref}, and look for a solution to meet the lower bound envelope $r_{i,\ min}\left(t,\ t^3_{ref}\right)$. With this logic, the minimum fault detection delay could be found until the residual could not meet the new low bound envelope $r_{i,\ min}$ with a smaller starting time. Then, an algorithm to minimize the fault detection delay t^1_{delay} could be obtained in Algorithm 4.1.

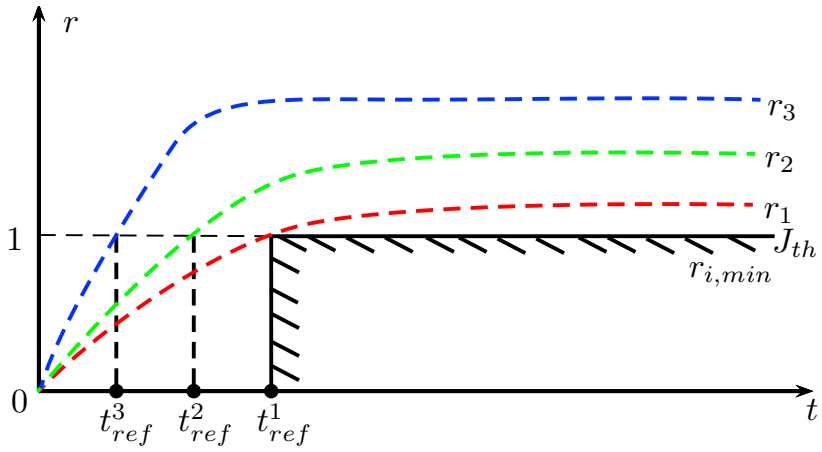

Figure 4.8.: Explanation of the minimization of fault detection delay t^1_{delay}

Algorithm 4.1 Minimization of fault detection delay t^1_{delay}

Step 1. Initialization: set a big step length λ, and choose a small positive value μ;
Step 2. Formulate the initial lower bound envelope $r_{i,\ min}$: select $L_0 = 0$ and Q_0 to satisfy $\|G_{rd}(L_0,\ Q_0)\|_\infty = 1$. The amplitude of the lower bound envelope $r_{i,\ min}$ is set to be lower than the steady value of the step response with L_0 and Q_0, which is also set as the threshold J_{th}. Note the time when the step response exceeds the threshold at the first time as δ, which determines the starting time of the lower bound envelope as $r_{i,\ min}(t,\ \delta)$;
Step 3. Formulate the new lower bound envelope $r_{i,\ min}$: update the lower bound envelope $r_{i,\ min}(t,\ \delta)$ with $\delta = \delta - \lambda$ and same amplitude;
Step 4. Search the solution: try to look for if there is a solution (L and Q) satisfying the following constraints:

$$r_{i,\ min}(t,\ \delta) \leqslant r_i(L,\ Q,\ t) \quad \underline{t} \leq t \leq \bar{t}$$
$$\|G_{rd}(L,\ Q)\|_\infty = 1$$

If the solution exists, go to Step 2. Otherwise, check whether $\lambda > \mu$, if yes, set $\lambda = \lambda/2$ and go to Step 2, otherwise exit.

4.3.1.2. Integrated design for general and specific cases

Generally, there are finite optimal solutions for the general case with (4.11), and there may be only one global optimal solution in some cases. As a result, the ability of fast fault detection for specific fault with the corresponding design is also determined. This ability is always not good, because the specifications for the specific fault are not considered in the design. In order to design general and specific cases together, a trade-off between the ability of fault detection in the worst case and the rapidity of fault detection for specific fault should be considered. A direct way to solve this multiobjective optimization problem is combing these two specifications as one objective function for minimization. The specification for general case is in frequency domain while the criterion for specific case is in time domain. Therefore, the combined objective function contains the specifications in frequency and time domain simultaneously. Unfortunately, the minimization of this kind of objection function is difficult to solve. Alternatively,

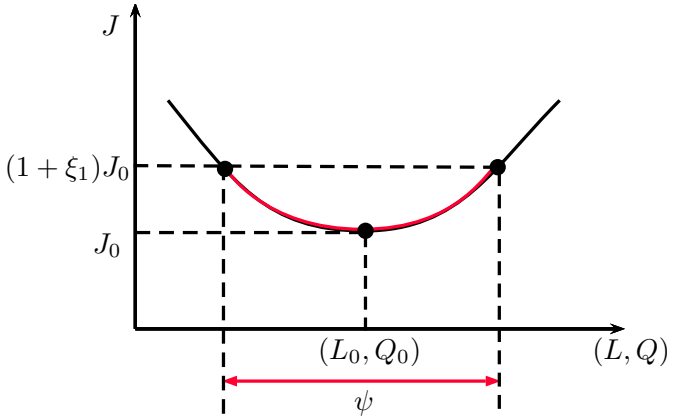

Figure 4.9.: Interpretation of integrated design

it is interesting to know how is the best design for the specific case with a tolerable degradation of the ability of fault detection for the general case, which could be explained with Fig. 4.9: assuming that the minimum value J_0 of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ is achieved with L_0 and Q_0. The integrated design extends the unique optimal design (L_0 and Q_0) for the minimization of $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ to a design region for L and Q with a tolerable degradation of the mixed specification $\|G_{rd}\|_\infty / \|G_{rf}\|_-$, as shown in the red part ψ in Fig. 4.9. The tolerable degradation can be represented by a parameter $\xi_1 = \frac{J_1 - J_0}{J_0}$, where J_1 is the biggest tolerable value of the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$. Then, the design problem is translated to find the solution with fastest fault detection for the specific fault when the design parameters L and Q belong to the design region ψ.

Remark 4.5. According to the definition of ξ_1, we have

$$\xi_1 = \frac{J_1 - J_0}{J_0} = \frac{J_1}{J_0} - 1 \tag{4.13}$$

which also means that

$$J_1 = J_0 (1 + \xi_1) \tag{4.14}$$

As introduced in Example 3.1

$$\frac{AMP_1}{AMP_0} = \frac{J_1}{J_0} \tag{4.15}$$

where AMP_0 and AMP_1 represents the lowest amplitudes of sinusoidal signal fault (the frequency ω_i of the sinusoidal signal meets $\|G_{rf}(L_i, Q_i)\|_- = G_{rf}(j\omega_i, L_i, Q_i)$, $i = 1,\ 2$) when the sinusoidal fault is strongly detectable with L_1, Q_1 and L_2, Q_2 respectively.

Then, the equation (4.13) could be

$$\xi_1 = \frac{J_1}{J_0} - 1 = \frac{AMP_1}{AMP_0} - 1 \tag{4.16}$$

which results in

$$AMP_1 = AMP_0\,(1+\xi_1) \tag{4.17}$$

Therefore, the physical meaning of (4.14) could be explained with (4.17): the degradations of J means the decrease of the set of strongly detectable faults.

With the aid of the iterative algorithm in Algorithm 4.1, the integrated design for general and specific cases could be realized with Algorithm 4.2.

Algorithm 4.2 Integrated fast fault detection observer design

Step 1. Select the interesting frequency ranges Φ for $\|G_{rf}\|_-$ and $\|G_{rd}\|_\infty$ (which could be different), ξ_1 and μ are small positive values, and λ is the adjusted step to change δ;

Step 2. Optimize for the unknown faults and disturbances (general case)

$$\begin{gathered}\underset{L,\ Q}{\text{minimize}}\frac{\|G_{rd}(L,\ Q)\|_\infty}{\|G_{rf}(L,\ Q)\|_-}\\ A-LC \text{ is asymptotically stable}\end{gathered} \tag{4.18}$$

and obtain the corresponding observer gain L_0, residual weighing matrix Q_0 and the minimized value $J_0 = \dfrac{\|G_{rd}(L_0,\ Q_0)\|_\infty}{\|G_{rf}(L_0,\ Q_0)\|_-}$. Then, choose a scalar v to make $Q_0 = vQ_0$, which satisfies $\|G_{rd}\,(L_0,\ Q_0)\|_\infty = 1$. The amplitude of the lower bound envelope $r_{i,\ min}$ is set to be under the steady value of the step response with L_0 and Q_0, which is also set as the threshold J_{th}. Set the time when the step response exceeds the threshold as δ, which determines the starting point of the lower bound envelope as $r_{i,\ min}\,(t,\ \delta)$;

Step 3. Set $\delta = \delta - \lambda$, and add the constraints in time domain for the optimization:

$$\begin{gathered}\underset{L,\ Q}{\text{minimize}}\frac{\|G_{rd}(L,\ Q)\|_\infty}{\|G_{rf}(L,\ Q)\|_-}\\ r_{i,\ min}\,(t,\ \delta) \leqslant |r_i\,(L,\ Q,\ t)| \quad \underline{t} \leq t \leq \bar{t}\\ \|G_{rd}\,(L,\ Q)\|_\infty = 1\\ A-LC \text{ is asymptotically stable}\end{gathered} \tag{4.19}$$

Obtain a minimized value $J_1 = \dfrac{\|G_{rd}(L_1,\ Q_1)\|_\infty}{\|G_{rf}(L_1,\ Q_1)\|_-}$ with designed L_1 and Q_1;

Step 4. If $\frac{J_1-J_0}{J_0} \leqslant \xi_1$, go to Step 3; Otherwise, check whether $\lambda > \mu$, if yes, set $\lambda = \lambda/2$ and go to Step 3, otherwise exit.

4.3.2. Design for fast detection and low false alarm rate

In this case, both the constraints of rise time with the lower bound envelope $r_{i,\ min}$ and overshoot with the upper bound envelope $r_{i,\ max}$ are used to detect step fault fast with a low false alarm rate. As in the previous design, the worst case with $\|G_{rf}\|_-$ and $\|G_{rd}\|_\infty$ is considered for the design of general case.

According to the analysis of fast fault detection and low false alarm rate, the following formulation can be proposed:

i) $A - LC$ is asymptotically stable;

ii) $\max\limits_{L,Q} \parallel G_{rf} \parallel_- = \max\limits_{L,Q} \inf\limits_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}(G_{rf})$

$$= \max_{L,Q} \inf_{\omega\in[\omega_1,\omega_2]} \underline{\sigma}\left(QD_f + QC(sI - A + LC)^{-1}(B_f - LD_f)\right)$$

iii) $\min\limits_{L,Q} \parallel G_{rd} \parallel_\infty = \min\limits_{L,Q} \sup\limits_{\omega'\in[\omega'_1,\omega'_2]} \bar{\sigma}(G_{rd})$

$$= \min_{L,Q} \sup_{\omega'\in[\omega'_1,\omega'_2]} \overline{\sigma}\left(QD_d + QC(sI - A + LC)^{-1}(B_d - LD_d)\right))^{-1}$$

iv) $r_{i,\ min}(t) \leqslant r_i(L,\ Q,\ t) \leqslant r_{i,\ max}(t)$, $\underline{t} \leq t \leq \bar{t}$ where $\underline{t}$ and $\bar{t}$ are the starting and ending time of the simulation in time domain.

The same ratio formulation of the specifications ii) and iii) is adopted to the multiobjective optimization for the general case:

$$\begin{aligned} &\underset{L,\ Q}{\text{minimize}} \frac{\|G_{rd}(L,\ Q)\|_\infty}{\|G_{rf}(L,\ Q)\|_-} \\ &\begin{cases} r_{i,\ min}(t) \leqslant r_i(L,\ Q,\ t) \\ r_i(L,\ Q,\ t) \leqslant r_{i,\ max}(t) \end{cases}, \quad \underline{t} \leq t \leq \bar{t} \\ &A - LC \text{ is asymptotically stable} \end{aligned} \tag{4.20}$$

4.4. Results

4.4.1. Model for simulation

To illustrate the effectiveness of the proposed integrated design methodology, the same example (vehicle lateral dynamic system) introduced in the previous chapter is considered here. In this example, the reference velocity is $v = 7m/s$, and the additive fault in above system is a fault in steering angle (input) measurement with $B_f = B_\delta$ and $D_f = D_\delta$. Thus, the corresponding B_u and D_u are the same as B_f and D_f respectively:

$$\left[\begin{array}{c|c|c} A & B_f & B_d \\ \hline C & D_f & D_d \end{array}\right] = \left[\begin{array}{cc|c|c} -20.7 & -0.46 & 10.1 & -1.4 \\ 21.2 & -27.3 & 63.7 & 0 \\ \hline -144 & 3.74 & 71 & -9.8 \end{array}\right]$$

All the discussed specifications in (4.19) and (4.20) could be formulated in the solver Sdotool. Especially, with selected lower bound envelope $r_{i,\ min}$ or upper bound envelope $r_{i,\ max}$, the specifications in time domain could be formulated as a hard constraint, while the mixed specifications $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ could be formulated with a custom objective module.

All the thresholds in the simulation are based on (4.6), where $\max(\|d\|_{rms})$ could be estimated with the worst disturbances acting on the plant before the simulation.

4.4.2. Simulation of observer design with fast fault detection for general and specific cases

Two designs are given to show the effectiveness of design: one is the optimal design for the worst case (unknown faults and disturbances), and the other design considers an additional constraint of fast fault detection in time domain for step fault.

1. As introduced before, the objective function to optimize for the worst case could be

$$\begin{aligned} &\underset{L,Q}{\text{minimize}} \quad \frac{\|G_{rd}\|_\infty}{\|G_{rf}\|_-} \\ &\text{subject to } A-LC \text{ is asymptotically stable} \end{aligned} \tag{4.21}$$

   where both frequency ranges for $\|G_{rd}\|_\infty$ and $\|G_{rf}\|_-$ are from 0 to infinity for this example.

2. Design for general and specific cases simultaneously with the algorithm in Algorithm 4.2

$$\begin{aligned} &\underset{L,\ Q}{\text{minimize}} \frac{\|G_{rd}(L,\ Q)\|_\infty}{\|G_{rf}(L,\ Q)\|_-} \\ &r_{i,\ min}(t,\ \delta) \leqslant r_i(L,\ Q,\ t) \quad \underline{t} \leq t \leq \bar{t} \\ &\|G_{rd}(L,\ Q)\|_\infty = 1 \\ &A-LC \text{ is asymptotically stable} \end{aligned} \tag{4.22}$$

4.4.2.1. Simulation

As introduced in the previous chapter, since the residual weighting matrix Q is a scalar, there will be no effects of the residual weighting matrix Q on the criteria H_-/H_∞. However, the residual weighting matrix Q will affect the transients of the residual from the step signal fault. In order to compare fault detection delay for different observer, Q is designed to meet $\|G_{rd}(L,\ Q)\|_\infty = 1$. In this case, the threshold for both cases will be the same if the disturbances d are the same.

With the optimization in (4.21), we get

$$\begin{aligned} L_1 = [0.1381, 0.1003]^T; Q_1 = [0.1020] \\ \|G_{rd}\|_\infty / \|G_{rf}\|_- = 0.1392 \end{aligned}$$

Considering the constraint of fast fault detection with lower bound envelope $r_{i,\ min}$, a fast fault detection observer could be obtained:

$$\begin{aligned} L_{fast} = [0.1394, -0.0933]^T; Q_{fast} = [0.1020] \\ \|G_{rd}\|_\infty / \|G_{rf}\|_- = 0.1397 \end{aligned}$$

According to the values of the criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ with the different designed observers, the effects of the integrated fault detection observer L_{fast} and Q_{fast} on the proposed criteria $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ is a little worse than the effects with the optimal design L_1 and Q_1 for general case. In other words, compared with the specific design L_1 and Q_1, the designed fast fault

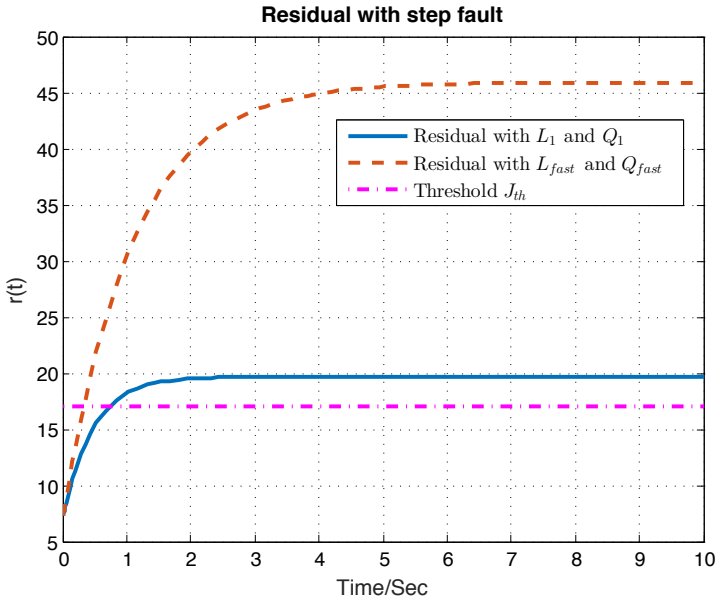

Figure 4.10.: Residual without disturbances (same threshold if disturbances are same)

detection observer L_{fast} and Q_{fast} gives nearly equivalent ability of robustness to disturbances and sensitivity to faults in the worst case.

In order to show the transients of residual from the step fault signal with different observers, the system is simulated without disturbances. With disturbances $d = 0$, Fig. 4.10 shows the transients of the residual responses with step fault signal due to two different observers. Remember that each Q achieves $\|G_{rd}(L,\ Q)\|_\infty = 1$, which also means that these two designs will have the same thresholds if the disturbances d are the same. Based on this condition, the rapidity of fault detection could be compared for these two observers. Since $D_f \neq 0$ in this example, the initial value of residual is not zero, which means that there will be a very small delay to detect fault if disturbance is small. When $J_{th} < 20$, the residual responses with L_{fast} and Q_{fast} is always faster than the design L_1 and Q_1. With the same threshold $J_{th} < 20$, the design L_{fast} and Q_{fast} can detect the step fault faster. When $45 > J_{th} > 20$, the design L_1 and Q_1 cannot detect any step fault any more, which means that the fault detection delay t^1_{delay} in this case is ∞. But the step fault could be detected by the design L_{fast} and Q_{fast} under the same conditions. Therefore, the design L_{fast} and Q_{fast} has better characteristics of fault detection for the step fault than the design L_1 and Q_1 : more sensitivity and faster transients.

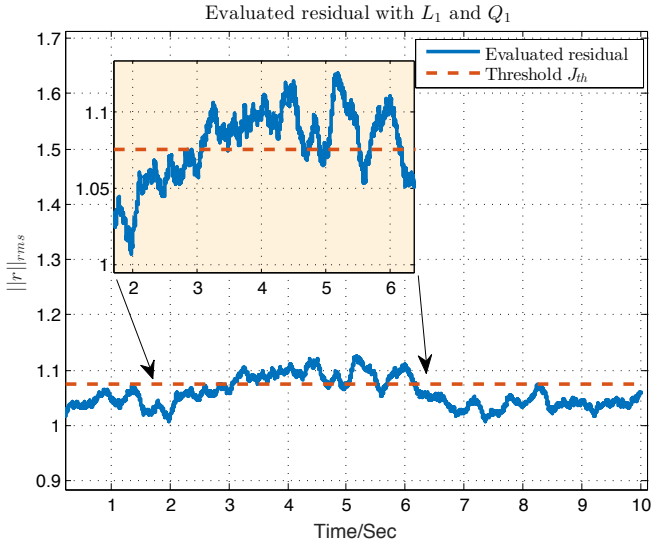

Figure 4.11.: Evaluated residual with L_1 and Q_1

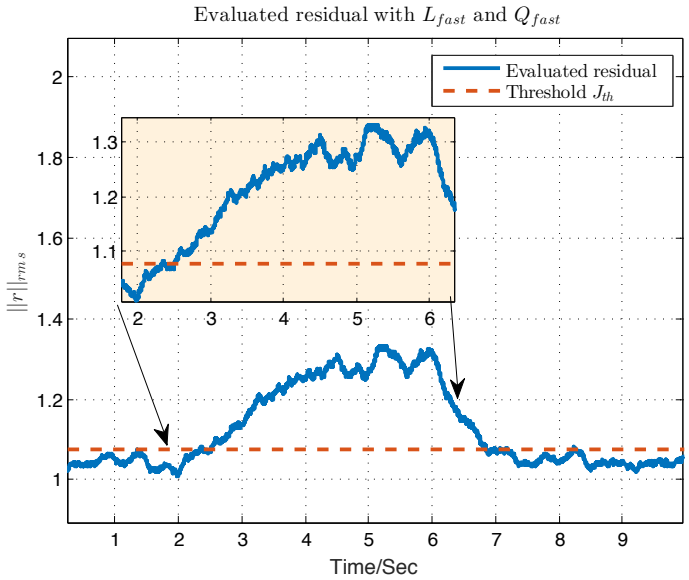

Figure 4.12.: Evaluated residual with L_{fast} and Q_{fast}

The biggest difference of the fault detection delay t^1_{delay} for two observers arises when $J_{th} = 20$ (faults could be detected by both observers), but in this case the step fault may be not detected by the L_1 and Q_1 under the effects of the disturbances d, which means that the fault detection delay t^1_{delay} is sensitive to the transients of disturbances. With Gaussian white noise, it is interesting to do a Monte Carlo simulation to check a statistical fault detection delay. With the amplitude of pulse fault be 0.01, Gaussian white noise with mean $\mu = 0$ and variance $\sigma^2 = 1$, time window for RMS in (3.22) $T = 0.1s$, 1500 Monte Carlo simulations show that the average fault detection delay t^1_{delay} for L_{fast} and Q_{fast} is $0.753s$ while the average fault detection delay t^1_{delay} for L_1 and Q_1 is $1.429s$. The ratio of the average fault detection delay t^1_{delay} between L_1, Q_1 and L_{fast}, Q_{fast} is $1.429/0.753 = 1.897$, which means that the fast fault detection design always can detect the step fault nearly 2 times faster than the design L_1 and Q_1. One of the

Monte Carlo simulation is shown in Fig. 4.11 and Fig. 4.12. The step fault starts at $2s$, and the fault is detected by L_{fast} and Q_{fast} at $2.34s$ while the fault is detected by L_1, Q_1 at $3.07s$.

4.4.3. Simulation of observer design with fast fault detection and low false alarm rate for general and typical cases

In order to show the effects of the observer design with fast fault detection and low false alarm rate, the design of worst case in (4.21) is also used. The evaluation frequency range $\Phi = [\omega_1,\ \omega_2]$ for the fault sensitivity $\|\cdot\|_-$ is specified, which is not full frequency range.

Different from the observer design only with fast fault detection, a constraint of the upper bound envelope $r_{i,\ max}$ is added into the optimization design (4.11) as (4.20).

4.4.3.1. Simulation

The frequency range $\Phi = [\omega_1,\ \omega_2]$ is set as $[0.01rad/s,\ 1rad/s]$ for fault sensitivity criterion H_-. With the H_-/H_∞ frequency design method (4.21) , we can get

$$L_2 = [0.1574, 0.3753]^T; \|G_{rd}\|_\infty / \|G_{rf}\|_- = 0.0329$$

In order to do the comparison with the design in time and frequency domain (4.20), the value of Q_2 is set as 0.002 to restrict the steady value of residual response of step fault signal to be 1.

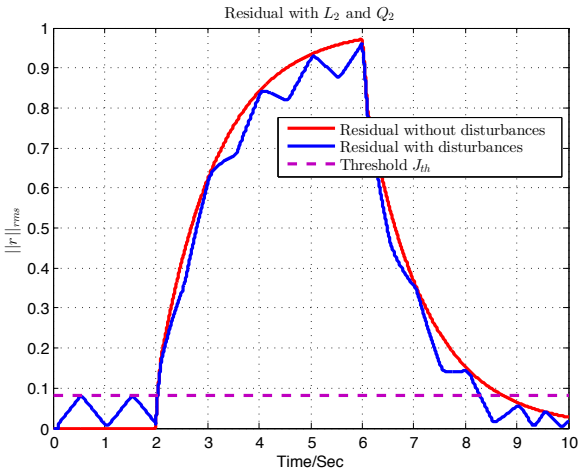

Figure 4.13.: Simulation with L_2 and Q_2

In the simulation, the fault signal is simulated as a pulse of unit amplitude that occurs from 2 to 6 seconds and is zero elsewhere. The disturbances considered in the example is a triangle wave with period T = 1 from 0 to 2. As shown in Fig. 4.13, with the aid of the initial value of the residual response from fault ($D_f \neq 0$), the designed observer L_2 and Q_2 can detect fault very fast (t^1_{delay} is very small) when the disturbances are small. However, when the fault disappears at $6s$, the evaluated residual from $6s$ to $8.3s$ is still bigger than the threshold J_{th}, which means that there are false alarms when fault disappears.

In order to improve the transients of the residual for the step fault, the second observer will consider the effects of short rise time and overshoot and simultaneously minimize the complex

criterion $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ as (4.20). The settings of the upper and lower bound envelopes in time domain are shown in Fig. 4.14. The generated signal could only stay the blank space between the upper and lower bound envelopes.

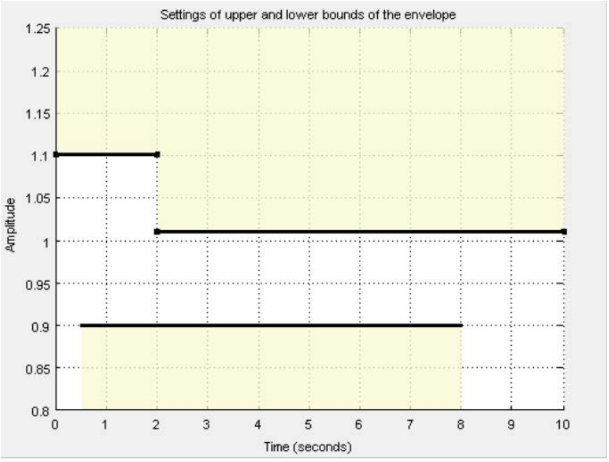

Figure 4.14.: Settings of upper and lower bound envelopes

Through the optimization, we obtain

$$L_3 = [0.1035, -0.5626]^T,$$
$$Q_3 = 0.0086,$$
$$\|G_{rd}\|_\infty / \|G_{rf}\|_- = 0.0870$$

where the residual weighting matrix Q_3 is designed to make the steady value of the residual response from step fault be 1.

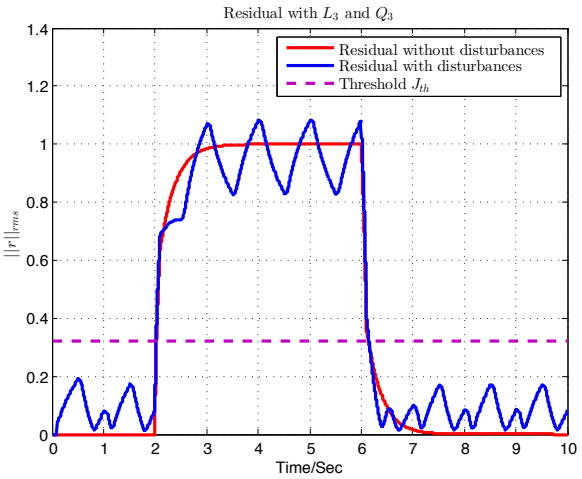

Figure 4.15.: Simulation with L_3 and Q_3

The effects of the design are shown in Fig. 4.15. The dynamics of the residual with L_3 and Q_3 when the fault appears and fault disappears are much better than the design L_2 and Q_2 while the ability of fault detection in the worst case with L_3 and Q_3 is worse than the design with L_2 and Q_2. The good transients of the residual for step signal fault will decrease the

ability of fault detection in the worst case. In practice, similar to the design in fast fault detection, a tolerable degradation of the mixed specification $\|G_{rd}\|_\infty / \|G_{rf}\|_-$ should be chosen first ($\frac{\|G_{rd}(L,Q)\|_\infty}{\|G_{rf}(L,Q)\|_-} < \eta$, and η is a value to express the tolerable degradation of the mixed specification $\|G_{rd}\|_\infty / \|G_{rf}\|_-$), then, the ability of fast fault detection with low false alarm rate could be optimized under the proposed constraint.

4.5. Conclusion

In this chapter, a method has been proposed to design an integrated fault detection observer for general case (unknown $\mathcal{L}_2$ bounded faults and disturbances) and specific case (some specific fixed fault) with performance indexes in frequency and time domain, where the design for general case is evaluated with mixed specifications H_-/H_∞ in frequency domain. With a lower bound envelope $r_{i,\,min}$, the first application of the specification in time domain is to adjust the transients of residual for a specific fault. Considering the effects of threshold selection and transients of residual on the fault detection delay t^1_{delay}, an algorithm is given with an iterative formulation to minimize the fault detection delay t^1_{delay}. Based on this iterative algorithm, an integrated design for general case and specific case is given. Furthermore, in order to reduce the false alarm when fault disappears, the overshoot of residual response is restricted into the design with an upper bound envelope $r_{i,\,max}$. The numerical simulation validates the proposed design methods.

5. Auxiliary signal design for active fault diagnosis

5.1. Introduction

There are two main approaches to achieve fault detection and isolation. One of them is the passive approach, where the detector observes the inputs and outputs of the system, and then tries to make a decision whether a fault has occurred. The inputs could be from the controller, or from the reference inputs to the system. A large amount of work on passive approach has been done in the field of fault detection and isolation [117, 14, 25, 35]. The major disadvantage of passive approach is that the faults could be masked by the operation [22], especially for the controlled systems. The controller is designed to keep the system working at some equilibrium points after faults occur. In other words, the abnormal behaviors of the system will be masked by the effects of the controller. Additionally, some kinds of faults may be hidden under certain operating states and could be found only under some specific conditions. Comparing with the passive approach, the active approach generates a series of auxiliary signals, which are used to excite the system as the inputs signal, to decide whether the system is in the nominal model or faulty model. In this way, the detection of faulty model with the active approach will be faster than the passive approach. The perturbations from the auxiliary signal on the system should be as small as possible.

Some work about active fault diagnosis has been done recently. With an assumption of energy bounded noise, [22, 83] introduced a framework to design the auxiliary signal with the aid of a quadratic optimization approach. A great deal of work has been done to improve the application of this framework: a direct approach for nonlinear system [5] and linear system with model uncertainty [6], a recursive approach for linear discrete time system [38] and a quadratic optimization approach for closed-loop systems [39] are introduced. With the aid of the quadratic optimization, the energy norm H_2 is used to evaluate the effects of auxiliary signal on the system. For parametric faults, [79, 81, 78, 80, 82] proposed another framework of active fault diagnosis based on the YJBK parameterization and the dual YJBK parameterization of an all stabilizing feedback controller, and it could connect with fault tolerant control (FTC). In this framework, a fault signature matrix is generated to reflect the fault directly, which can be activated by injecting a sinusoidal auxiliary signal. A stochastic case is improved in [88]. In this typical formulation, the perturbations of the auxiliary signal on the system do not need to be considered since the fault signature matrix is separated from the system. In the application with this method, it needs to translate the system into the proposed formulation. Using the predicted information, [97] proposed a new unified formulation of the active fault detection and control problem for discrete-time stochastic systems with a closed loop strategy. The relationship between active fault detection and optimal control is given with a stochastic framework. For linear discrete time hybrid systems, [102] introduced an algorithm to generate appropriate test signals based on reach set computation for faulty and normal systems. However, the

optimization in the design may be infeasible when the polytope of the tolerable performance region is used to approximate the tolerable area. In [103], model predictive approach was applied on the active fault diagnosis of hybrid systems. The objective is to find an auxiliary signal with a shortest sequence, and the effects of auxiliary signal on the system is not considered. In order to achieve active fault diagnosis under the pointwise bounded disturbances, [96] designed auxiliary signal based on zonotope. Using the feedback of the closed-loop, [90] introduced two closed-loop approaches to reduce the length (the time duration) and the cost (in the view of norm) of the auxiliary signal. With the aid of the set-valued estimation of the states by set-membership approaches, a method to generate optimal auxiliary signals with a finite time horizon was introduced in [101]. Since the stochastic approach aims to maximize the probability of a correct diagnosis at a certain time, while the deterministic approach aims to guarantee diagnosis within at a certain time, [73] introduced a hybrid stochastic-deterministic approach to design the auxiliary signal. The proposed method guarantees diagnosis at a given time. Same as in ([22]), the energy norm H_2 is used to measure the effects of auxiliary signal on the system in [96, 90, 101, 73]. Considering to the maximum detection performance, minimum system performance degradation and minimum targeted detection time, [20] chose an optimal frequency for the single frequency periodic signals and optimal gain for an estimator. The proposed framework cannot be applied on MIMO system directly.

A disadvantage of active fault diagnosis is that the injection of auxiliary signal will disturb the system. The casued perturbations by auxiliary signal may result in a performance degradation of the system, if the auxiliary signal is not well designed. Hence, the influences of designed auxiliary signal on the system should be as small as possible. The introduced methods in [83, 90, 96] always proposed the energy norm H_2 (A.2 or A.4) to evaluate the effects of the auxiliary signal on the output of the system y:

$$\min \ \|y^v\|_2 \tag{5.1}$$

and the injected signal v should be also as "small" as possible:

$$\min \ \|v\|_2 \tag{5.2}$$

where y^v represents the response from the auxiliary signal v to the output of the system y. A constraint of the auxiliary signal v with pointwise bound for safety or other considerations could be solved by a direct optimization formulation [6, 5]:

$$L_{\min} \leqslant v \leqslant L_{\max} \tag{5.3}$$

where $L_{\min}$ and $L_{\max}$ are the lower and the upper bound of the auxiliary signal in time domain respectively.

However, from a practical point of view, we want to know how are the maximum perturbations on the system caused by the auxiliary signal. In other words, the "worst" effects of the injected signal on the system should be considered, which could be represented by the peak amplitude of the response from the auxiliary signal. In the view of the evaluation function, the energy norm H_2 measures the average effects of the auxiliary signal. The H_2 value of a signal is not affected by the transient. As a result, even if H_2 norm of a signal is small, the signal may have large peaks, which are not too frequent and do not contain too much energy [19]. In this case, a small H_2 value of signal fails to guarantee a small peak amplitude of the signal. Therefore, the objectives (5.1) and (5.2) fail to evaluate the worst effects of the auxiliary signal on the

residual. A criterion to evaluate the worst effects of auxiliary signal on the system should be considered. Furthermore, in the traditional methods, the perturbations of auxiliary signal on the system are represented by the changes on the output y, and the energy of auxiliary signal v is designed to be as small as possible. It is interesting to consider the effects of the auxiliary signal v on the control signal of the plant models. This chapter presents a method of auxiliary signal design for active fault diagnosis in the presence of unknown disturbances. In Section 5.2, the problem of auxiliary signal design is formulated with a new framework based on fault detection theory. Then, with the new active fault diagnosis framework, the discrimination for two models case (also could be called as model detection) is illustrated in Section 5.3. A model discrimination condition in the worst case is given to formulate the auxiliary signal for successful model discrimination. A peak norm, $\|\cdot\|_{\text{peak}}$, is introduced to evaluate the worst effects of the designed auxiliary signal on the system. Different from the classical methods, the influence of auxiliary signal on the control signal is also considered. In order to restrict the transients of the auxiliary signal, a constraint of decay rate is also introduced. Section 5.4 extends the two models discrimination case to the multiple models discrimination case (more than two models to distinguish, called as model isolation) with a discrimination logic. At last, both two models discrimination case and multiple models discrimination case are simulated with a DC motor control system DR300 to illustrate the effectiveness of the proposed design methods.

5.2. Problem formulation

A LTI system includes m true system models and $N-m$ models of various faults:

$$P_i : \begin{cases} \dot{x} = A_i x + B_u^i u + B_d^i d + B_v^i v \\ y = C_i x + D_u^i u + D_d^i d + D_v^i v \end{cases} \tag{5.4}$$

for $i \in \{0, \ldots, N-1\}$. P_i with $i \in \{0, \ldots, m-1\}$ represent the m nominal system models, and P_i with $i \in \{m, \ldots, N-1\}$ mean the $N-m$ faulty models. $x_i \in \mathbb{R}^{n_x}$ is the state vector, $y \in \mathbb{R}^{n_y}$ is the output, $v \in \mathbb{R}^{n_v}$ is the auxiliary signal, $u \in \mathbb{R}^{n_u}$ is the control input, and $d \in \mathbb{R}^{n_d}$ represents noises, perturbations and unmeasured inputs. A_i, B_d^i, B_u^i, C_i, D_d^i, D_u^i are known constant matrices of appropriate dimensions. Particularly, if the auxiliary signal v is injected into the input u directly, matrices B_v^i and D_v^i are the same as B_u^i and D_u^i respectively.

Without loss of generality, disturbance d is assumed to be $\mathcal{L}_2$ norm bounded.

5.2.1. Framework to discriminate models

Different from the logic in [83], a framework of fault detection [25, 35] is proposed to discriminate the models, which is shown in Fig. 5.1. The auxiliary signal v is generated by an auxiliary signal generator. A reference model is used to estimate the output of the system, and the estimated output $\hat{y}$ is compared with the output of the system y to generate a typical signal r, which is used to indicate whether the system stays at the nominal or faulty model. In the ideal case, this signal r should be zero when the system is in the nominal model, while it should be nonzero when the system in the faulty model. Due to the fact that this kind of special signal has the similar characteristics of the residual in the field of fault detection and isolation, the signal r is also called as residual in this chapter.

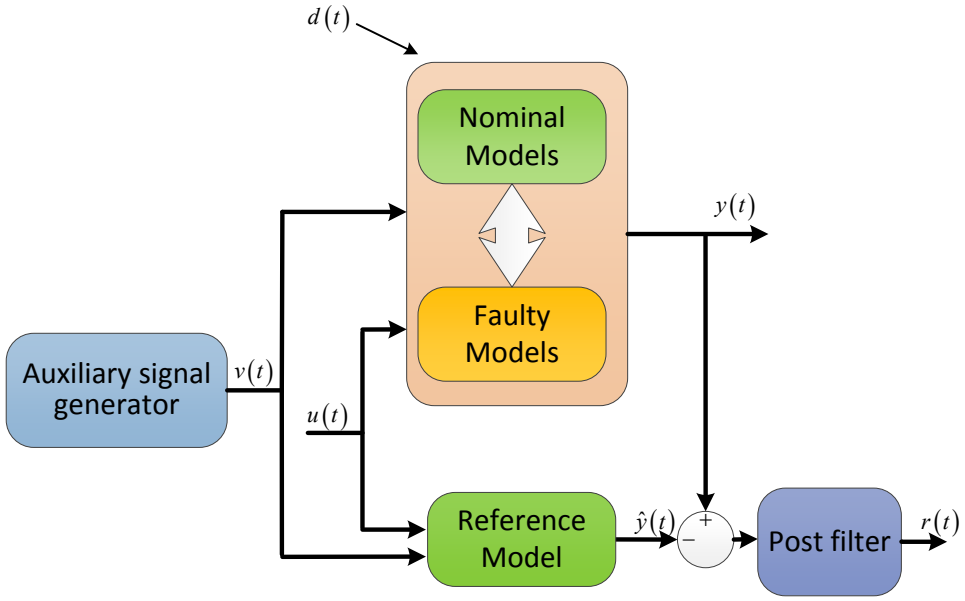

Figure 5.1.: Framework with series of auxiliary signals to discriminate models

In the LTI system, the designed auxiliary signal should be continuous. Then, the auxiliary signal v in Fig. 5.1 can be generated by filtering a reference signal δ with a filter Q:

$$Q = \left[\begin{array}{c|c} A_q & B_q \\ \hline C_q & D_q \end{array}\right] = \begin{cases} \dot{x}_q = A_q x_q + B_q \delta \\ v = C_q x_q + D_q \delta \end{cases} \tag{5.5}$$

where $A_q \in \mathbb{R}^{n_q \times n_q}$, $B_q \in \mathbb{R}^{n_q \times 1}$, $C_q \in \mathbb{R}^{n_v \times n_q}$ and $D_q \in \mathbb{R}^{n_v \times 1}$. Among them, n_q depends on the order of selected Q. Signal δ is the selected reference signal, which could be impulse signal, step signal, etc.

Residual r could be improved with a well designed post filter F, which could have similar formulation as the filter Q in (5.5):

$$F = \left[\begin{array}{c|c} A_F & B_F \\ \hline C_F & D_F \end{array}\right] = \begin{cases} \dot{x}_F = A_F x_F + B_F (y - \hat{y}) \\ r = C_F x_F + D_F (y - \hat{y}) \end{cases} \tag{5.6}$$

Then, the residual r is generated as:

$$r = F (y - \hat{y}) \tag{5.7}$$

The residual r could be unique when the reference model is either an unchanged reference model or a series of switching reference models, and it also can be a bank of residuals when a bank of reference models are used for the "reference model" in Fig. 5.1.

In order to show the effects of controller on active fault diagnosis, the framework in Fig. 5.1 can be extended with an addition of controller. The new framework is shown in Fig. 5.2. In the new framework, the auxiliary signal v is introduced into the system as a kind of perturbations on the output from controller u, and the control signal for the plant system in the closed loop

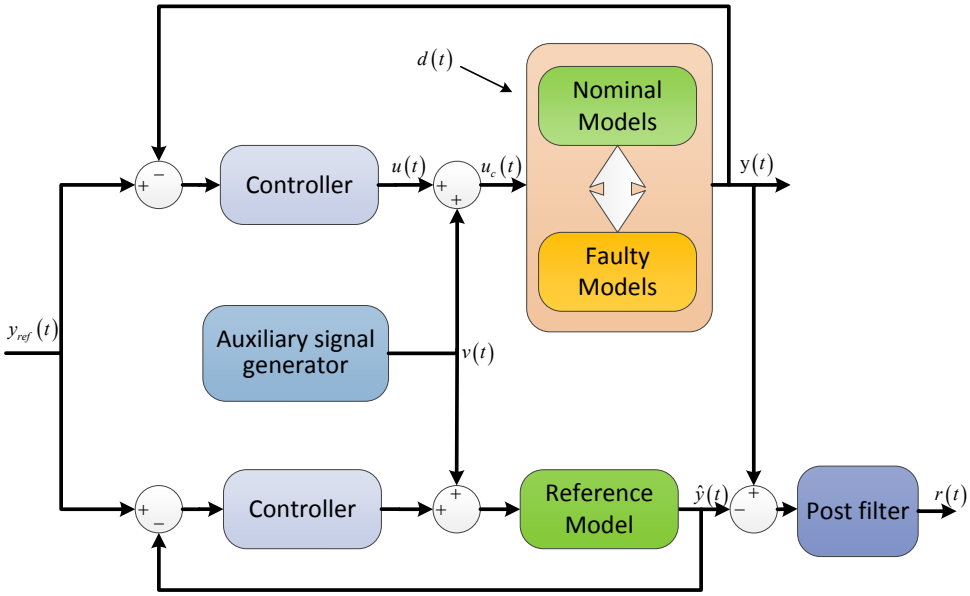

Figure 5.2.: Framework with series of auxiliary signals to discriminate models

will be $u_c = u + v$. In this case, $B_u^i = B_v^i$ and $D_u^i = D_v^i$. Then, the models in (5.4) change to be

$$P_i : \begin{cases} \dot{x} = A_i x + B_u^i (u + v) + B_d^i d \\ y = C_i x + D_u^i (u + v) + D_d^i d \end{cases} \tag{5.8}$$

Owing to the shortages of the H_2 norm to evaluate the effects of auxiliary signal on the system, the worst effects of the auxiliary signal v on the control signal u_c and output y are evaluated with the peak amplitude criterion, $\|\cdot\|_{peak}$ (A.3), in this chapter. Then, the following objectives should be considered:

$$\min \ \|u_c^v\|_{\text{peak}}, \ \min \ \|y^v\|_{\text{peak}}$$

where u_c^v represents the response signal from the auxiliary signal v to the control signal u_c.

Since the auxiliary signal design for two models ($N = 2$) case is a typical case of the active fault diagnosis for multiple models ($N \geqslant 3$) case, the following part starts with the auxiliary signal design for two models case, and then extends to the auxiliary signal design for multiple models case.

5.3. Active fault diagnosis for two models (Model detection)

For two models case, from a FDI point of view, faulty model and nominal model should be distinguished. Assuming that the system is in the nominal model at the beginning, the objective of design is to find whether a switching from nominal model to faulty model arises or not. In the following part, the active fault diagnosis focuses on this switching.

5.3.1. Residual generation with auxiliary signal

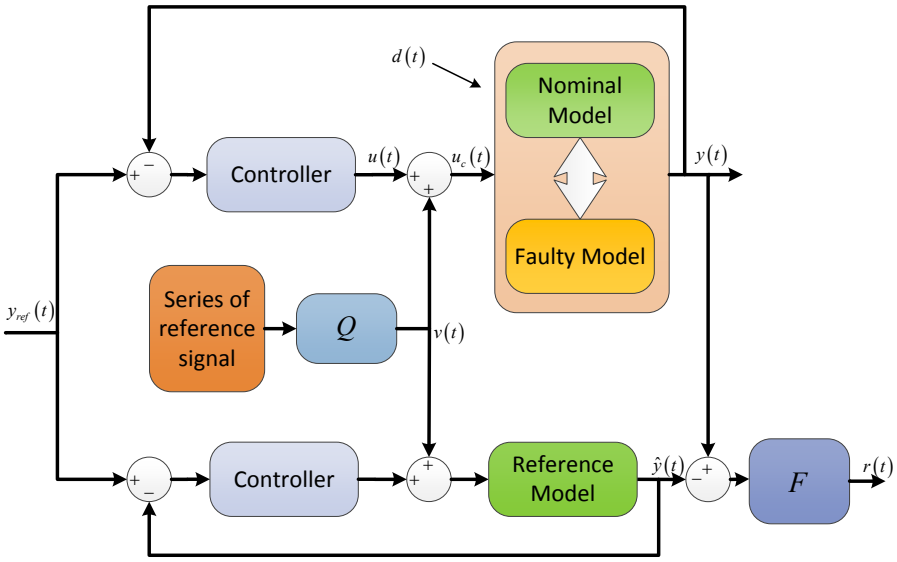

Figure 5.3.: Active fault diagnosis of two models with auxiliary signal

For the two models case, there is only one nominal model P_0 and one faulty model P_1. As shown in Fig. 5.3, the system could be either in nominal model P_0 or in faulty model P_1. The ideal nominal model without the part of the disturbances (P_0 with $d = 0$) is selected as the unique reference model:

$$G_{ref} : \begin{cases} \dot{\hat{x}} = A_0\hat{x} + B_u^0 (u + v) \\ \hat{y} = C_0\hat{x} + D_u^0 (u + v) \end{cases} \tag{5.9}$$

The framework in Fig. 5.3 could be applied for MIMO system and SISO system. For the simplicity of expression, the SISO case is used to express the methodology of active fault diagnosis with this new framework. Note G_0 and G_1 as the transfer functions from control signal u_c to output y when system is in nominal model P_0 and faulty model P_1 separately. In this case, transfer function G_0 will be the same as transfer function G_{ref}. G_c represents the transfer function of controller.

Therefore, the formulation of the residual could be (since the reference model is fixed, an assumption with zero initial state $\hat{x}(0)$ of the reference model in (5.9) does not introduce any conservativeness):

- If system is in nominal model $P = P_0$ (The system is in P_0 at the beginning, and assume that the initial state $x(0) = 0$ when the system starts to work):

$$\begin{aligned} r &= \mathfrak{L}^{-1} \left(F (y - \hat{y}) \right) \\ &= \mathfrak{L}^{-1} \left(F G_{yd}^0 d \right) = r^0 \end{aligned} \tag{5.10}$$

  where

$$G_{yd}^0 = C_0 (sI - A_0)^{-1} B_d^0 + D_d^0 \tag{5.11}$$

- If the system switches to faulty model $P = P_1$:

$$\begin{aligned} r &= \mathfrak{L}^{-1}\left(F\left(y-\hat{y}\right)\right) \\ &= \mathfrak{L}^{-1}\Bigg(F\left(\frac{G_cG_1}{1+G_cG_1}-\frac{G_cG_0}{1+G_cG_0}\right)y_{ref}+F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v \\ &\qquad + FG_{yd}^1 d\Bigg)+\Psi_1\left(x\left(t\right)\right) \\ &= r^1 \end{aligned} \tag{5.12}$$

  where

$$G_{yd}^1 = C_1\left(sI-A_1\right)^{-1}B_d^1+D_d^1 \tag{5.13}$$

  and $v = Q\delta$, $\Psi_1\left(x\left(t\right)\right)$ is the natural response caused by the switching from nominal model to faulty model with the initial state $x\left(t\right)$ when switching occurs.

The differences between r^0 and r^1 should be used to separate the nominal model and faulty model by designing the auxiliary signal generator Q and the post filter F. As shown in equations (5.10) and (5.12), residual r^0 is only affected by disturbances d. By contrast, residual r^1 is affected by four parts: reference input part y_{ref}, auxiliary signal part v, the disturbances part d and initial state part $x\left(0\right)$ during the switching.

- The effects of the reference input part y_{ref}: for the controlled systems, the controller G_c is designed to keep the system working at some equilibrium points when the behavior of system changes. In other words, the controller is designed to decrease the differences when the system switches to the faulty model from the nominal model. As shown in Fig. 5.3, the output of the system y will be equal to the reference input y_{ref} because of the feedback. In this way, the switching (or fault) is hidden if the controller is well designed. It is possible to use the transients of the part $F\left(\frac{G_cG_1}{1+G_cG_1}-\frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$ to find whether there is a switching in the system. However, the changes of the part $F\left(\frac{G_cG_1}{1+G_cG_1}-\frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$ may be covered by the unknown disturbances. Furthermore, reference input y_{ref} always changes with the objective of control, and it cannot be changed with the objective of model discrimination in the active way.
- The effects of the disturbances part d: in some cases, the differences between G_{yd}^0 and G_{yd}^1 could be used to discriminate two models with some kinds of typical disturbances. However, this kind of discrimination is not reliable because of the unknown characteristics of the disturbances, which will be detailed in the next part.
- The effects of auxiliary signal part v: signal v is generated by the reference signal δ and the filter Q in (5.5). With a suitable design of filters F and Q, residual r^1 could be separated from residual r^0.
- The effects of transients caused by switching ($\Psi_1\left(x\left(0\right)\right)$): Since the common state vector x in (5.4) is used for nominal model and faulty model, the state vector x will not jump during the switching but changes smoothly. For the reason that the transients of $\Psi_1\left(x\left(0\right)\right)$ affect r^1 (5.12) but do not affect r^0, this part could be used to separate r^1 from r^0. However, this discrimination is also unreliable, which will be detailed in the next part.

In order to discriminate nominal model P_0 and faulty model P_1 with the difference between r^1

(5.12) and r^0 (5.10), a method to distinguish between r^1 and r^0 should be provided. With the aid of the function of the threshold in fault detection field, a suitable selected threshold could be used to distinguish r^1 and r^0.

5.3.2. Logic to separate r^1 from r^0 with a threshold

Due to fact that r^0 is only affected by the disturbances, an estimation of the maximum effects of disturbances on the nominal model P_0 could be an appropriate candidate to determine the threshold:

$$J_{th} = \sup_{d} \left\| r^0 \right\|_{\text{rms}} \tag{5.14}$$

where the discussed time window RMS (3.22) is used to evaluate the residual r.

According to the RMS norm relationship [19], we have

$$\begin{aligned} \left\| r^0 \right\|_{\text{rms}} &= \left\| \mathfrak{L}^{-1} \left(F G_{yd}^0 d \right) \right\|_{\text{rms}} \\ &\leqslant \left\| F G_{yd}^0 \right\|_{\infty} \left\| d \right\|_{\text{rms}} \\ &\leqslant \left\| F G_{yd}^0 \right\|_{\infty} \cdot \max \left(\left\| d \right\|_{\text{rms}} \right) \end{aligned} \tag{5.15}$$

As a consequence, a threshold can be obtained

$$J_{\text{th}} = \left\| F G_{yd}^0 \right\|_{\infty} \cdot \max \left(\left\| d \right\|_{\text{rms}} \right) \tag{5.16}$$

which is dependent on the dynamics of G_{yd}^0 and the design of the post filter F.

Then, a discrimination conclusion can be obtained with the following discrimination logic:

$$\begin{cases} \left\| r \right\|_{\text{rms}} > J_{th} & \text{Faulty model} \\ \left\| r \right\|_{\text{rms}} < J_{th} & \text{Nominal model} \end{cases} \tag{5.17}$$

5.3.3. Model discrimination without auxiliary signal

Without auxiliary signal, the residual will be affected by the transients $\Psi_1 \left(x \left(0 \right) \right)$ caused by the switching, reference inputs y_{ref} and the disturbances d. Since the characteristics of these three factors are different, the possibilities of model discrimination with these three factors are analyzed respectively.

- Model discrimination only with disturbances part

  In this case, with the discrimination logic (5.17) and threshold (5.16), the discrimination depends on

  $$\left\| r^1 \right\|_{\text{rms}} = \left\| \mathfrak{L}^{-1} \left(F G_{yd}^1 d \right) \right\|_{\text{rms}} > J_{th} \geqslant \left\| r^0 \right\|_{\text{rms}} = \left\| \mathfrak{L}^{-1} \left(F G_{yd}^0 d \right) \right\|_{\text{rms}} \tag{5.18}$$

  Therefore, it is possible to discriminate two models when the part $F G_{yd}^1 d$ in r^1 is big enough to make the evaluated residual $\left\| r^1 \right\|_{\text{rms}}$ exceed the threshold J_{th} by suitably choosing the filter F. However, this kind of design cannot guarantee a successful model dis-

crimination in all cases. For example, let us consider two models of a system having the same characteristics of the bode diagram in high frequency range, but different characteristics in low frequency range. If low frequency disturbances attack the system, the effects of $G_{yd}^1 d$ in r^1 will be different from $G_{yd}^0 d$ in r^0, then it may permit to separate the faulty model from the nominal model even there is no auxiliary signal. However, if the disturbances are in the high frequency, the effects of $G_{yd}^0 d$ in r^0 and $G_{yd}^1 d$ in r^1 will be the same. As a result, the discrimination fails. In this case, the model discrimination depends on the knowledge of the disturbances; however, the disturbances are always unknown in reality. Therefore, the discrimination only with disturbances needs to wait for some typical disturbances to achieve model discrimination.

- Model discrimination with disturbances and reference input parts

  With the addition of reference input part, the discrimination is more complex. The faulty model P_1 could be separated from the nominal model P_0 only when

$$\left\|r^1\right\|_{\text{rms}} = \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref} + FG_{yd}^1 d\right)\right\|_{\text{rms}} > J_{th} \tag{5.19}$$

  where the discrimination depends on the dynamics of disturbances d and effects of the reference input y_{ref} on r^1.

  If $F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$ satisfies

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}\right)\right\|_{\text{rms}} > \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1 d\right)\right\|_{\text{rms}} + J_{th} \tag{5.20}$$

  we have

$$\begin{aligned}&\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref} + FG_{yd}^1 d\right)\right\|_{\text{rms}}\\ \geqslant &\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}\right)\right\|_{\text{rms}} - \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1 d\right)\right\|_{\text{rms}} > J_{th}\end{aligned}$$

  which means that a reference input y_{ref} meeting (5.20) can discriminate nominal model and faulty model with any disturbances. This discrimination depends on the reference input y_{ref}, which is chosen with the requirements of control.

  Furthermore, if the part $F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$ is small, the disturbances part is more important for the residual. The transients of $F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$ may increase or decrease the effects of disturbances on r^1 because of the unknown characteristics of disturbances. That's the reason that the disturbances, which meet the discrimination condition in (5.18), may not discriminate models during the transients of $F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref}$. As a consequence, the discrimination is unreliable, which depends on reference input y_{ref} and dynamics of unknown disturbances d.

- The effects of switching on model discrimination

With the addition of $\Psi_1(x(0))$, model discrimination can be achieved when

$$\left\|r^1\right\|_{\text{rms}} = \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_cG_1}{1+G_cG_1} - \frac{G_cG_0}{1+G_cG_0}\right)y_{ref} + FG_{yd}^1 d\right) + \Psi_1(x(0))\right\|_{\text{rms}} > J_{th} \tag{5.21}$$

The part of $\Psi_1(x(0))$ directly implies that a switching from nominal model P_0 to faulty model P_1 occurs. However, the effects of $\Psi_1(x(0))$ on the residual r^1 may be covered by disturbances and the responses caused by the reference input y_{ref} part. Furthermore, the part $\Psi_1(x(0))$ depends on the states of system when switching arises. Therefore, this discrimination is also unreliable.

All the discussed discrimination cases cannot guarantee a successful model discrimination, which depend on the unknown disturbances d, the selected reference input y_{ref} and the initial state when switching arises. Since the model discrimination without auxiliary signal is unreliable and even unavailable for some cases, it is necessary to generate an auxiliary signal to achieve model discrimination.

Since transients $\Psi_1(x(0))$ during the switching is beneficial to model discrimination, the effects of $\Psi_1(x(0))$ on the residual r^1 should be utilized in the design of auxiliary signal v. However, if there is no information about initial state $x(0)$ when switching occurs, the transients of the residual during the switching are difficult to use. An alternative way is to achieve active fault diagnosis with auxiliary signal v during the steady state after switching. A shortage of this setting is that this kind of auxiliary signal may not achieve an successful model discrimination during the switching. In this case, a series of auxiliary signals with a fixed period could be injected into the system, then the faulty model could be distinguished until an alarm arises. This chapter considers to design the auxiliary signal v in the steady state of residual. In following part, the transient part $\Psi_1(x(0))$ in (5.12) will not be considered in the design of auxiliary signal.

Remark 5.1. In the determination of results with the discrimination logic in (5.17), the alarm does not need to be caused by the auxiliary signal. For the two models case, any alarm implies that the system switches from the nominal model to the faulty model. By contrast, for the multiple models case ($N > 2$), not every alarm means that a switching arises, which will be explained in the multiple models part.im

Remark 5.2. In the previous analysis, disturbances d are assumed to be unknown. However, in most cases, disturbances d are not completely unknown. If there is some prior information on disturbances d, this kind of prior information should be contained into the system to design the auxiliary signal. The disturbances with prior information can be formulated by filtering a white noise with a filter. A new system could be obtained by combining the original system and the filter together. Then, the design of the original system with prior information of disturbances is equivalent to the design of the new formulated system with white noise.

5.3.4. Model discrimination with auxiliary signal

Since the transients of reference input y_{ref} affect the residual r^1, this part should be considered into the design of auxiliary signal. If y_{ref} is known, the reference input y_{ref} could be combined into the design of auxiliary signal v with following remark directly.

Remark 5.3. According to the residual r^1 (5.12) in the steady state after switching ($\Psi_1(x(0)) = 0$), we have

$$\begin{aligned} r^1 &= \mathfrak{L}^{-1}\left(F\left(y-\hat{y}\right)\right) \\ &= \mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)\left(G_cy_{ref}+v\right)+FG_{yd}^1d\right)+\Psi_1\left(x\left(0\right)\right) \\ &= \mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v'+FG_{yd}^1d\right)+\Psi_1\left(x\left(0\right)\right) \end{aligned}$$

We can design v' first to achieve model discrimination. Then, the auxiliary signal can be obtained by $v = v' - G_cy_{ref}$. Therefore, the design of v is equivalent to the design of v'. For the simplicity of expression, in the design of auxiliary signal in the following part, the reference input y_{ref} part in (5.12) is omitted.

Due to the existence of FG_{yd}^0d, the faulty model is discriminated only if the evaluated residual surpasses the threshold J_{th} when the system is in the faulty model. In the case of "without any knowledge on the disturbances", the designed auxiliary signal should guarantee the following inequality in the worst case of FG_{yd}^1d:

$$\inf_d \left\|r^1\right\|_{\text{rms}} > J_{\text{th}} \tag{5.22}$$

Note that

$$\begin{aligned} \inf_d \left\|r^1\right\|_{\text{rms}} &= \inf_d \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v+FG_{yd}^1d\right)\right\|_{\text{rms}} \\ &\geqslant \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}} - \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1d\right)\right\|_{\text{rms}} \end{aligned} \tag{5.23}$$

As a result, the faulty model could be detected for any possible disturbances d if

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}} > \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1d\right)\right\|_{\text{rms}} + J_{th} \tag{5.24}$$

where

$$J_{\text{th}} = \sup_d \left\|r^0\right\|_{\text{rms}} = \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^0d\right)\right\|_{\text{rms}} \tag{5.25}$$

Then we can get

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}} > \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^0d\right)\right\|_{\text{rms}} \tag{5.26}$$

$$+ \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1d\right)\right\|_{\text{rms}} \tag{5.27}$$

With the definition of the threshold in (5.16):

$$\sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1d\right)\right\|_{\text{rms}} \leqslant \left\|FG_{yd}^1\right\|_\infty \max\left(\left\|d\right\|_{\text{rms}}\right) \tag{5.28}$$

$$\sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^0d\right)\right\|_{\text{rms}} \leqslant \left\|FG_{yd}^0\right\|_\infty \max\left(\left\|d\right\|_{\text{rms}}\right) \tag{5.29}$$

Therefore, equation (5.26) could be met when

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}} > \left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)\max\left(\|d\|_{\text{rms}}\right) \tag{5.30}$$

Introducing one quantity φ_{detec} to evaluate the sufficient condition for model discrimination, an auxiliary signal is strong detectable in the presence of disturbances if

$$\begin{aligned}\varphi_{\text{detec}} &= \frac{\sup\limits_d\left\|FG_{yd}^0 d\right\|_{\text{rms}}+\sup\limits_d\left\|FG_{yd}^1 d\right\|_{\text{rms}}}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}} \\ &\leqslant \frac{\left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)\max\left(\|d\|_{\text{rms}}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}} \\ &< 1\end{aligned} \tag{5.31}$$

If $\varphi_{detec} \geqslant 1$, designed auxiliary signal can be used to discriminate models in the presence of disturbances, however, the design will fail to discriminate models in the worst case.

Remark 5.4. In order to discriminate models for all possible disturbances, the quantity φ_{detec} should be smaller than 1 for any possible disturbances d. Since the part $\max\left(\|d\|_{\text{rms}}\right)$ can be estimated with the worst disturbances acting on the plant off-line, the model discrimination condition (5.31) could be transformed to

$$\begin{aligned}\varphi_{\text{detec}} &= \frac{\left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)\max\left(\|d\|_{\text{rms}}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}} \\ &= \frac{\left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}/\max\left(\|d\|_{\text{rms}}\right)} \\ &= \frac{\left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v/\max\left(\|d\|_{\text{rms}}\right)\right)\right\|_{\text{rms}}} \\ &= \frac{\left(\left\|FG_{yd}^0\right\|_\infty+\left\|FG_{yd}^1\right\|_\infty\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v'\right)\right\|_{\text{rms}}} < 1\end{aligned} \tag{5.32}$$

where $v' = v/\max\left(\|d\|_{\text{rms}}\right)$. Once an auxiliary signal v' is found in the design, the auxiliary signal used in practice could be $v = v'\max\left(\|d\,(t)\|\right)$ when the estimation of $\max\left(\|d\|_{\text{rms}}\right)$ is obtained.

5.3.5. Some practical aspects in optimization

After introducing the model discrimination condition of auxiliary signal, this part gives some notes about the operations in practice to design auxiliary signal with above discussed framework: the selection of reference signal, calculation of $\|u_c^v\|_{peak}$ and $\|y^v\|_{peak}$ in time domain, and a constraint to restrict the transients of auxiliary signal.

5.3.5.1. Selection of reference signal $\delta(t)$ for application

In order to decrease the conservativeness of design, it is necessary to select a reference signal such that arbitrary signal could be generated with the filter Q in (5.5). On the other hand, this reference signal should be easy to obtain. The impulse signal has the characteristic that all the frequency spectrum is flat, which provides equivalent weights for the design at any frequency. With a suitable selection of filter Q for the impulse signal, the produced auxiliary signal could be any signal in theory.

The impulse signal, also known as Dirac impulse, is defined by

$$\delta(t - t_0) = \begin{cases} \infty & t = t_0 \\ 0 & t \neq t_0 \end{cases} \tag{5.33}$$

where t_0 represents the time when the impulse signal occurs.

However, from the practical point of view, it is impossible to find such an ideal impulse signal. An alternative method is to approximate the impulse signal as

$$h_\Delta(t - t_0) = \begin{cases} 1/\Delta & \text{if } t_0 < t < t_0 + \Delta \\ 0 & \text{otherwise} \end{cases} \tag{5.34}$$

which could be

$$H_\Delta(s) = \mathfrak{L}^{-1}(h_\Delta(t))$$

In the view of application, periodic auxiliary signal, which contains a series of auxiliary signal with a fixed period, is more useful in the monitoring of system. This could be realized with a series of approximated impulse signal h_Δ:

$$\sigma(t) = \begin{cases} 1/\Delta & \text{if } kT_1 < t < kT_1 + \Delta \\ 0 & \text{otherwise} \end{cases} \tag{5.35}$$

where $k \in \{0, 1 \ldots\}$ and T_1 is the period of auxiliary signal injection.

5.3.5.2. Calculation of other specifications

According to Fig. 5.3, the response signal from auxiliary signal to the control signal u_c can be represented as

$$u_c^v = \left(\frac{1}{1 + G_c G_i} Q\right) H_\Delta$$

and the response signal from auxiliary signal to the output of the system y are

$$y^v = \left(\frac{G_i}{1 + G_c G_i} Q\right) H_\Delta$$

As introduced in the previous part, the constraint of "disturbing the system as little as possible" could be transformed to minimize the criteria $\|u_c^v\|_{\text{peak}}$ and $\|y^v\|_{\text{peak}}$ simultaneously. Since both

specifications have to be minimized, a combined formulation for two criteria could be considered:

$$\min_{Q} \left(\|u_c^v\|_{\text{peak}} + \alpha \|y^v\|_{\text{peak}} \right) \tag{5.36}$$

Because the norm $\|\cdot\|_{\text{peak}}$ is evaluated in time domain directly, the interval ($\underline{t} \leq t \leq \bar{t}$) of simulation to estimate $\|\cdot\|_{\text{peak}}$ should be chosen in advance, which can not be infinite in practice. This implement to evaluate $\|\cdot\|_{\text{peak}}$ may introduce a problem that the peak of signal is small during the simulation interval ($\underline{t} \leq t \leq \bar{t}$), but it is unstable in the infinite time range. In order to overcome this shortage, a constraint of stability for the approximated signal h_Δ should be added. Since the approximated signal h_Δ consists of two opposite step signals, the stability for the approximated signal h_Δ can be achieved by adding a bound to the peak value of the step responses:

$$\left\| \mathfrak{L}^{-1} \left(\frac{1}{s} \frac{G_i}{1 + G_c G_i} Q \right) \right\|_{\text{peak}} \leqslant \xi \tag{5.37}$$

where ξ is a constant value.

Using the periodic impulse signals $\sigma(t)$ in (5.35), the period T_1 of $\sigma(t)$ should be carefully selected. A big period T_1 will delay the time to detect faulty model when the switching from nominal model to faulty model arises. By contrast, a small period T_1 will cause a problem when the convergence rate of the output responses y from auxiliary signal is not high: if the output responses of the present auxiliary signal do not decay to be zero yet, the output responses of a new auxiliary signal will be added into the rest of output responses of the present auxiliary signal . In this case, the output responses caused by auxiliary signals will accumulate, and they become bigger when more auxiliary signals are injected into the system, which means that the perturbations on the system caused by auxiliary signals will also be more. This phenomenon will cause performance degradation of the system, or even damage the system. Therefore, a constraint of convergence rate on the output responses of auxiliary signal should be added into the design:

$$\max(\text{real}(\text{eig}(\frac{G_i}{1 + G_c G_i} Q))) \leqslant \varsigma \tag{5.38}$$

where ς is a negative value, which is selected to meet the requirements of the rapidity of auxiliary signal. The period T_1 in (5.38) should bigger than the settling time of output responses caused by the designed auxiliary signal with (5.38).

5.3.6. Optimization problem with all proposed constraints

Then, we have following required specifications to design a strongly detectable auxiliary signal for model discrimination:

1. Minimizing the worst effects of designed auxiliary signal on the control signal u_c and the output y:

$$\min_{Q} \left(\max_{i=0,1} \left(\left\| \mathfrak{L}^{-1} \left(\frac{1}{1 + G_c G_i} Q H_\Delta \right) \right\|_{\text{peak}} \right) + \max_{i=0,1} \left(\left\| \mathfrak{L}^{-1} \left(\frac{G_i}{1 + G_c G_i} Q \right) \right\|_{\text{peak}} \right) \right) \tag{5.39}$$

2. Discrimination condition of auxiliary signal:

$$\frac{\left(\left\|FG_{yd}^{0}\right\|_{\infty}+\left\|FG_{yd}^{1}\right\|_{\infty}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)QH_{\Delta}\right)\right\|_{\mathrm{rms}}}<1 \tag{5.40}$$

3. Stability of the responses on the output y for the approximated signal h_{Δ}:

$$\left\|\mathfrak{L}^{-1}\left(\frac{1}{s}\frac{G_i}{1+G_cG_i}Q\right)\right\|_{\mathrm{peak}}\leqslant\xi_i \tag{5.41}$$

4. Constraints of fast decay transients of the designed auxiliary signal:

$$\max(\mathrm{real}(\mathrm{eig}(\frac{G_i}{1+G_cG_i}Q)))\leqslant\varsigma \tag{5.42}$$

5. Stability of the filter F:

$$F \text{ should be stable} \tag{5.43}$$

From the optimization point of view, the peak norm $\|\cdot\|_{\mathrm{peak}}$ is calculated in time domain, and the (sub)gradients of this specification is difficult to obtain. As introduced in chapter 2, it is more suitable to solve this kind of specification by some derivative-free methods, such as genetic algorithm. In the optimization with genetic algorithm, the peak norm in (5.39) and (5.41) could be estimated with the command "lsim" in matlab. In practice, it is not necessary to guarantee an alarm in all duration of injection of auxiliary signal. An alarm caused by the some part of the residual response caused by auxiliary signal is enough. Therefore, the discrimination condition (5.40) could be changed to

$$\frac{\left(\left\|FG_{yd}^{0}\right\|_{\infty}+\left\|FG_{yd}^{1}\right\|_{\infty}\right)}{\max\limits_{t}\left(\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)QH_{\Delta}\right)\right\|_{\mathrm{rms}}\right)}<1 \tag{5.44}$$

where the calculation of the time window RMS (3.22) should be also in time domain. All the above objective functions and constraints can be formulated as the form in genetic algorithm, and the needed filters Q and F could be obtained to achieve two models discrimination.

However, for the multiple models case, there will be $N>2$ models to discriminate. In this case, only one reference model is not enough anymore, because the decision logic (3.28) only can separate two models. Therefore, a bank of reference model should be considered, which will introduce more than one threshold. Different from the two models case, too many combinations of thresholds and reference models for the multiple models case increase the difficulty to achieve model discrimination. In the following part, a decision logic to achieve model discrimination for multiple models case is introduced.

5.4. Active fault diagnosis for multiple models (Model isolation)

For the multiple models case P_i (i:$0, \ldots, N-1$, $N > 2$) , there are a number of variations:

1. One nominal model P_0 and $N-1$ faulty models P_i (i:$1, \ldots, N-1$, $N > 2$);

2. m nominal models P_i (i:$0, \ldots, m-1$, $N > m \geqslant 2$) and $N-m$ faulty models P_i (i:$m, \ldots, N-1$, $N > 2$).

In the view of FDI, it is important to detect the switching from nominal model to faulty model. In the first case, there is only one nominal model and $N-1$ faulty models. With the same framework for two models case in Fig. 5.3, the nominal model P_0 without the disturbances part ($d = 0$) is considered as the reference model. Assuming that the system is in the nominal model P_0 at the beginning, any alarm caused by any factor (disturbances part, reference inputs part, the natural response caused by switching or the auxiliary signal part) implies that the system moves to one of the faulty models from the nominal model. From this aspect, the first case is a typical extension of the discussed two models case. Therefore, this case is also in the field of model detection.

Different from the first case, there is more than one nominal model in the second case. The alarms caused by the switchings among the nominal models should be separated from the alarms caused by the switching from nominal models to faulty models. In order to determine whether the system is in one of the m nominal models or the $N-m$ faulty models, all the m nominal models should be distinguished. It is the reason why the active fault diagnosis for multiple models is also called as model isolation.

Typically, if there are $N-1$ nominal models and one faulty model, it is necessary to distinguish all the models. Thus, the following part introduces a method for the most ordinary case: discriminate all the N models.

5.4.1. Active fault diagnosis with a bank of reference models

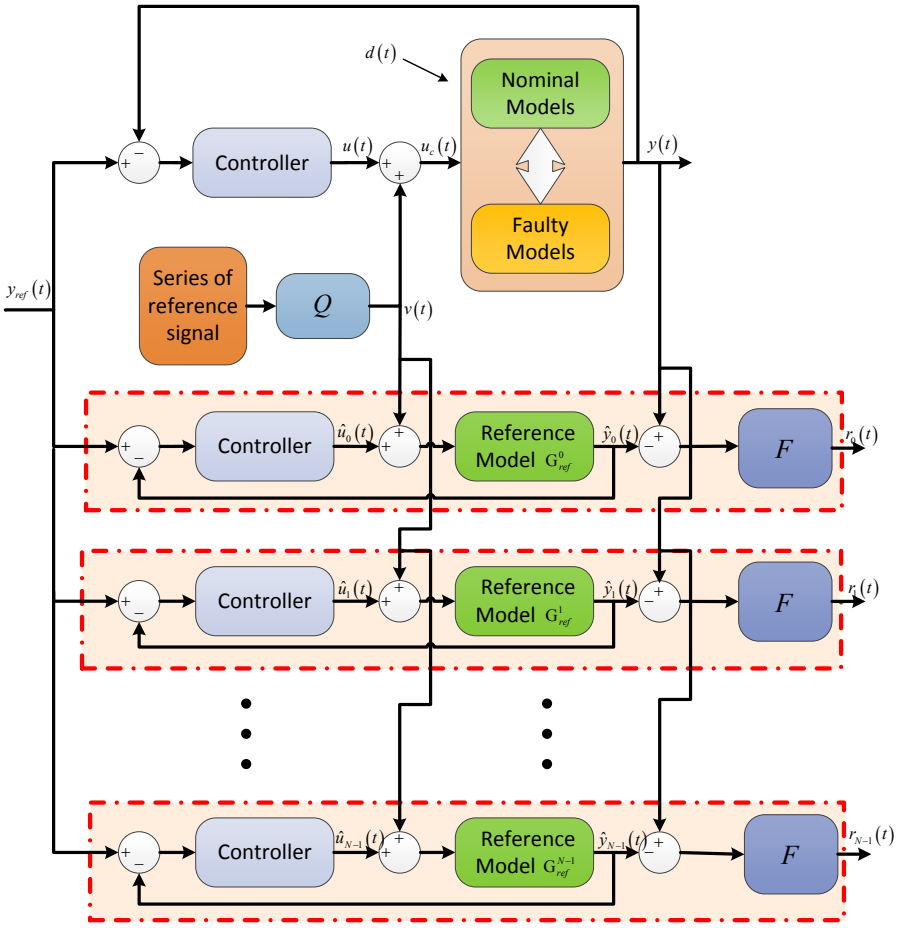

Figure 5.4.: Active fault diagnosis for multiple models with a bank of reference models G^j_{ref}

In the multiple models case, as shown in (5.4), there are i plant models (i:$0, \ldots, N-1$, $N > 2$) to discriminate, which are noted as P_i (i:$0, \ldots, N-1$). For the simplicity of expression, the multiple models case also illustrates the design with an SISO system. Note G_i as the transfer function from input u_c to output y when system is in P_i. In order to discriminate all the N models, a bank of reference models is used to generate a bank of residuals, as shown in Fig. 5.4. Similar to the two models case, all the plant models P_i without the part of disturbances are considered as the bank of reference models G^j_{ref} (j:$0, \ldots, N-1$):

$$G^j_{ref} : \begin{cases} \dot{\hat{x}}_j = A_j \hat{x}_j + B^j_u (\hat{u}_j + v) \\ \hat{y}_j = C_j \hat{x}_j + D^j_u (\hat{u}_j + v) \end{cases} \tag{5.45}$$

where the transfer function G^j_{ref} will be the same as the transfer function G_i if $i = j$.

In this framework, N residuals r_j (j:$0, \ldots, N-1$) are generated (assuming that the switching can occur among all the models):

- If $j = i$:

$$\begin{aligned} r_j &= \mathfrak{L}^{-1}\left(F\left(y - \hat{y}_j\right)\right) \\ &= \mathfrak{L}^{-1}\left(FG_{yd}^i d\right) + \Psi_i\left(x\left(0\right)\right) = r_i^i \end{aligned} \tag{5.46}$$

  where

$$G_{yd}^i = C_i\left(sI - A_i\right)^{-1} B_d^i + D_d^i \tag{5.47}$$

  and $\Psi_i\left(x\left(0\right)\right)$ is the natural response of G_i with the initial state $x\left(0\right)$ when the system switches to P_i.

- If $j \neq i$:

$$\begin{aligned} r_j &= \mathfrak{L}^{-1}\left(F\left(y - \hat{y}_j\right)\right) \\ &= \mathfrak{L}^{-1}\Bigg(F\Big(\frac{G_cG_i}{1+G_cG_i} - \frac{G_cG_{ref}^j}{1+G_cG_{ref}^j}\Big)y_{ref} + F\Big(\frac{G_i}{1+G_cG_i} - \frac{G_{ref}^j}{1+G_cG_{ref}^j}\Big)v \\ &\qquad + FG_{yd}^i d\Bigg) + \Psi_i\left(x\left(0\right)\right) \\ &= \mathfrak{L}^{-1}\left(F\Big(\frac{G_i}{1+G_cG_i} - \frac{G_{ref}^j}{1+G_cG_{ref}^j}\Big)\left(G_c y_{ref} + v\right) + FG_{yd}^i d\right) + \Psi_i\left(x\left(0\right)\right) \\ &= \mathfrak{L}^{-1}\left(F\Big(\frac{G_i}{1+G_cG_i} - \frac{G_{ref}^j}{1+G_cG_{ref}^j}\Big)v' + FG_{yd}^i d\right) + \Psi_i\left(x\left(0\right)\right) \\ &= r_j^i \end{aligned} \tag{5.48}$$

  where

$$G_i = C_i\left(sI - A_i\right)^{-1} B_u^i + D_u^i \tag{5.49}$$

  and $\Psi_i\left(x\left(0\right)\right)$ is the natural response of G_i with the initial state $x\left(0\right)$ when the system switches to P_i.

Similar to the analysis in the part of two models case, without any information of disturbances, the discrimination only with disturbances is unreliable. Reference input y_{ref} is generated with the aim of control, which cannot be used to discriminate the models in the active way. The transients of the part $F\left(\frac{G_cG_i}{1+G_cG_i} - \frac{G_cG_{ref}^j}{1+G_cG_{ref}^j}\right)y_{ref}$ provide a possibility to discriminate models; however, the discrimination is also unreliable. The natural response caused by switching also could be used to distinguish models. Nonetheless, this discrimination only could be used during the switching, and it depends on the states of system when switching arises.

A better alternative way to distinguish models under the effects of disturbances is by introducing an auxiliary signal. Then, the objective is changed to design auxiliary signal generator Q and post filter F. Since there is no information about the initial state $x\left(0\right)$ of switching, auxiliary signal is also designed for the steady state case.

As introduced in Remark 5.3, a design with v' in (5.48) is equivalent to the design of auxiliary signal v if the reference inputs are known. The part of the reference input y_{ref} in (5.48) is also omitted in the following part.

An important issue of this framework is to obtain the alarms from N residuals when the auxiliary signal is injected into the system, and make a decision that the system is in which model according to these alarms. As the discrimination for two models, the first step is to set thresholds J_{th}^i such that

$$\begin{cases} \|r\|_{\text{rms}} > J_{th}^i & \text{Alarm} \\ \|r\|_{\text{rms}} \leqslant J_{th}^i & \text{No Alarm} \end{cases} \tag{5.50}$$

Similarly, without the consideration of natural response caused by the switching, because the residual is only affected by the disturbances when $i = j$ (5.46), the thresholds could be selected as the estimation of maximum $\|r_i^i\|_{\text{rms}}$:

$$\begin{aligned} \left\|r_i^i\right\|_{\text{rms}} &= \left\|\mathfrak{L}^{-1}\left(FG_{yd}^i d\right)\right\|_{\text{rms}} \\ &\leqslant \left\|FG_{yd}^i\right\|_\infty \|d\|_{\text{rms}} \\ &\leqslant \left\|FG_{yd}^i\right\|_\infty \cdot \max\left(\|d\|_{\text{rms}}\right) \end{aligned} \tag{5.51}$$

As a consequence, thresholds J_{th}^i could be:

$$J_{th}^i = \left\|FG_{yd}^i\right\|_\infty \cdot \max\left(\|d\|_{\text{rms}}\right) \tag{5.52}$$

Different from two models discrimination case, an alarm in (5.50) for multiple models case does not mean a switching arises. Without auxiliary signal, there may be some alarms in some residual r_j caused by any factors of disturbances or reference input, even when there is no model switching. For example, the system is in the nominal model P_0. Without model switching and auxiliary signal, the residuals r_j $(j \neq 0)$ are:

$$r_j = \mathfrak{L}^{-1}\left(F\left(\frac{G_c G_0}{1 + G_c G_0} - \frac{G_c G_{ref}^j}{1 + G_c G_{ref}^j}\right) y_{ref} + FG_{yd}^i d\right)$$

and the corresponding $\|r_j\|_{\text{rms}}$ may be higher than the threshold in (5.52). Therefore, it is necessary to find a discrimination logic about the alarms in (5.50) for different residual r_j and different thresholds J_{th}^i $(i,\ j : 0, \ldots, N\text{-}1)$ to achieve a successful discrimination for the multiple models case. In the next section, a logic to make decision and the corresponding requirements of auxiliary signal design are introduced.

5.4.2. Logic to discriminate multiple models with unique F and unique Q

For simplicity of the expression, we consider $N = 3$ to illustrate the logic of multiple models discrimination with unique F and unique Q. As introduced in the previous part, $G_{ref}^0 = G_0$, $G_{ref}^1 = G_1$ and $G_{ref}^2 = G_2$. We note that (P_i, r_j) means the residual r_j when system is in model P_i (5.8).

As shown in (5.46), residual r_j is only affected by the disturbances if $i = j$ when the system is in steady state. With the definition of threshold J_{th}^i in (5.52) and the alarm logic in (5.50),

threshold J_{th}^i is the estimated upper bound of the effects of the disturbances on residual r_i^i:

$$\left\|r_i^i\right\|_{\text{rms}} = \left\|\mathfrak{L}^{-1}\left(F\left(y-\hat{y}\right)\right)\right\|_{\text{rms}} \leqslant J_{th}^i$$

Therefore, there will be no alarm with thresholds J_{th}^0, J_{th}^1 and J_{th}^2 for r_0, r_1 and r_2 respectively. Then, the following table could be obtained:

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.1.: Logic to discriminate

where $\times$ means that there is an alarm with the corresponding selected threshold J_{th}^i, red $\times$ implies the required alarm with threshold J_{th}^i caused by the auxiliary signal, and $\circ$ represents that there is no alarm with threshold J_{th}^i. The same definitions of the symbols $\times$ and $\circ$ are used in the following part.

There are some relationships between the selected thresholds J_{th}^0, J_{th}^1 and J_{th}^2. For convenience, we assume that $J_{th}^0 \geqslant J_{th}^1 \geqslant J_{th}^2$. If $J_{th}^0 \leqslant J_{th}^1$, a similar table like Tab. 5.1 can be obtained by exchanging the places of P_0 and P_1, r_0 and r_1, J_{th}^0 and J_{th}^1 simultaneously. Tab. 5.1 shows that there is no alarm for (P_1, r_1) with threshold J_{th}^1, then, we can obtain following equation with the residual in (5.46):

$$\left\|r_1^1\right\|_{\text{rms}} = \left\|F\left(y-\hat{y}\right)\right\|_{\text{rms}} \leqslant J_{th}^1 \leqslant J_{th}^0 \tag{5.53}$$

which means that there is also no alarm for (P_1, r_1) with threshold J_{th}^0. With the similar derivations for (P_2, r_2) with J_{th}^0 and J_{th}^1, Tab. 5.1 could be extended to:

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.2.: logic to discriminate

In order to guarantee a reliable model discrimination, a concept of Hamming distance is introduced into the design, which is widely used in the field of communications. In information theory, the Hamming distance means the minimum number substitutions required to change one string into the other string. Hamming distance in this chapter means the number of the different alarms with same residual r_i $(i : 0,\ 1,\ 2)$ and threshold J_{th}^j $(\ j : 0,\ 1,\ 2)$ between P_k and P_l $(k \neq l,\ k,\ l : 0,\ 1,\ 2)$. With the aim to discriminate models with enough Hamming distances $(\geqslant 2)$, the existed information in Tab. 5.2 should be fully utilized. For the model P_0 with residuals r_0, r_1 and r_2, because

$$\left\| r_0^0 \right\|_{\text{rms}} = \left\| \mathfrak{L}^{-1} \left(F \left(y - \hat{y} \right) \right) \right\|_{\text{rms}} \leqslant J_{th}^0 \tag{5.54}$$

and $J_{th}^0 \geqslant J_{th}^1 \geqslant J_{th}^2$, the alarms for (P_0, r_0) with thresholds J_{th}^1 and J_{th}^2 are not guaranteed. It only can be sure that there is no alarm for (P_0, r_0) with threshold J_{th}^0. In order to separate model P_0 from P_1 and P_2 with the auxiliary signal, this unique information should be utilized to design the auxiliary signal. Therefore, there should be an alarm for (P_1, r_0) and (P_2, r_0) with threshold J_{th}^0 if auxiliary signal is injected into the system.

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.3.: logic to discriminate

Due to the fact that there is an alarm for (P_1, r_0) with threshold J_{th}^0 and $J_{th}^0 \geqslant J_{th}^1 \geqslant J_{th}^2$, the

following relationship could be obtained

$$\left\|r_0^1\right\|_{\text{rms}} = \left\|\mathfrak{L}^{-1}\left(F\left(y-\hat{y}\right)\right)\right\|_{\text{rms}} > J_{th}^0 \geqslant J_{th}^1 \geqslant J_{th}^2 \tag{5.55}$$

which means that there will also be alarms for the case (P_1, r_0) with thresholds J_{th}^1 and J_{th}^2. Same conclusions could be gotten for (P_2, r_0) with thresholds J_{th}^1 and J_{th}^2.

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.4.: logic to discriminate

Since the designed auxiliary signal should guarantee that there is an alarm for (P_1, r_0) with threshold J_{th}^0, we have

$$\inf_d \left\|r_0^1\right\|_{\text{rms}} > J_{th}^0 \tag{5.56}$$

Note that the following equation could be obtained with the theory in the part of discrimination in the worst case (G_i represents the transfer function in (5.49)):

$$\begin{aligned} \inf_d \left\|r_0^1\right\|_{\text{rms}} &= \inf_d \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1} - \frac{G_{ref}^0}{1+G_cG_{ref}^0}\right)v + FG_{yd}^1 d\right)\right\|_{\text{rms}} \\ &\geqslant \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1} - \frac{G_{ref}^0}{1+G_cG_{ref}^0}\right)v\right)\right\|_{\text{rms}} - \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1 d\right)\right\|_{\text{rms}} \end{aligned} \tag{5.57}$$

The condition (5.56) could be met if

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1} - \frac{G_{ref}^0}{1+G_cG_{ref}^0}\right)v\right)\right\|_{\text{rms}} - \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1 d\right)\right\|_{\text{rms}} > J_{th}^0 \tag{5.58}$$

Therefore

$$\begin{aligned} \left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1} - \frac{G_{ref}^0}{1+G_cG_{ref}^0}\right)v\right)\right\|_{\text{rms}} &> \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^1 d\right)\right\|_{\text{rms}} + J_{th}^0 \\ &= J_{th}^1 + J_{th}^0 \\ &= J_{th}^1 + \sup_d \left\|\mathfrak{L}^{-1}\left(FG_{yd}^0 d\right)\right\|_{\text{rms}} \end{aligned} \tag{5.59}$$

Then (5.58) will be

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G^0_{ref}}{1+G_cG^0_{ref}}\right)v\right)\right\|_{\text{rms}}-\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}>J^1_{th} \tag{5.60}$$

Because of $G^0_{ref}=G_0$, $G^1_{ref}=G_1$ and non-negativity of the RMS norm in (3.22),

$$\begin{aligned}\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G^0_{ref}}{1+G_cG^0_{ref}}\right)v\right)\right\|_{\text{rms}}&=\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G^0_{ref}}{1+G_cG^0_{ref}}-\frac{G_1}{1+G_cG_1}\right)v\right)\right\|_{\text{rms}}\\&=\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_0}{1+G_cG_0}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v\right)\right\|_{\text{rms}}\end{aligned} \tag{5.61}$$

With the condition in (5.60), we have

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_0}{1+G_cG_0}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v\right)\right\|_{\text{rms}}-\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}>J^1_{th} \tag{5.62}$$

Because

$$\begin{aligned}\inf_d\left\|r^1_0\right\|_{\text{rms}}&=\inf_d\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_0}{1+G_cG_0}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v+FG^0_{yd}d\right)\right\|_{\text{rms}}\\&\geqslant\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_0}{1+G_cG_0}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v\right)\right\|_{\text{rms}}-\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}\end{aligned} \tag{5.63}$$

Then

$$\inf_d\left\|r^0_1\right\|_{\text{rms}}>J^1_{th} \tag{5.64}$$

which means that there is an alarm for (P_0,r_1) with threshold J^1_{th}. With the same rule, an alarm for (P_2,r_0) with threshold J^0_{th} implies an alarm for (P_0,r_2) with threshold J^2_{th}. Because there is an alarm for (P_0,r_1) with threshold J^1_{th} and $J^1_{th}\geqslant J^2_{th}$, there will be an alarm for (P_0,r_1) with threshold J^2_{th}. Then, we have Tab. 5.5:

Until now, from Tab. 5.5 with 3 residuals (r_0, r_1 and r_2), we can find that the Hamming distance to separate model P_0 from P_1 and P_2 is 2. In order to separate models P_1 from P_2 with 2 Hamming distances, there should be at least 2 alarms for (P_1,r_2) with thresholds J^0_{th}, J^1_{th} , J^2_{th} and (P_2,r_1) with thresholds J^0_{th} and J^1_{th}. Due to the fact that $J^0_{th}\geqslant J^1_{th}\geqslant J^2_{th}$, alarms for (P_1,r_2) with thresholds J^0_{th} and J^1_{th} are more conservative than with threshold J^2_{th}, and an alarm for (P_2,r_1) with threshold J^0_{th} is more conservative than with threshold J^1_{th}. It is interesting to find that an alarm for (P_1,r_2) with threshold J^2_{th} is equivalent to an alarm for (P_2,r_1) with threshold J^1_{th}. Therefore, with least conservativeness, the auxiliary signal design should guarantee that there will be an alarm for the case (P_2,r_1) with thresholds J^1_{th}, which also results in an alarm for the case (P_1,r_2) with threshold J^2_{th}. Consequently, we get the following table:

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.5.: logic to discriminate

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\begin{cases} \circ & J_{th}^0 \\ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \times & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$
r_1	$\begin{cases} & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ \times & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$
r_2	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} & J_{th}^0 \\ & J_{th}^1 \\ \times & J_{th}^2 \end{cases}$	$\begin{cases} \circ & J_{th}^0 \\ \circ & J_{th}^1 \\ \circ & J_{th}^2 \end{cases}$

Table 5.6.: logic to discriminate

After the unknown alarm cases are cleared from Tab. 5.6, the decision logic of multiple models discrimination is shown in Tab. 5.7.

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	$\circ \quad J_{th}^0$	$\times \quad J_{th}^0$	$\times \quad J_{th}^0$
r_1	$\times \quad J_{th}^1$	$\circ \quad J_{th}^1$	$\times \quad J_{th}^1$
r_2	$\times \quad J_{th}^2$	$\times \quad J_{th}^2$	$\circ \quad J_{th}^2$

Table 5.7.: Decision logic for 3 models discrimination

Remark 5.5. Without auxiliary signal, models can be discriminated only if alarms caused by any factors of disturbances, reference inputs and natural response caused by switching meet the decision logic in Tab. 5.7 for multiple models case. Comparing with the models discrimination condition for two models case, the discriminination conditon for multiple models case without auxiliary signal is much stricter. In other words, comparing with the two models case, the auxiliary signal is more important for the models discrimination in the multiple models case.

Remark 5.6. The decision logic in Tab. 5.7 is also appropriate for the case when it is not necessary to discriminate all the models. For example, it might be desirable to separate model

0 from models 1 and 2, but it does not need to distinguish models 1 and 2. Then, the auxiliary signal does not need to achieve all the alarms in Tab. 5.7, but just guarantees the alarms for (P_1, r_0) and (P_2, r_0) with threshold J^0_{th}. The corresponding designed auxiliary signal will not be as conservative as the design with requirements of all the alarms in Tab. 5.7.

Remark 5.7. Since the natural response caused by switching affects the residual r_j in (5.46) and (5.48), the switching will disturb the decision logic in Tab. 5.7 if the auxiliary signal is injected into the system when switching occurs. In this case, a series of auxiliary signals with a fixed period could be injected into the system, then models can be discriminated with one of the auxiliary signals when the natural response caused by switching converges to steady state.

From Tab. 5.7 with 3 residuals (r_0, r_1 and r_2) and three thresholds, the Hamming distance for either pair among P_0, P_1 and P_2 is 2 . Then, the key of model discrimination with 2 Hamming distance is to guarantee that there will be alarms for (P_1, r_0) with threshold J^0_{th}, (P_2, r_0) with threshold J^0_{th} and (P_2, r_1) with threshold J^1_{th} by designing the auxiliary signal generator Q and the post filter F. Then, the following inequalities should be satisfied for the three models discrimination:

$$\inf_d \left\| r_0^1 \right\|_{\text{rms}} > J^0_{th} \tag{5.65}$$

$$\inf_d \left\| r_0^2 \right\|_{\text{rms}} > J^0_{th} \tag{5.66}$$

$$\inf_d \left\| r_1^2 \right\|_{\text{rms}} > J^1_{th} \tag{5.67}$$

Note that

$$\begin{aligned} \inf_d \left\| r_0^1 \right\|_{\text{rms}} &= \inf_d \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_1}{1+G_cG_1} - \frac{G^0_{ref}}{1+G_cG^0_{ref}} \right) v + FG^1_{yd} d \right) \right\|_{\text{rms}} \\ &\geqslant \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_1}{1+G_cG_1} - \frac{G^0_{ref}}{1+G_cG^0_{ref}} \right) v \right) \right\|_{\text{rms}} - \sup_d \left\| \mathfrak{L}^{-1} \left(FG^1_{yd} d \right) \right\|_{\text{rms}} \end{aligned} \tag{5.68}$$

$$\begin{aligned} \inf_d \left\| r_0^2 \right\|_{\text{rms}} &= \inf_d \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_2}{1+G_cG_2} - \frac{G^0_{ref}}{1+G_cG^0_{ref}} \right) v + FG^2_{yd} d \right) \right\|_{\text{rms}} \\ &\geqslant \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_2}{1+G_cG_2} - \frac{G^0_{ref}}{1+G_cG^0_{ref}} \right) v \right) \right\|_{\text{rms}} - \sup_d \left\| \mathfrak{L}^{-1} \left(FG^2_{yd} d \right) \right\|_{\text{rms}} \end{aligned} \tag{5.69}$$

$$\begin{aligned} \inf_d \left\| r_1^2 \right\|_{\text{rms}} &= \inf_d \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_2}{1+G_cG_2} - \frac{G^1_{ref}}{1+G_cG^1_{ref}} \right) v + FG^2_{yd} d \right) \right\|_{\text{rms}} \\ &\geqslant \left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_2}{1+G_cG_2} - \frac{G^1_{ref}}{1+G_cG^1_{ref}} \right) v \right) \right\|_{\text{rms}} - \sup_d \left\| \mathfrak{L}^{-1} \left(FG^2_{yd} d \right) \right\|_{\text{rms}} \end{aligned} \tag{5.70}$$

The conditions in (5.65), (5.66) and (5.67) are satisfied when

$$\left\| \mathfrak{L}^{-1} \left(F \left(\frac{G_1}{1+G_cG_1} - \frac{G^0_{ref}}{1+G_cG^0_{ref}} \right) v \right) \right\|_{\text{rms}} - \sup_d \left\| \mathfrak{L}^{-1} \left(FG^1_{yd} d \right) \right\|_{\text{rms}} > J^0_{th} \tag{5.71}$$

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G^0_{ref}}{1+G_cG^0_{ref}}\right)v\right)\right\|_{\text{rms}}-\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}>J^0_{th} \tag{5.72}$$

$$\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v\right)\right\|_{\text{rms}}-\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}>J^1_{th} \tag{5.73}$$

Due to $G_i=G^i_{ref}$ where $i=0,\ 1,\ 2$, then we can simplify above three conditions:

$$\begin{aligned}\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G^0_{ref}}{1+G_cG^0_{ref}}\right)v\right)\right\|_{\text{rms}}&=\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}\\&\geqslant\sup_d\left\|\mathfrak{L}^{-1}\left(FG^1_{yd}d\right)\right\|_{\text{rms}}+J^0_{th}\\&=\sup_d\left\|\mathfrak{L}^{-1}\left(FG^1_{yd}d\right)\right\|_{\text{rms}}+\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}\end{aligned} \tag{5.74}$$

$$\begin{aligned}\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G^0_{ref}}{1+G_cG^0_{ref}}\right)v\right)\right\|_{\text{rms}}&=\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}\\&\geqslant\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}+J^0_{th}\\&=\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}+\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}\end{aligned} \tag{5.75}$$

$$\begin{aligned}\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G^1_{ref}}{1+G_cG^1_{ref}}\right)v\right)\right\|_{\text{rms}}&=\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_1}{1+G_cG_1}\right)v\right)\right\|_{\text{rms}}\\&\geqslant\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}+J^1_{th}\\&=\sup_d\left\|\mathfrak{L}^{-1}\left(FG^2_{yd}d\right)\right\|_{\text{rms}}+\sup_d\left\|\mathfrak{L}^{-1}\left(FG^1_{yd}d\right)\right\|_{\text{rms}}\end{aligned} \tag{5.76}$$

Then, the following model discrimination conditions should be met to design the auxiliary signal:

$$\begin{aligned}\varphi^{10}_{\text{detec}}&=\frac{\sup_d\left\|\mathfrak{L}^{-1}\left(FG^1_{yd}d\right)\right\|_{\text{rms}}+\sup_d\left\|\mathfrak{L}^{-1}\left(FG^0_{yd}d\right)\right\|_{\text{rms}}}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}}\\&\leqslant\frac{\left(\left\|FG^1_{yd}\right\|_\infty+\left\|FG^0_{yd}\right\|_\infty\right)\max\left(\|d\|_{\text{rms}}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}}\\&<1\end{aligned} \tag{5.77}$$

$$\begin{aligned}\varphi_{\text{detec}}^{20} &= \frac{\sup\limits_{d}\left\|\mathfrak{L}^{-1}\left(FG_{yd}^{2}d\right)\right\|_{\text{rms}}+\sup\limits_{d}\left\|\mathfrak{L}^{-1}\left(FG_{yd}^{0}d\right)\right\|_{\text{rms}}}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}} \\ &\leqslant \frac{\left(\left\|FG_{yd}^{2}\right\|_{\infty}+\left\|FG_{yd}^{0}\right\|_{\infty}\right)\max\left(\|d\|_{\text{rms}}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_0}{1+G_cG_0}\right)v\right)\right\|_{\text{rms}}} \\ &< 1\end{aligned} \tag{5.78}$$

$$\begin{aligned}\varphi_{\text{detec}}^{21} &= \frac{\sup\limits_{d}\left\|\mathfrak{L}^{-1}\left(FG_{yd}^{2}d\right)\right\|_{\text{rms}}+\sup\limits_{d}\left\|\mathfrak{L}^{-1}\left(FG_{yd}^{1}d\right)\right\|_{\text{rms}}}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_1}{1+G_cG_1}\right)v\right)\right\|_{\text{rms}}} \\ &\leqslant \frac{\left(\left\|FG_{yd}^{2}\right\|_{\infty}+\left\|FG_{yd}^{1}\right\|_{\infty}\right)\max\left(\|d\|_{\text{rms}}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_2}{1+G_cG_2}-\frac{G_1}{1+G_cG_1}\right)v\right)\right\|_{\text{rms}}} \\ &< 1\end{aligned} \tag{5.79}$$

where the quantities of $\varphi_{\text{detec}}^{10}$, $\varphi_{\text{detec}}^{20}$ and $\varphi_{\text{detec}}^{21}$ show the ability of the auxiliary signal to discriminate the models P_0, P_1 and P_2. In the case of 3 models, there are 3 model discrimination conditions to meet.

When there are N models to distinguish, with the Remark 5.4, we have

$$\varphi_{\text{detec}}^{ij} = \frac{\left(\left\|FG_{yd}^{i}\right\|_{\infty}+\left\|FG_{yd}^{j}\right\|_{\infty}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_j}{1+G_cG_j}-\frac{G_i}{1+G_cG_i}\right)v\right)\right\|_{\text{rms}}} < 1 \tag{5.80}$$

where $j \in \{0,1,\ldots,N-1\}$ and $i : \{j+1,\ldots,j+N-1\}$, $N \geqslant 2$. Thus, $\left(N^2-N\right)/2$ model discrimination conditions should be met to design the auxiliary signal for N multiple models case.

In the optimization design for the multiple models case with unique F and unique Q, the model discrimination condition (5.44) in Section. 5.3.6 should be replaced by

$$\varphi_{\text{detec}}^{ij} = \frac{\left(\left\|FG_{yd}^{i}\right\|_{\infty}+\left\|FG_{yd}^{j}\right\|_{\infty}\right)}{\max\limits_{t}\left(\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_j}{1+G_cG_j}-\frac{G_i}{1+G_cG_i}\right)QH_{\Delta}\right)\right\|_{\text{rms}}\right)} < 1 \tag{5.81}$$

where $j \in \{0,1,\ldots,N-1\}$ and $i : \{j+1,\ldots,j+N-1\}$, $N \geqslant 2$.

In the two models case, (5.38) is added into the optimization to restrict the convergence rate of output responses caused by auxiliary signals. By contrast, the constraint (5.38) is not enough for the multiple models case anymore. The constraint (5.38) only considers the effects of the transients of output responses caused by auxiliary signals, but does not deal with the effects of the transients of residual responses from auxiliary signals to residual r. If the period T_1 of the periodic impulse signals $\sigma(t)$ is too small, the residual responses in residual r_j^i $(i \neq j)$ caused by auxiliary signals will accumulate. The evaluated residual $\left\|r_j^i\right\|_{\text{rms}}$ will increase when more auxiliary signals are applied, and an alarm will be caused in residual r_j^i when the accumulated evaluated residual exceeds the threshold even there is no injection of auxiliary signal during

the period. This phenomenon does not cause any problem for the two models case, because accumulated residual only appears in residual r^1 when system is in faulty model. However, this kind of alarm will disturb the decision logic in Tab. 5.7. For example, assume that there are three models P_0 , P_1 and P_3. The system stays at model P_0 for some time, then switches to P_1. The periodic auxiliary signals (5.35) are injected into this system when the system is in P_0. According to the definition of residuals in (5.46) and (5.48) and the effects of designed auxiliary signal, there will be alarms in r_1 and r_2 but not in r_0. Since the period T_1 is too small, the residual responses in r_1 and r_2 caused by the periodic auxiliary signals accumulate to be very big. When the system switches to P_1, there will be alarms in r_0 and r_2 when the new auxiliary signal attacks the system. However, the residual responses in r_1 are still big because of the accumulation with the periodic auxiliary signals when the system is in p_0, which may cause an alarm in r_1. As a consequence, there will be three alarms in the three residuals r_0, r_1 and r_2 respectively, which is not any case in Tab. 5.7. The correct result appears when the residual r_1 decays until the alarm in r_1 disappears. This kind of phenomenon will delay the model discrimination, and may cause wrong conclusion in some cases. Therefore, a constraint of convergence rate on the residual responses caused by auxiliary signal should be added into the design:

$$\max(\text{real}(\text{eig}(F\frac{G_i}{1+G_cG_i}Q))) \leqslant \varsigma \tag{5.82}$$

where ς is a negative value, which is selected to meet the requirements of the rapidity of auxiliary signal. The period T_1 in (5.35) should be bigger than the settling time of residual responses caused by the designed auxiliary signal with (5.82). In the optimization for the multiple models case, the constraints (5.38) should be replaced by (5.82).

5.5. Results

This section applies the proposed active fault diagnosis methods on a DC motor control system DR 300. Both two models case and multiple models case are given to illustrate the effectiveness of the design.

5.5.1. Speed control of a DC motor

DC (Direct Current) motor converts electrical energy into mechanical energy. The model of DC motor control system DR300 from [35] is considered for the application.

5.5.1.1. Model of DC motor

As shown in Fig. 5.5, the DC motor contains a mechanical part and an electrical part. Define that the loop current I_A and the armature frequency (rotation frequency) Ω as state variables, the terminal voltage U_A as input and the (unknown) load M_L as disturbances. The state space description of model could be

$$\begin{bmatrix} \dot{I}_A \\ \dot{\Omega} \end{bmatrix} = \begin{bmatrix} -\frac{R_A}{L_A} & -\frac{C\Phi}{L_A} \\ \frac{K_M}{J} & 0 \end{bmatrix} \begin{bmatrix} I_A \\ \Omega \end{bmatrix} + \begin{bmatrix} \frac{1}{L_A} \\ 0 \end{bmatrix} U_A + \begin{bmatrix} 0 \\ -\frac{1}{J} \end{bmatrix} M_L \tag{5.83}$$

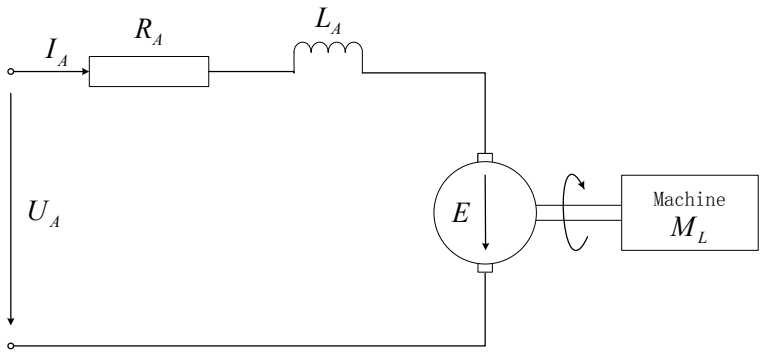

Figure 5.5.: Diagram of a DC motor

where the parameters are given in Tab. 5.8.

Parameters	Symbol	Value	Unit
Total Inertia	J	$80.45 \cdot 10^{-6}$	$kg \cdot m^2$
Voltage constant	$C\Phi$	$6.27 \cdot 10^{-3}$	V/Rpm
Motor constant	K_M	0.06	Nm/A
Armature Inductance	L_A	0.003	H
Resistance	R_A	3.13	Ohm
Tacho output voltage	K_T	$5 \cdot 10^{-3}$	V/Rpm
Tacho filter time constant	T_T	5	ms

Table 5.8.: Parameters of DC motor DR300

With the equation in (5.83), the structure diagram of motor with load is given in Fig. 5.6.

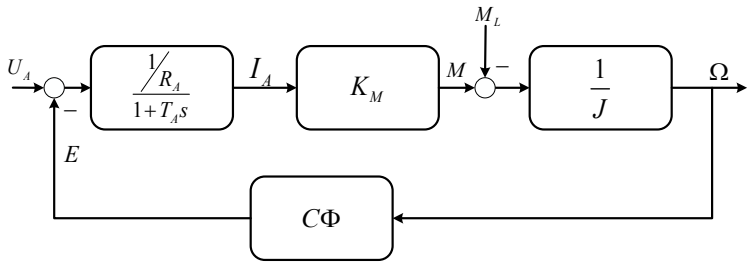

Figure 5.6.: Block diagram of the DC motor with load

5.5.1.2. Models of DC motor cascade control system

For the purpose of speed control, a speed control loop and a current control loop are contained in the cascade control scheme. In the framework of cascade control, the DC motor together with the current control loop (I_A) is considered as the plant, which will be controlled by a PI speed controller (C_Ω controller). The framework of cascade current control is shown in Fig. 5.7.

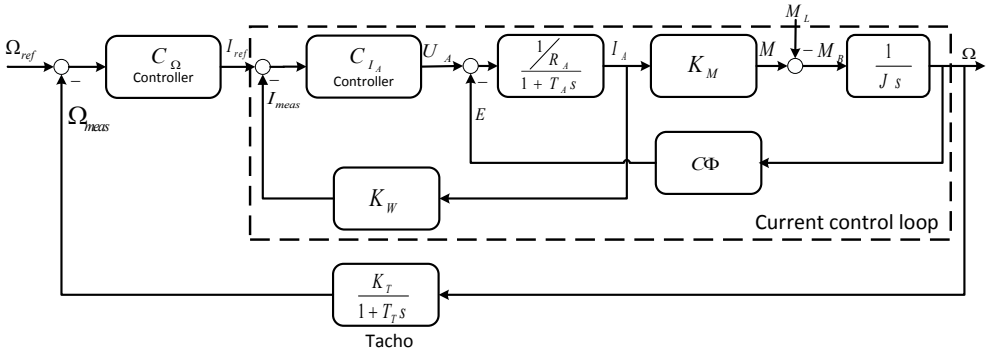

Figure 5.7.: Cascade current control

The transfer functions of the plant (DC motor with the current control loop) are:

$$\frac{\Omega}{I_{ref}} = \frac{K_M C_{I_A}}{JL_A s^2 + (C_{I_A} K_W J + JR_A)\, s + K_M C\phi}$$
$$\frac{\Omega}{M_L} = \frac{R_A + R_A T_A s + C_{I_A} K_W}{JR_A T_A s^2 + (R_A J + C_{I_A} K_W J)\, s + C\Phi K_M}$$

Then, closed loop model can be achieved by

$$y = G_{yu} u + G_{yd} d$$
$$G_{yu} = \frac{\Omega_{meas}}{\Omega_{ref}} = \frac{K_M C_{I_A} K_T C_\Omega}{K_M C_{I_A} K_T C_\Omega + (1 + T_T s)\,(JL_A s^2 + (C_{I_A} K_W J + JR_A)\, s + K_M C\Phi)}$$
$$G_{yd} = \frac{\Omega_{meas}}{M_L} = \frac{1}{K_M \frac{-C_{I_A} C_\Omega - C\Phi \frac{1+K_T s}{K_T}}{R_A + R_A T_A s + C_{I_A} K_W} - \frac{1+K_T s}{K_T} Js}$$

with $y = \Omega_{meas}$ as output, the control input singal $u = \Omega_{ref}$ and $d = M_L$ as disturbance.

The designed auxiliary signal could be injected into the system in the form of extra inputs Ω_{ref}, which can also be injected into the system at the output of C_Ω controller I_{ref}. Both of the designs are equivalent. In order to match the design framework in Fig. 5.2, Ω_{ref} is considered as the reference input y_{ref}, and auxiliary signal is injected into the system at I_{ref}. Then, the closed loop system could be modified as

$$y = G_{yu} u + G_{yd} d$$
$$G_{yu} = \frac{\Omega_{meas}}{\Omega_{ref}} = \frac{K_M C_{I_A} K_T}{K_M C_{I_A} K_T C_\Omega + (1 + T_T s)\,(JL_A s^2 + (C_{I_A} K_w J + JR_A)\, s + K_M C\Phi)}$$
$$G_{yd} = \frac{\Omega_{meas}}{M_L} = \frac{1}{K_M \frac{-C_{I_A} C_\Omega - C\Phi \frac{1+K_T s}{K_T}}{R_A + R_A T_A s + C_{I_A} K_W} - \frac{1+K_T s}{K_T} Js}$$

with $y = \Omega_{meas}$ as output, the input $u = I_{ref}$ (the output of the speed controller C_Ω) and $d = M_L$ as disturbance.

According to the transfer functions given in [35], the approximated transfer functions are given

(I_A controller is a P controller) with

$$K_W = 0.2611,\ C_{I_A} = 214.7$$

The PI speed controller (Ω controller, signed as C_Ω) is set to be

$$C_\Omega = 1.6\frac{1 + 1.225s}{s}$$

Three different models are obtained by changing the value of total inertia J:

- Model 0: $J_0 = 80.45 \cdot 10^{-6} kg \cdot m^2$.
- Model 1: $J_1 = 10^{-5} kg \cdot m^2$.
- Model 2: $J_2 = 10^{-6} kg \cdot m^2$.

5.5.2. Two models case (Model detection)

In two models case, Model 0 with J_0 is considered the nominal model, while Model 2 with J_2 is considered the faulty model.

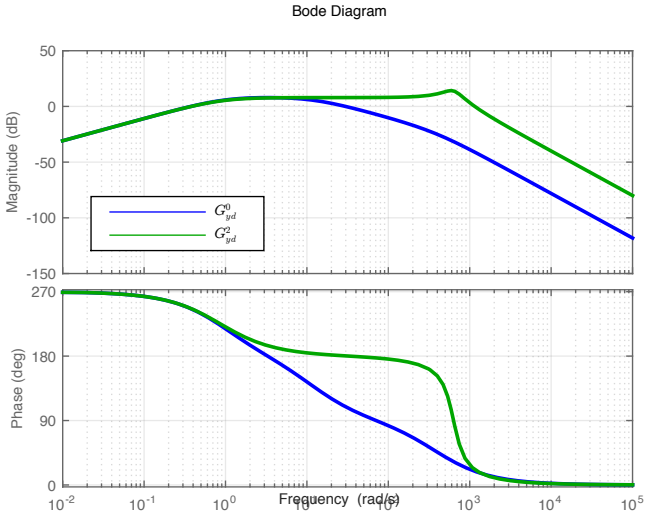

Figure 5.8.: Bode diagram of transfer function G^0_{yd} and G^2_{yd}.

Fig. 5.8 shows that, if the disturbances contain some high frequency parts ($\omega \geqslant 10\, rad/s$), the effects of disturbances on the output y for nominal model and faulty model are different. Then, it is possible to discriminate these two models using the high frequency part of the disturbances. However, Fig. 5.8 also shows that for the low frequency disturbances ($\omega < 10 rad/s$), the effects of these disturbances on the output y for the nominal case and the faulty case are the same. In other words, the low frequency disturbances cannot be used to discriminate the two models.

In order to validate the analysis, a kind of normalized low frequency disturbances ($\omega < 10 rad/s$) is generated in Fig. 5.9.

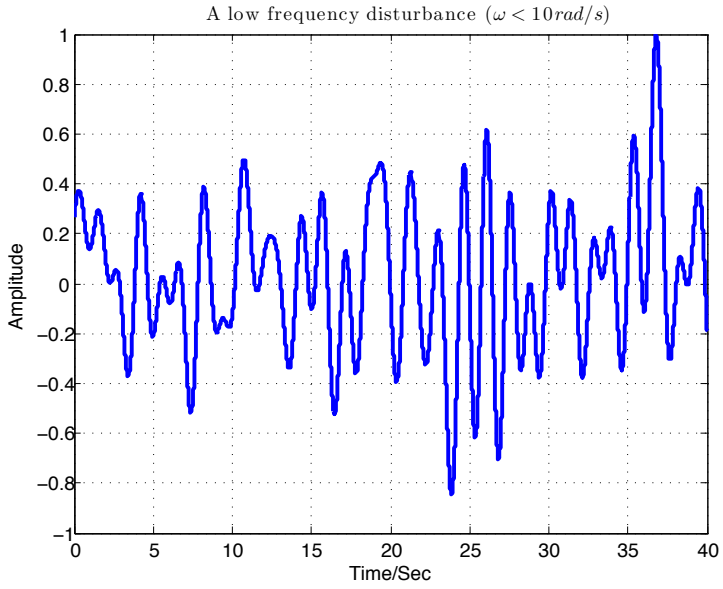

Figure 5.9.: Normalized low frequency disturbances ($\omega < 10\, rad/s$)

Without the auxiliary signal, Fig. 5.10 shows the effects of low frequency disturbances ($\omega < 10\, rad/s$) with $y_{ref} = 1$. The system will switch from the nominal model to the faulty model at 20s.

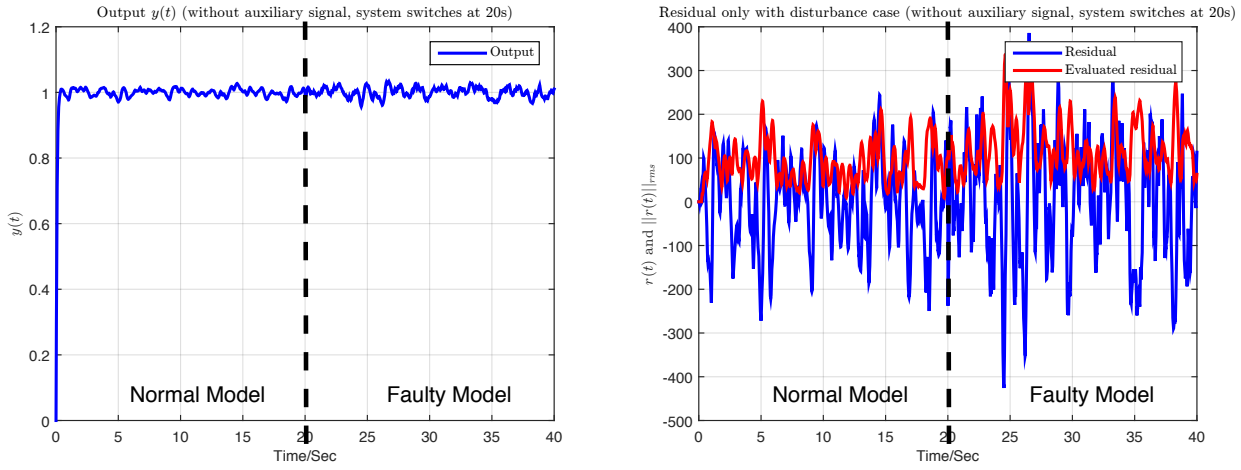

Figure 5.10.: Fail to detect the faulty model only using disturbances (system switches at 20s)

In this case, as shown in Fig. 5.10 (right), only with low frequency disturbances, it is impossible to discriminate these two models. Therefore, it is reasonable to design an auxiliary signal to distinguish models:

$$\min_{Q} \left(\max_{i=0,1} \left(\text{Peak} \left(\left\| \mathfrak{L}^{-1} \left(\frac{1}{1+G_c G_i} Q H_\Delta \right) \right\|_{\text{peak}} \right) \right) + \max_{i=0,1} \left(\text{Peak} \left(\left\| \mathfrak{L}^{-1} \left(\frac{G_i}{1+G_c G_i} Q \right) \right\|_{\text{peak}} \right) \right) \right)$$

under the constraints of

$$\frac{\left(\left\|FG_{yd}^{0}\right\|_{\infty}+\left\|FG_{yd}^{1}\right\|_{\infty}\right)}{\left\|\mathfrak{L}^{-1}\left(F\left(\frac{G_1}{1+G_cG_1}-\frac{G_0}{1+G_cG_0}\right)QH_{\Delta}\right)\right\|_{\text{rms}}}<1$$

$$\max(\text{real}(\text{eig}(\frac{G_i}{1+G_cG_i}Q)))\leqslant\varsigma=-0.7,\ \|F\|_{\infty}\leqslant 1$$

$$\left\|\mathcal{L}^{-1}\left(\frac{1}{s}\frac{G_0}{1+G_cG_0}Q\right)\right\|_{\text{peak}}\leqslant\xi_0=3$$

$$\left\|\mathcal{L}^{-1}\left(\frac{1}{s}\frac{G_1}{1+G_cG_1}Q\right)\right\|_{\text{peak}}\leqslant\xi_1=3$$

$$\max(\text{real}(\text{eig}(F)))<0$$

where the constraint $\|F\|_{\infty}\leqslant 1$ is used to normalize the parameter in F.

For this Multiple Input Single Output (MISO) example, the state space settings for Q in (5.5) and F in (5.6) could be transformed to be the transfer function expressions. In order to decrease the abrupt changes when the proposed auxiliary signal attacks the system, the filters Q and F are set as:

$$Q(s)=\frac{a_1s+a_2}{s^2+a_3s+a_4}$$
$$F(s)=\frac{b_1s^2+b_2s+b_3}{s^2+b_4s+b_5}$$

where a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 and b_5 are the parameters to design.

The initial points are $Q(s)=\frac{s+1}{s^2+s+1}$ and $F(s)=\frac{s^2+s+1}{s^2+s+1}$. Through the optimization with genetic algorithm, we get

$$Q^*(s)=\frac{-174.2s-201.7}{s^2+1.4s+410.3}$$
$$F^*(s)=\frac{0.4885s^2+17.5531s+236.7766}{s^2+17.56s+458}$$

The characteristics of the designed filers Q^* and F^* are shown in Fig. 5.11 (left). Since $\max\left(\|d\|_{\text{rms}}\right)=0.0253$ in the simulation, the auxiliary signal is generated with amplitue of reference signal be 0.0253. Then, the reference signal and the auxiliary signal (the signal generated by the auxiliary signal generator Q^*) used in the simulation are shown in Fig. 5.11 (right) with the parameter $\triangle=0.2s$ for the reference signal.

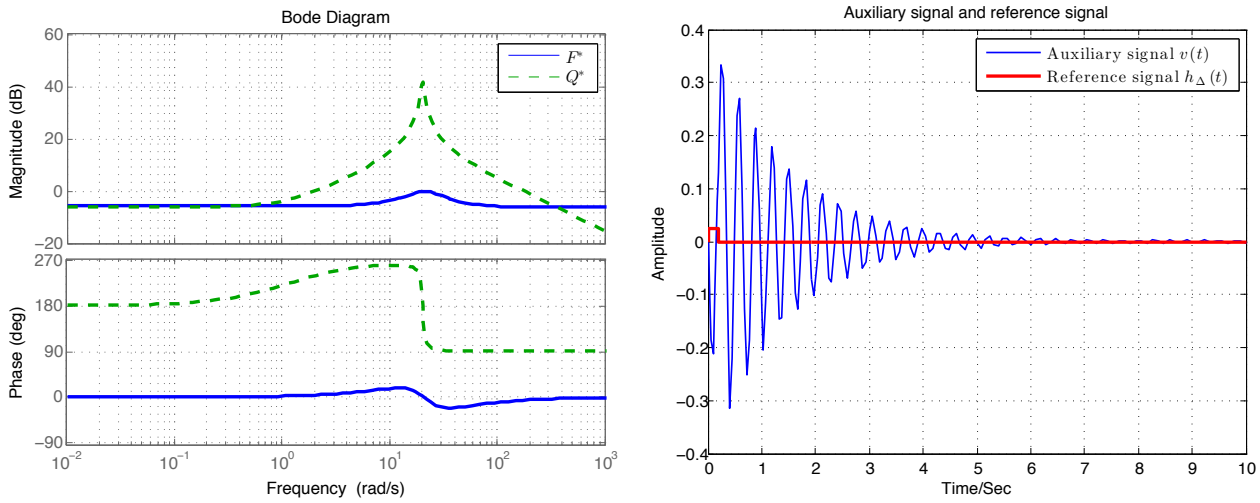

Figure 5.11.: Bode diagram of filters F^* and Q^*, and the designed auxiliary signal

In the setting of simulation, the system will switch from the nominal model to the faulty model at 50s, and the designed auxiliary signal will attack the system at 3s, 13s, 23s, 33s, 43s, 53s, 63s, 73s, 83s, 93s respectively. The disturbances in the simulation are kinds of low frequency noises ($\omega < 10\, rad/s$), which have the similar effects on the nominal and faulty models, as shown in Fig. 5.10. Fig. 5.12 (right) shows that comparing with the effects of disturbances on the control signal u_c, the effects of auxiliary signal on the control signal u_c are acceptable.

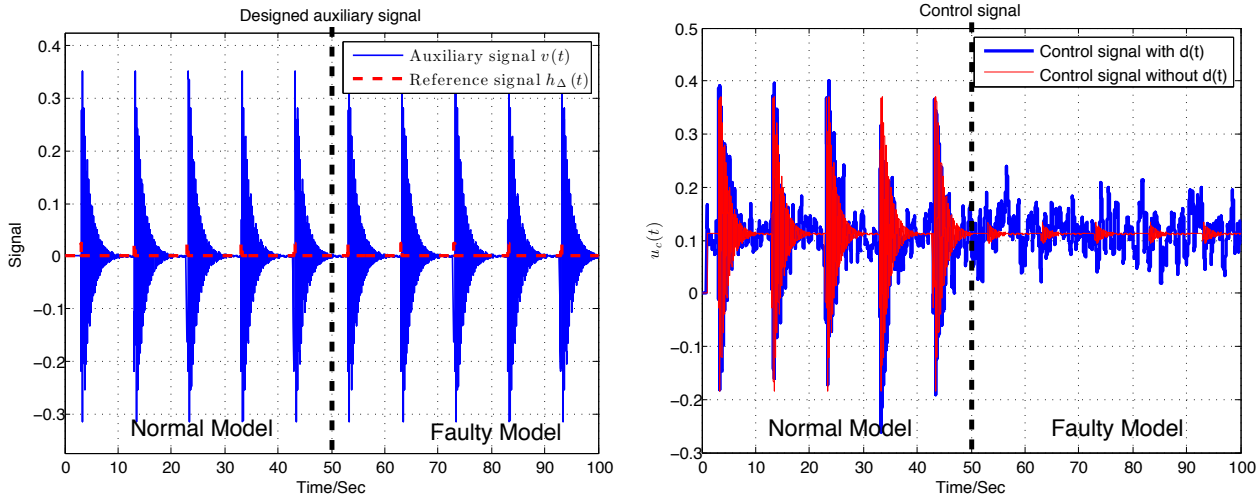

Figure 5.12.: Auxiliary signal, and the control signal u_c

As discussed in the previous part for the effects of the auxiliary signal on the outputs, the peak value of the responses from the auxiliary signal is used to evaluate the worst effects of the auxiliary signal on the outputs y. Fig. 5.13 shows that comparing with the effects of disturbances on the system outputs y, the effects of designed auxiliary signal on the system outputs y are not obvious.

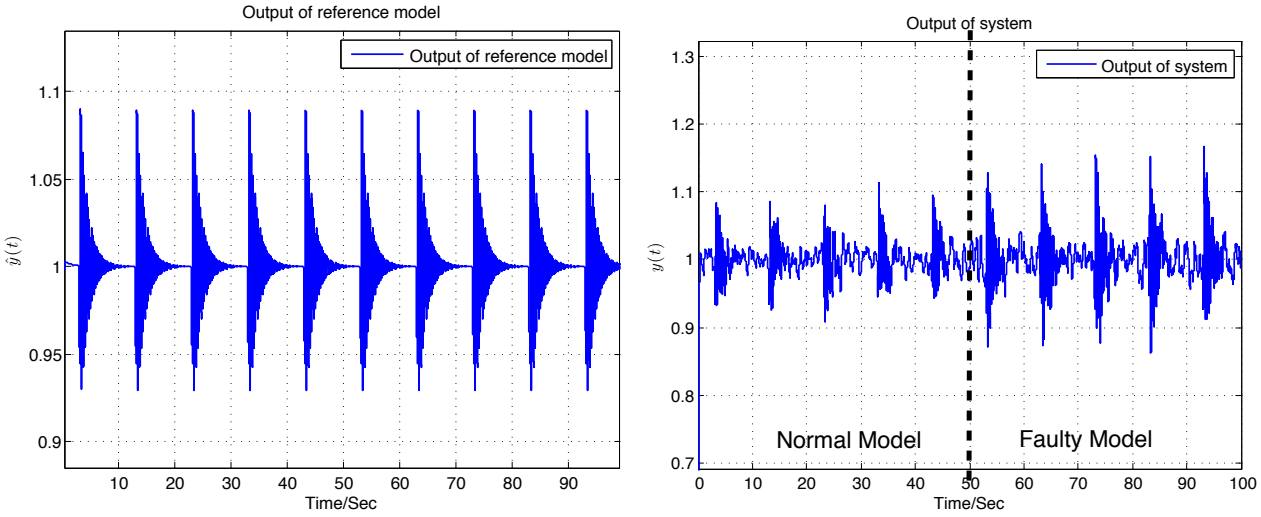

Figure 5.13.: The effects of auxiliary signal on the reference output $\hat{y}$ and system output y

With finite time window $T = 0.2s$, the evaluated residual $\|r\|_{\text{rms}}$ is shown in Fig. 5.14. When the system is in nominal model, there is no alarm. By contrast, alarms are geterated at 53s, 63s, 73s, 83s and 93s when auxilairy signals are injected into the system, which means that the faulty model is detected when the auxiliary signal attacks the system.

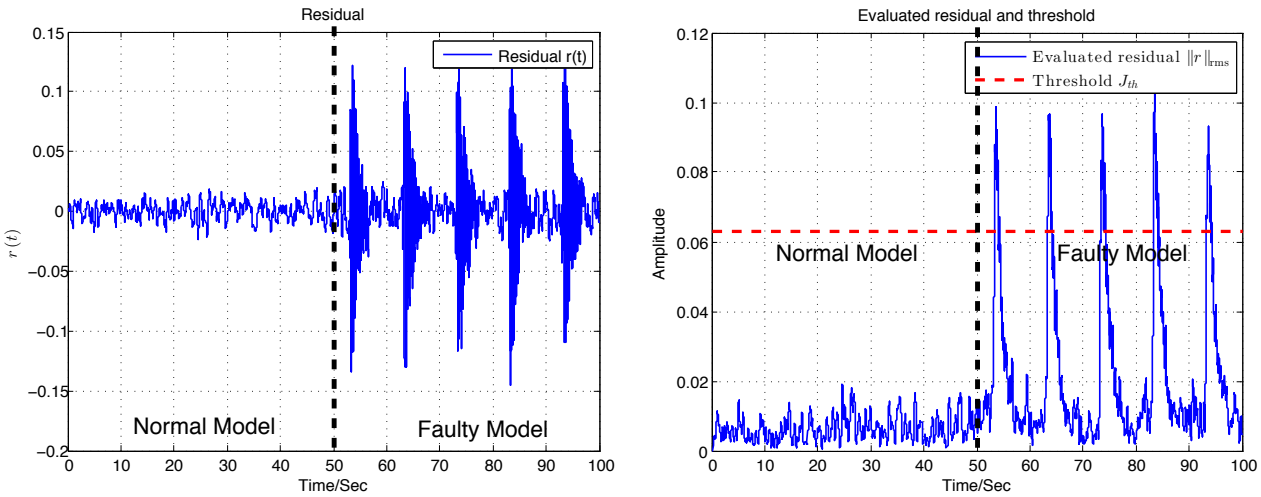

Figure 5.14.: Residual and threshold

5.5.3. Multiple models case (Model isolation)

For the multiple models case, all of the three models (Model 0, Model 1 and Model 2) are considered in the following part.

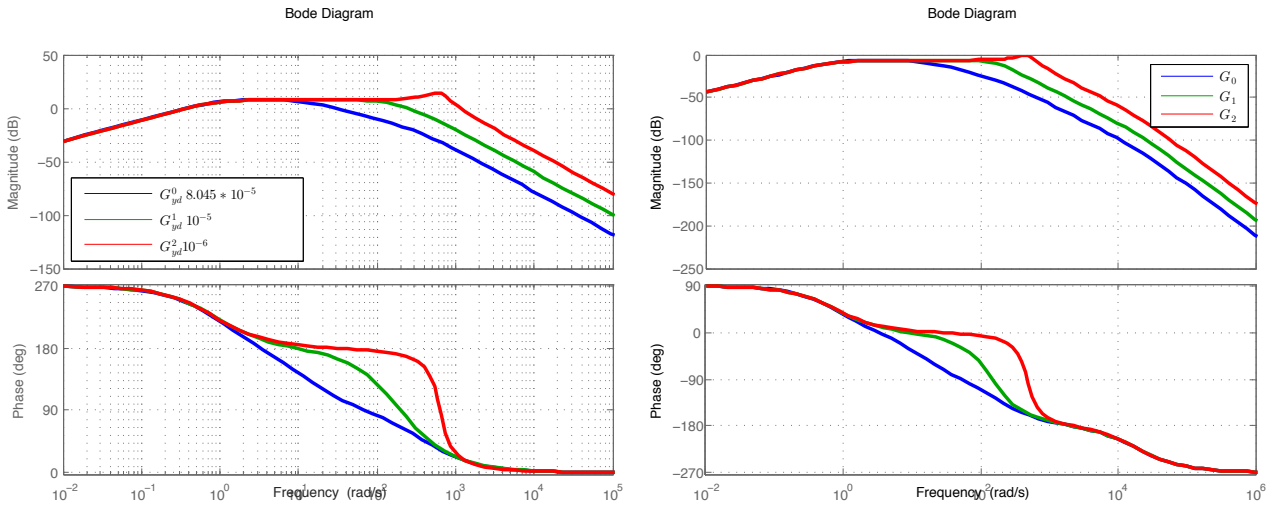

Figure 5.15.: Diagrams of different G^i_{yd} (5.47) and G_i (5.49) with different J_i $(i : 1,\ 2,\ 3)$

Fig. 5.15 (left) shows the bode diagrams of different G^i_{yd} with different J_i. As the similar analysis in the example of two models case, these three models show the same dynamics with low frequency disturbances ($\omega < 5rad/s$), while the effects of disturbances ($5 < \omega < 100rad/s$) on the models P_1 and P_2 are the same. If the disturbances work on the low frequency ($\omega < 5rad/s$), it is impossible to discriminate these three models by the disturbances. The disturbances ($5 < \omega < 100rad/s$) could be used to separate the model P_0 from the models P_1 and P_2, but could not be used to discriminate the models P_1 and P_2. By contrast, the active fault diagnosis with some suitable auxiliary signals could distinguish all three models directly.

The prior information of disturbances is not used for the design in two models case. If some information of disturbances is known, they should be included into the model, as introduced in Remark 5.2. In this example, it is difficult to discriminate models only by the low frequency disturbances ($\omega < 5rad/s$). In the design, the frequency range of the disturbances will affect the settings of the threshold (5.16) and the model discrimination condition (5.81). Therefore, the prior information of the disturbances should also be included into the optimization for the model discrimination condition in (5.81).

Therefore, the objective function of disturbances rejection

$$\min_{F} \left\| FG^i_{yd} \right\|_\infty + \left\| FG^j_{yd} \right\|_\infty \tag{5.84}$$

could be transformed to

$$\min_{F,\omega\in[\omega_1,\omega_2]} \left\| FG^i_{yd} \right\|_\infty + \left\| FG^j_{yd} \right\|_\infty \tag{5.85}$$

where the frequency range could be set as $\omega \in (0, 5]$ in this example.

The settings of the filters Q and F are the same as in the example of two models case:

$$Q(s) = \frac{a_1 s + a_2}{s^2 + a_3 s + a_4}$$
$$F(s) = \frac{b_1 s^2 + b_2 s + b_3}{s^2 + b_4 s + b_5}$$

where a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 and b_5 are the parameters to design.

The initial points settings are $Q(s) = \frac{s+1}{s^2+s+1}$ and $F(s) = \frac{s^2+s+1}{s^2+s+1}$. After the optimization with genetic algorithms for the introduced specifications for the multiple models case, we get

$$Q_{prior}(s) = \frac{424.5s + 449}{s^2 + 3631s + 7830}$$
$$F_{prior}(s) = \frac{0.2778s^2 + 0.0555s + 4.0}{s^2 + 16.63s + 3613}$$

Then, the auxiliary signal generator and post filter considered in practice are

$$Q_{prior}(s) = \frac{424.5s + 449}{s^2 + 3631s + 7830}\max\left(\|d\|_{\text{rms}}\right)$$
$$F_{prior}(s) = \frac{0.2778s^2 + 0.0555s + 4.0}{s^2 + 16.63s + 3613}$$

where $\max\left(\|d\|_{\text{rms}}\right)$ could be evaluated off-line. In this example, $\max\left(\|d\|_{\text{rms}}\right) = 0.9838$.

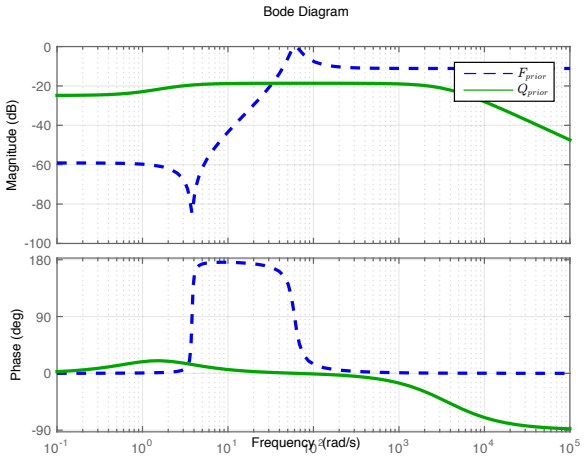

Figure 5.16.: Bode diagram of F_{prior} and Q_{prior}

The bode diagram in Fig. 5.16 shows that the filter F is a high-pass filter. Since the threshold J_{th}^i depends on $\left\|FG_{yd}^i\right\|_\infty$ with the frequency range $\omega \in (0, 5\text{rad/s}]$, the designed high-pass filter F will give smaller thresholds J_{th}^i. As shown in Fig. 5.15 (right), the bode diagrams of different G_i with different J_i are similar in low frequency but different in high frequency. The high-pass filter F will also increase the high frequency part of auxiliary signal on the residual. It is reason why the filter F is a high-pass filter.

The alarms from Fig. 5.17 is collected as a form in Tab. 5.9. The conditions of the alarms in Tab. 5.9 are the same as in Tab. 5.7, which means that filters F_{prior} and Q_{prior} realize the objective of the design. System models P_i could be discriminated by the proposed auxiliary signal and post filter.

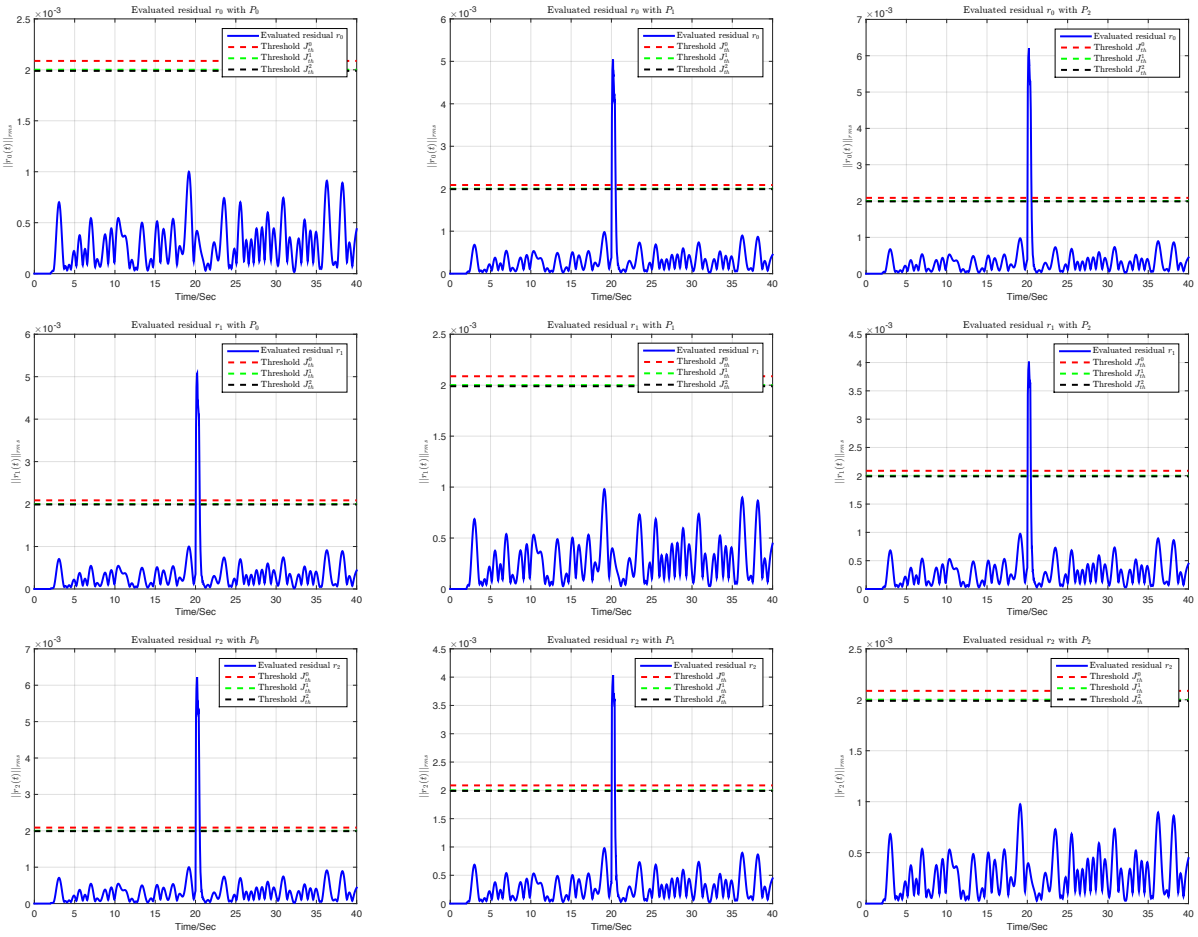

Figure 5.17.: Evaluated residual with F_{prior} and Q_{prior} for different reference models and plant models

$r_j \backslash P_i$	P_0	P_1	P_2
r_0	○ J^0_{th}	× J^0_{th}	× J^0_{th}
r_1	× J^1_{th}	○ J^1_{th}	× J^1_{th}
r_2	× J^2_{th}	× J^2_{th}	○ J^2_{th}

Table 5.9.: Reformulated logic table with improved threshold

Fig. 5.18 shows the effects of the post filter without the auxiliary signal for these three models (P_0, P_1 and P_2). In this case, it is rather difficult to discriminate the system models P_i only by disturbances.

The effects of the designed auxiliary signal on the system with different combinations of reference models G^j_{ref} and models P_i are shown in Fig. 5.19 and Fig. 5.20. Comparing with the effects of disturbances, the effects of the auxiliary signal on the system are small (small $\|u^v_c\|_{peak}$ and small $\|y^v\|_{peak}$), which implies that the design can discriminate the models P_i only with small perturbations.

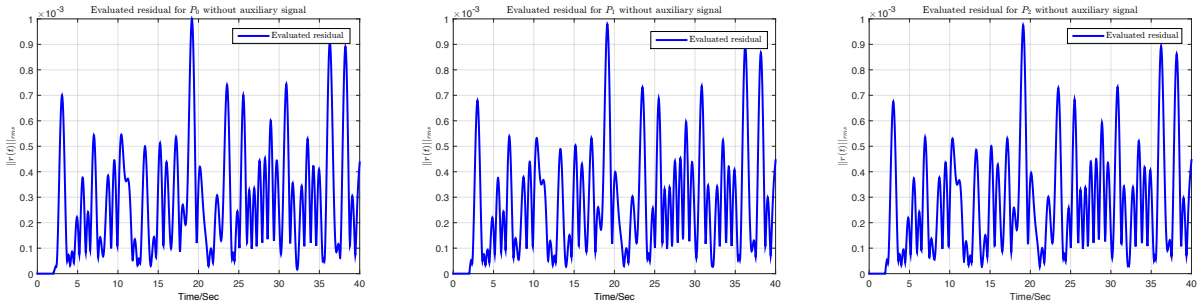

Figure 5.18.: Evaluated residual with F_{prior} and Q_{prior} for different referent models and plant models without auxiliary signal

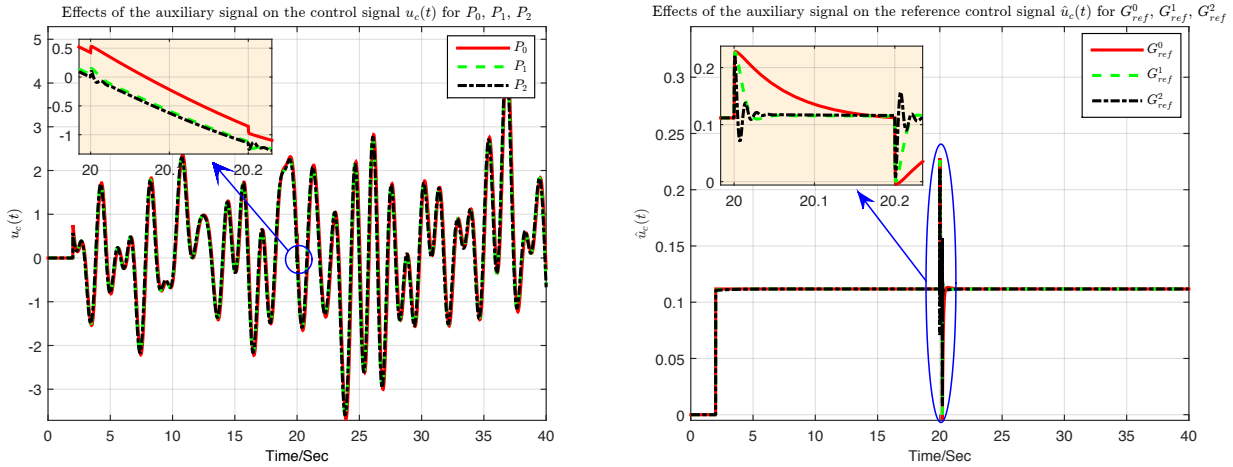

Figure 5.19.: Effects of the auxiliary signal on the control signal u_c and $\hat{u}_c$ for P_i and G^j_{ref}

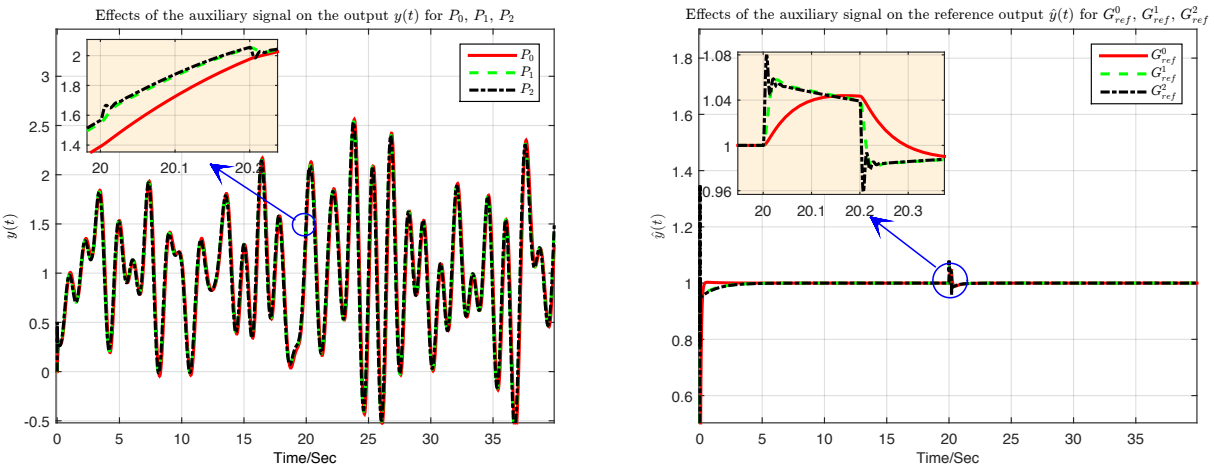

Figure 5.20.: Effects of the auxiliary signal on the output y and $\hat{y}$ for P_i and G^j_{ref}

5.6. Conclusions

In this chapter, a new framework of auxiliary signal design for active fault diagnosis has been presented. An auxiliary signal generator Q and a post filter F are designed to discriminate models with reference models. Different from the previous work in the literature, the following specifications are used to measure the perturbations of the auxiliary signal on the system:

- The peak amplitude of the response from the auxiliary signal to the control signal of the system;
- The peak amplitude of the response from the auxiliary signal to the outputs of the system y.

A model discrimination condition in the worst case is also given to evaluate the effectiveness of the designed auxiliary signal. According to the requirements of model discrimination for two models case, a multiobjective optimization problem is formulated, which could be solved by genetic algorithm. Furthermore, in order to achieve model discrimination for multiple models case, a bank of reference models are introduced to generate a bank of residuals. An decision logic is also given to discriminate the multiple models. The effectiveness of the designs for two models case and multiple models case are proven by simulations with a DC motor control system DR300.

6. Conclusion and perspective

This chapter summarizes the results obtained in the previous chapters, and presents some remarks based on these results. Some future directions for further developments are highlighted next.

6.1. Conclusions

Prior work has documented the effectiveness of different performance indices with different approaches in improving the ability of fault detection, either passive or active fault detection approach. To meet different requirements of a "good" fault detection, different performance indices in time and frequency domain were added into the design, which could be solved by nonsmooth optimization methods. The main work in this dissertation could be concluded with following three aspects:

H_-/H_∞ filter design

Owing to the development of nonsmooth optimization method, the criteria in the worst case H_-/H_∞ are solved by the proposed method. To improve the transients of residual, a constraint of eigenvalue is added into the optimization. The attractive advantages of this method are that the nonsmooth optimization method works faster and could provide a less conservative result than the classical methods. Due to the advantages of nonsmooth optimization method, the nonsmooth optimization method provides an alternative method to design a H_-/H_∞ fault detection filter with some other complex specifications. For a system with multiple models, two different frameworks are proposed to design the H_-/H_∞ fault detection filter:

- A unique observer gain and a unique residual weighting matrix are designed for the system with multiple models. A complex multiobjective functions are formulated to contain the trade-off among different models.
- If the exact information of model is unknown, a new framework of deconvolution filter is introduced. The generation of residual does not depend on the model of system, but is affected by the inputs of system. A new complex multiobjective functions are proposed to optimize. Furthermore, a unique threshold is introduced to work with the robust fault detection filter.

Both formulated problems are solved by nonsmooth optimization approach.

H_-/H_∞ filter design with mixed time and frequency specifications

In order to achieve fast fault detection for some specific faults, a specification in time domain is introduced, which is difficult to be represented by LMI. Comparing with the classical per-

formance index, the specification in time domain adjusts the response in time domain directly with a lower bound envelope and an upper bound envelope, which shows better characteristics than the indirect method. Based on different functions of these two envelopes, two different designs in mixed time and frequency domain are given:

- In order to achieve faster fault detection for step fault and better ability of fault detection in the worst case, an iterative algorithm is introduced to minimize fault detection delay for step fault under a constraint of good ability of fault detection in the worst case. The minimization of fault detection delay is achieved with the lower bound envelope and the ability of fault detection in the worst is measured by the mixed specification H_-/H_∞.
- Furthermore, with the aim of detecting step signal fault faster when fault appears with lower false alarm rate when fault disappears, an upper bound envelope is added into the design to restrict the overshoot of residual responses. The mixed criteria H_-/H_∞ in frequency domain is also considered to improve the ability of fault detection in the worst case. A compromise between these two characteristics is given in the design.

Active fault diagnosis (Model detection and model isolation)

The third objective is to design an auxiliary signal to detect faults, which are difficult to detect in the passive way. The faults are not the additive formulation in system, but change the parameters of the system. In this case, an auxiliary signal is designed to discriminate all the models. One shortage of the active fault diagnosis is that the auxiliary signal will affect the characteristics of the system, therefore, the influence of auxiliary signal on the system should be as small as possible. With the benefits of the introduced theory of fault detection in the previous part, a new framework of active fault diagnosis is given in Chapter 5. An application to the DC motor control system DR300 shows the effectiveness of the proposed design method. From the discussion presented in the previous chapter, the following conclusions can be drawn:

- The active fault diagnosis (model detection for 2 models case and model isolation for multiple models case) can be achieved with the proposed new framework of auxiliary signal design, which has a simple design procedure and requires little on-line computation,
- A peak amplitude specification of response is used to replace the H_2 norm of the response to evaluate the worst effects of auxiliary signal on the system,
- In the closed loop case, the effects of auxiliary signal on the control signal is considered,
- With a bank of reference models and a decision making logic, multiple models can be discriminated.

6.2. Perspectives

The results obtained in this dissertation are summarized in the previous section. The proposed techniques and applications with different specifications to improve the performance of fault detection were briefly described. Besides on the results above, several possible future research directions could be:

1. Chapter 3 proposed H_-/H_∞ filter design by nonsmooth optimization. For the case when G_{rf} is invertible, (3.10) is used to calculate the value of H_-. Although Morre-Penrose

pseudo inverse could be used to extend the application of the proposed method, a direct analysis should be given on the criterion H_-: calculate the Clarke subdifferental of H_-. Then, the proposed nonsmooth optimization could be used to solve all the problems about H_-/H_∞ filter design for LTI system.

2. Notice that in Chapter 4, a specification in time domain was proposed to improve the transients of residual response, where the lower bound envelope was used to specify the transients of response and the threshold is defined with the specification (3.26) in frequency domain. In the second part of Chapter 4, the upper bound envelope was used to decrease the overshoot of response. Another possible application of the upper bound envelope may specify the maximum effects of disturbances on the residual. In this case, the threshold could be defined with the maximum effects of disturbances on residual directly. Then, all the optimization could be realized in the time domain.

3. Chapter 5 designed a "least perturbations" auxiliary signal to discriminate multiple models for SISO LTI systems. One direct extension could be formulating the framework of active fault diagnosis into MIMO case. Since a bank of reference models are used in multiple models case, an idea is to design a unique reference model G_{ref} such that the implements of model discrimination will be easier and more convenient.

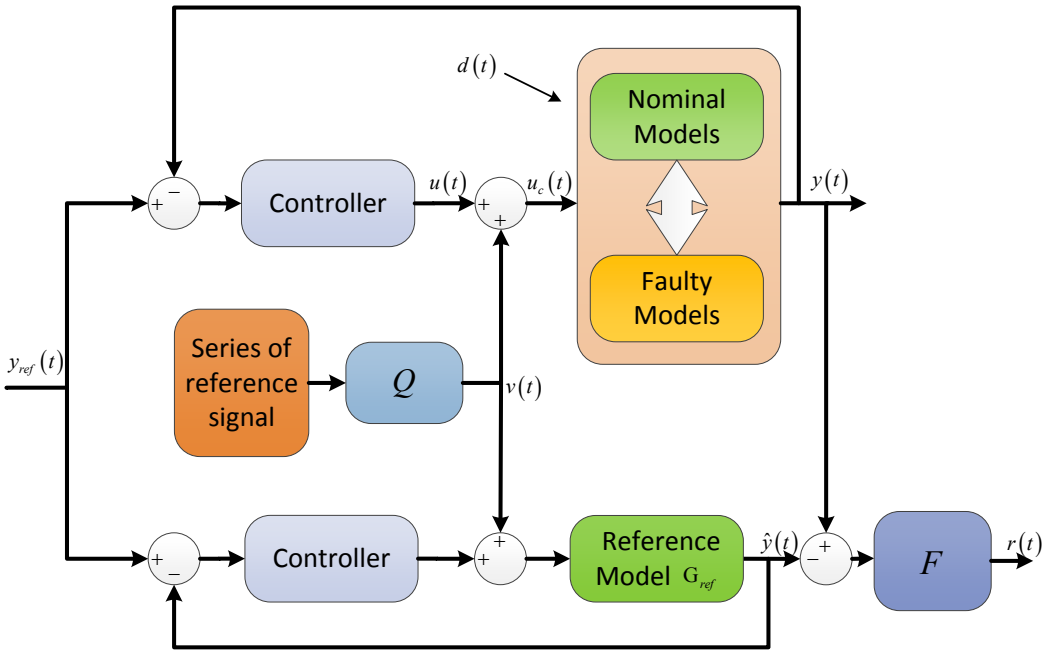

Figure 6.1.: Active diagnosis for multi models with designing referenced model G_{ref}

One typical application of designing reference model is that there are n faulty models P_N^i $(0 \leqslant i \leqslant n-1)$ and m nominal models P_F^j $(0 \leqslant j \leqslant m-1)$, and the user might only care that the system is in nominal models or faulty models but not the exact model P_N^i or P_F^j.

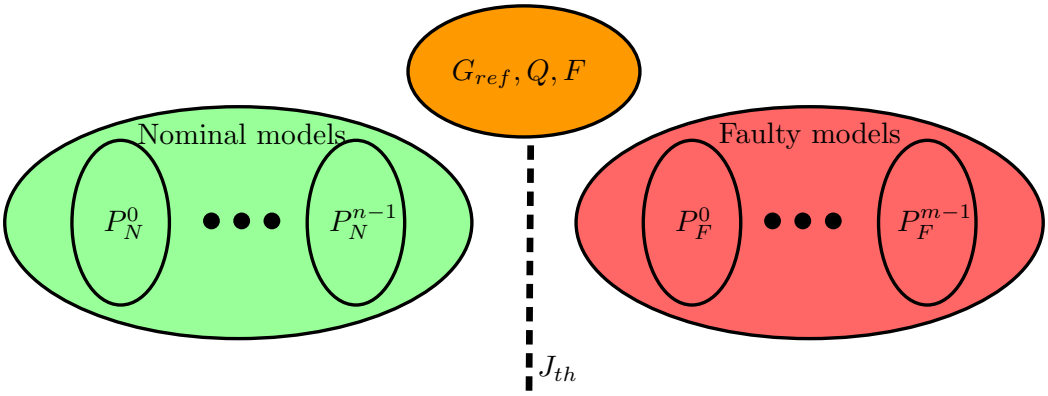

Figure 6.2.: Typical case

In this case, it is necessary to discriminate n (assume $n < m$) faulty models P_N^i to obtain the conclusions with a bank of reference models. By contrast, with a designable reference model, a decision, whether the system is in nominal model or faulty model, could be made with one residual.

4. In the above work, the effects of the finite time window T for RMS evaluation function in (3.22) on the design are not investigated systematically. The selection of finite time window T should be also contained into the design, which may provide additional freedom to improve the obtained results.

A. Appendix: Signal norms and system norms

A.1. Signal norms

A.1.1. Peak ($\mathcal{L}_\infty$ norm)

Generally speaking, one simple and direct interpretation of "the signal is small" is that the signal is bounded by a small value all the time, and the corresponding value is the maximum or peak absolute value (which is also called as the least upper bound of absolute value). The definition of peak or $\mathcal{L}_\infty$ norm of the signal $u(t)$ is

- Scalar case

$$\|u\|_\infty = \sup_{t\geqslant 0} |u(t)|$$

- Vector case

$$\|u\|_\infty = \max_{1\leqslant i\leqslant n} \|u_i\|_\infty = \sup_{t\geqslant 0} \max_{1\leqslant i\leqslant n} |u_i(t)|$$

The peak of a signal is dependent on the biggest values the signal takes on. Therefore, this norm is always used to describe an unknown signal, but some bound on the peak. This kind of description is also called an unknown but bounded model of a signal.

In [19], a variation of peak is proposed to evaluate the peak for the steady state case:

$$\|u\|_{ss\infty} = \lim_{t\to\infty} \sup |u(t)| = \lim_{T\to\infty} \sup_{t\geqslant T} |u(t)|$$

As discussed in [35], the peak amplitude of a vector valued residual is used to compare with the given threshold in FDI study. Then, the peak is improved as the peak norm:

- Scalar case

$$\|u\|_{peak} = \|u\|_\infty$$

- Vector case

$$\|u\|_{peak} = \sup_t \left(u^T(t)\, u(t)\right)^{1/2} \tag{A.1}$$

A.1.2. $\mathcal{L}_1$: Resource consumption

- Scalar case

$$\|u\|_1 = \int_0^\infty |u(t)|\, dt$$

- Vector case

$$\|u\|_1 = \int_0^\infty \sum_{i=1}^{n} |u_i(t)|\, dt = \sum_{i=1}^{n} \|u_i\|_1$$

The $\mathcal{L}_1$ norm could be considered as a measurement of a total resource consumption. It is a kind of criterion for transient signals, which decay to zero as time progresses.

A.1.3. $\mathcal{L}_2$: Square root total energy

- Scalar case

$$\|u\|_2 = \left(\int_0^\infty u(t)^2\, dt \right)^{1/2}$$

- Vector case

$$\|u\|_2 = \left(\int_0^\infty u^T(t)\, u(t)\, dt \right)^{1/2}$$

or

$$\|u\|_2 = \left(\sum_{i=1}^{\infty} u^T(i)\, u(i) \right)^{1/2}$$

This norm is usually used to evaluated the finite energy of the signal. While $\sum_{i=1}^{\infty} u^T(i)\, u(i)$ represents the instantaneous power, $\|u\|_2^2$ means the total energy.

According to Parseval's theorem, the computation of a signal with $\mathcal{L}_1$ norm could be carried out in the frequency domain:

$$\|u\|_2 = \left(\frac{1}{2\pi} \int_0^\infty u^T(-j\omega)\, u(j\omega)\, d\omega \right)^{1/2} \tag{A.2}$$

for the continuous-time signal and

$$\|u\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} u^T \left(e^{-j\theta} \right) u \left(e^{j\theta} \right) d\theta \right)^{1/2}$$

for the discrete-time signal.

A.1.4. Root mean square (RMS)

Comparing with $\mathcal{L}_2$ norm, the average size of a signal, named as root mean square (RMS) value, is more important in practice:

- Scalar case

$$\|u\|_{rms} = \left(\lim_{T \to \infty} \frac{1}{T} \int_0^T u(t)^2 \, dt \right)^{1/2}$$

- Vector case

$$\|u\|_{rms} = \left(\lim_{T \to \infty} \frac{1}{T} \int_0^T u(t)^T u(t) \, dt \right)^{1/2}$$

which are widely used in many areas of engineering.

While the $\mathcal{L}_2$ norm measures the total energy for a signal, the RMS norm can be thought of as measuring its average power. Even a signal with large peaks, whose RMS norm may be small. Thus, $\|u\|_\infty$ is more affected than $\|u\|_{rms}$ by the large but infrequent values of the signal.

Remark A.1. From a practical point of view, it is impossible to evaluate the signal over the whole time or frequency domain in reality, a window could be introduced. Then, the previous criterion could be:

$$\|u\|_{peak} = \sup_{t \in \tau} |u(t)| \tag{A.3}$$

$$\|u\|_{\infty,\tau} = \sup_{t \in \tau} \max_{1 \leqslant i \leqslant n} |u_i(t)|$$

$$\|u\|_{1,\tau} = \int_{t_1}^{t_2} \sum_{i=1}^{n} |u_i(t)| \, dt$$

$$\|u\|_{2,\tau} = \left(\int_{t_1}^{t_2} u^T(t) \, u(t) \, dt \right)^{1/2} \tag{A.4}$$

$$\|u\|_{2,\varphi} = \left(\frac{1}{2\pi} \int_{\omega_1}^{\omega_2} u^T(-j\omega) \, u(j\omega) \, d\omega \right)^{1/2}$$

where $\tau = (t_1, t_2)$ and $\varphi = (\omega_1, \omega_2)$ represent the interesting time and frequency domain windows.

A.2. System norms

In this section, the performance index to evaluate the size of a system with input u, output y, and transfer matrix G, are introduced.

Consider LTI systems $y(p) = G(p)\,u(p)$, which are causal and stable. A causal system means $G(t) = 0$ for $t < 0$ with $G(t)$ as impulse response. Causality requires that $G(p)$ is proper:

$$\lim_{p\to\infty} G(p) < \infty$$

A system is strictly proper if

$$\lim_{p\to\infty} G(p) = 0$$

A.2.1. $\mathcal{H}_2$ norm

If $G(s) \in \mathcal{L}_2$, the $\mathcal{L}_2$ norm of G is defined as

$$\begin{aligned}\|G\|_2 &= \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\left\{G^*(j\omega)\,G(j\omega)\right\} d\omega\right)^{1/2} \\ &= \|g\|_2 \\ &= \left(\int_{-\infty}^{\infty} \operatorname{trace}\left\{g^*(t)\,g(t)\right\} dt\right)^{1/2}\end{aligned}$$

$\mathcal{H}_2$ norm is widely used in the control theory. One application of $\mathcal{H}_2$ norm is to design the optimal Kalman filter.

A.2.2. $\mathcal{H}_\infty$ norm

If $G(s) \in \mathcal{RL}_\infty$, the $\mathcal{L}_\infty$ norm of G is defined as

$$\|G\|_\infty = \sup_{\omega\in[0,\infty]} \overline{\sigma}\left\{G(j\omega)\right\} \tag{A.5}$$

which also could be interpreted from the mathematical viewpoint:

$$\sup_{u\neq 0}\frac{\|y\|_2}{\|u\|_2} = \sup_{u\neq 0}\frac{\|Gu\|_2}{\|u\|_2} = \sup_{\omega\in[0,\infty]} \overline{\sigma}\left\{G(j\omega)\right\}$$

which means that the $\mathcal{H}_\infty$ norm is the amplification of a transfer function matrix G that maps the inputs signal u with finite energy into the output signal. In the view of FDI, the $\mathcal{H}_\infty$ norm means the biggest effects of the inputs on the outputs.

It is also interesting to note that if $\|u\|_{\text{rms}}$ is bounded even $\|u\|_2$ is not bounded, there is a relationship:

$$\sup_{u\neq 0}\frac{\|y\|_{\text{rms}}}{\|u\|_{\text{rms}}} = \sup_{\omega\in[0,\infty]} \overline{\sigma}\left\{G\left(j\omega\right)\right\} = \|G\|_\infty$$

the proof could be seen in[19].

A.2.3. Generalized $\mathcal{H}_2$ norm

$$\|G\|_g = \sup_{u\neq 0}\frac{\|y\|_{peak}}{\|u\|_2}$$

which is used to evaluate the effects of the bounded energy input on the power changer in the output.

B. Appendix: Basic knowledge to calculate the subdifferential

[10, 7] apply the nonsmooth programming techniques for controller design. A series of improvement has been done by Apkarian and his team. Several solvers in matlab, like Systune, Hinfstruct and Looptune are proposed with the nonsmooth programming techniques. Because of the comparability between the control and FDI, the nonsmooth programming techniques could be used to solve some FDI problems with the aid of nice characteristics of these techniques.

The following notions are discussed in [31], which also could be considered in [16, 98].

B (x,r) represents an open sphere with the center $x \in \mathbb{R}^n$ and the radius $r > 0$:

$$B(x,r) \triangleq \{y \in T^n : \|y - x\| < r\} \tag{B.1}$$

Definition B.1. Let Y be a subset of X. A function $f : Y \to \mathbb{R}$ is said to satisfy a Lipschitz condition (on Y) provided that, for some nonnegative scalar K, one has

$$|f(y) - f(y') \leqslant K \|y - y'\|| \tag{B.2}$$

for all points y, y' in Y; this is also referred to as a Lipschitz condition of rank K. The function f is Lipschitz (of rank K) near x if, for some $\varepsilon > 0$, f satisfies a Lipschitz condition on the set $B(x, \varepsilon)$.

A function of a real variable, having Lipschitz condition near a point need not to be differentiable there, nor need it admit directional derivatives.

Definition B.2. The directional derivative of the function $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ if

$$\lim_{\substack{t \to 0 \\ t>0}} \frac{f(x + td) - f(x)}{t}$$

exits and is finite, denoted as $f'(x, d)$.

If f is differentiable at x, it contains the directional derivatives in all the directions d, and we have $f'(x, d) = f'(x)(d) = \nabla f(x)^T$. The inverse is not true except the directional derivatives are continues.

Definition B.3. A function $\sigma : \mathbb{R}^n \mapsto \mathbb{R}$ is sublinear if it satisfies

1. Subadditivity:

$$\sigma(x + y) \leq \sigma(x) + \sigma(y) \text{ for all } x, y \in \mathbb{R}^n$$

2. Positive homogeneity:

$$\sigma(tx) \leq t\sigma(x) \text{ for all } x \in \mathbb{R}^n \text{and } t > 0$$

Theorem B.1. *If $\sigma : \mathbb{R}^n \mapsto \mathbb{R}$ is sublinear, then the set*

$$S_\sigma \triangleq \{s \in \mathbb{R}^n : \langle s, x\rangle \leq \sigma(x) \text{ for all } x \in \mathbb{R}^n\}$$

is not empty, compact and convex. It has the relation

$$\sigma(x) = \sup\{\langle s, x\rangle : s \in S_\sigma\}$$

Vice versa.

Convex analysis

With the definition of convex, all the convex function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is local Lipschitz at all the points $\mathbb{R}^n$.

Theorem B.2. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a convex function. The classical directional derivative $f'(x, d)$ exists in every direction $d \in \mathbb{R}^n$ and it satisfies*

$$f'(x, d) = \inf_{t>0} \frac{f(x + td) - f(x)}{t}$$

for all points $x \in \mathbb{R}^n$.

Then, the definition of subdifferential could be got with directional derivative:

Theorem B.3. *The subdifferential of a convex function $f : \mathbb{R}^n \mapsto \mathbb{R}$ at $x \in \mathbb{R}^n$ is the set $\partial_c f(x)$ of vectors $s \in \mathbb{R}^n$ such that*

$$\partial_c f(x) = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq f'(x, d) \text{ for all } d \in \mathbb{R}^n\}$$

each vector $s \in \partial_c f(x)$ is called a subgradient of f at x.

There is a relationship between the subdifferential and the directional derivative:

$$f'(x, d) = \max\{\langle s, d\rangle : s \in \partial_c f(x)\}$$

When the function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex and differential at x, $\partial_c f(x)$ will be $\nabla f(x)$.

Nonconvex analysis

For the locally Lipschitz continuous functions, may be the classical directional derivatives do not exist. The definition of generalized directional derivative is introduced:

Definition B.4. $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a locally Lipschitz continuous function at $x \in \mathbb{R}^n$. The generalized directional derivative of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f^\circ (x, d) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f (x + td) - f (x)}{t}$$

The generalized directional derivative always exists for locally Lipshcitz continuous functions, and it is sublinear. Then, the subdifferential for nonconvex locally Lipschitz continuous functions with the directional derivative replaced by the generalized directional derivative.

Definition B.5. $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a locally Lipschitz continuous function at $x \in \mathbb{R}^n$. Then the subdifferential of f at x is the set $\partial f (x)$ of vectors $s \in \mathbb{R}^n$ such that

$$\partial f (x) = \{s \in \mathbb{R}^n | f^\circ (x, d) \geq \langle s, d \rangle \text{ for all } d \in \mathbb{R}^n\}$$

Each vector $s \in \partial f (x)$ is called a subgradient of f at x.

Some properties of the subdifferential both in convex and nonconvex cases are summarized:

- The subdifferential $\partial_c f (x)$ for a convex function f is a nonempty, convex, and compact set such that $\partial_c f (x) \in B (0, L)$;
- The subdifferential $\partial f (x)$ for a a locally Lipschitz continuous function f is a nonempty, convex, and compact set such that $\partial f (x) \in B (0, L)$, where $L > 0$ is Lipschitz constant of f at x. Moreover, $f^\circ (x, d) = \max \{\langle s, d \rangle : s \in \partial f (x)\}$ for all $d \in \mathbb{R}^n$;
- The subdifferential for locally Lipschitz continuous functions is a generalization of the subdifferential for convex functions: If $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function, then $f^{'} (x, d) = f^\circ (x, d)$ for all $d \in \mathbb{R}^n$ and $\partial_c f (x) = \partial f (x)$.
- The subdifferential for locally Lipschitz continuous functions is a generalization of the classical derivative: If $f : \mathbb{R}^n \mapsto \mathbb{R}$ is both locally Lipschitz continuous and differentiable at $x \in \mathbb{R}^n$, then $\nabla f (x) \in \partial f (x)$. If, in addition, $f : \mathbb{R}^n \mapsto \mathbb{R}$ is continuously differentiable at $x \in \mathbb{R}^n$, then $\partial f (x) = \{\nabla f (x)\}$.

Bibliography

[1] (2013). *Robust control toolbox 5.0*. MathWorks, Natick, MA, USA.

[2] Abdo, A. (2013). *Fault Detection Schemes for Switched Systems*. Shaker.

[3] Abdo, A., Damlakhi, W., Saijai, J., and Ding, S. X. (2010). Design of robust fault detection filter for hybrid switched systems. In *Proceedings of the 2010 Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 161–166, Nice, France,.

[4] Abdo, A., Ding, S. X., Damlakhi, W., and Saijai, J. (2011). Robust fault detection filter design for uncertain switched systems with adaptive threshold setting. In *Proceedings of the 45th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 5467–5472.

[5] Andjelkovic, I. (2008). *Auxiliary Signal Design for Fault Detection for Nonlinear Systems: Direct Approach*. PhD thesis, North Carolina State university.

[6] Andjelkovic, I. and Campbell, S. L. (2011). Direct optimization determination of auxiliary test signals for linear problems with model uncertainty. In *Proceedings of the 50th Conference on Decision and Control and European Control Conference (CDC-ECC)*.

[7] Apkarian, P. (2012). Tuning controllers against multiple design requirements. In *Proceedings of the 16th International Conference on System Theory, Control and Computing (ICSTCC)*, pages 1–6, Sinaia, Romania.

[8] Apkarian, P. and Noll, D. (2005). Nonsmooth optimization for multidisk H_∞ synthesis. *European Journal of Control*.

[9] Apkarian, P. and Noll, D. (2006a). Controller design via nonsmooth multidirectional search. *SIAM Journal on Control and Optimization*, 44(6):1923–1949.

[10] Apkarian, P. and Noll, D. (2006b). Nonsmooth H_∞ synthesis. *IEEE Transactions on Automatic Control*, 51(1):71–86.

[11] Apkarian, P., Noll, D., and Rondepierre, A. (2009). Mixed H_2/H_∞ control via nonsmooth optimization. In *Proceedings of the Annual Conference on Decision and Control and Chinese control conference (CDC-CCC)*, pages 6460–6465, Shanghai, P.R. China.

[12] Arturo, M. C. (2005). *Multiobjective Control: Linear Matrix Inequality Techniques and Genetic Algorithms Approach*. PhD thesis, University of Sheffield.

[13] Bagirov, A., Karmitsa, N., and Mäkelä, M. M. (2014). *Introduction to Nonsmooth Optimization*. Springer.

[14] Basseville, M. and Nikiforov, I. V. (1993). *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs.

[15] Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. SIAM.

[16] Bompart, V. (2007). *Optimisation non lisse pour la commande des systèmes de l'aéronautique*. PhD thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier.

[17] Bompart, V., Apkarian, P., and Noll, D. (2008). Control design in the time and frequency domain using nonsmooth techniques. *Systems & Control Letters*, 57(3):271–282.

[18] Bouattour, M., Chadli, M., Chaabane, M., and Hajjaji, A. E. (2011). Design of robust fault detection observer for Takagi-Sugeno models using the descriptor approach. *International Journal of Control, Automation and Systems*, 9(5):973–979.

[19] Boyd, S. and Barratt, C. (1991). *Linear controller design: limits of performance*. Prentice Hall PTR.

[20] Busch, R. and Peddle, I. K. (2014). Active fault detection for open loop stable LTI SISO systems. *International Journal of Control, Automation and Systems*, 12(2):324–332.

[21] Campbell, S. L., Drake, K., and Nikoukhah, R. (2002). Auxiliary signal design for multi-model identification in systems with multiple delays. In *Proceedings of the IEEE Mediterranean Conference on Control and Automation*.

[22] Campbell, S. L. and Nikoukhah, R. (2004). *Auxiliary signal design for failure detection*. Princeton University Press.

[23] Casavola, A., Famularo, D., and Franze, G. (2005). A robust deconvolution scheme for fault detection and isolation of uncertain linear systems: an LMI approach. *Automatica*, 41:1463–1472.

[24] Casavola, A., Famularo, D., and Franze, G. (2008). Robust fault detection of uncertain linear systems via quasi-LMIs. *Automatica*, 44:289–295.

[25] Chen, J. and Patton, R. J. (2012). *Robust model-based fault diagnosis for dynamic systems*. Springer Publishing Company, Incorporated.

[26] Chen, J., Patton, R. J., and Liu, G. P. (1996a). Optimal residual design for fault diagnosis using multi-objective optimization and genetic algorithms. *International Journal of Systems Science*, 27(6):567–576.

[27] Chen, J. and Patton, T. J. (1999). *Robust Model-Based Fault Diagnosis For Dynamic Systems*. Kluwer.

[28] Chen, J., Patton, T. J., and Liu, G. P. (1996b). Optimal residual design for fault diagnosis using multi-objective optimization and genetic algorithms. *International Journal of Systems Science*, 27(6):567–576.

[29] Chiang, L. H., Braatz, R. D., and Russell, E. L. (2001). *Fault detection and diagnosis in industrial systems*. Springer.

[30] Chow, E. and Willsky, A. S. (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control*, 29(7):603–614.

[31] Clarke, F. H. (1990). *Optimization and nonsmooth analysis*, volume 5. SIAM.

[32] Cui, Y., Huang, X. H., and Wang, M. (2009). Multi-objective robust fault detection filter design in a finite frequency range. In *Proceedings of the 6th Advances in Neural Networks*, pages 733–743.

[33] Dennis, J., John, E., and Schnabel, R. B. (1996). *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16. SIAM.

[34] Ding, S. X. (2009). Integrated design of feedback controllers and fault detectors. *Annual reviews in control*, 33(2):124–135.

[35] Ding, S. X. (2013). *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer.

[36] Ding, S. X., Jeinsch, T., Frank, P. M., and Ding, L. (2000). A unified approach to the optimization of fault detection systems. *International Journal of Adaptive Control and Signal Processing*, 14(7):725–745.

[37] Ding, X., Guo, L., and Jeinsch, T. (1999). A characterization of parity space and its application to robust fault detection. *IEEE Transactions on Automatic Control*, 44(2):337–343.

[38] Esna, A. A., Nikoukhah, R., and Campbell, S. L. (2011). Auxiliary signal design for robust active fault detection of linear discrete-time systems. *Automatica*, 47(9):1887–1895.

[39] Esna, A. A., Nikoukhah, R., and Campbell, S. L. (2012). Active robust fault detection in closed-loop systems: Quadratic optimization approach. *IEEE Transactions on Automatic Control*, 57(10):2532–2544.

[40] Frank, P. M. and Ding, X. (1997). Survey of robust residual generation and evaluation methods in observer-based fault detection systems. *Journal of process control*, 7(6):403–424.

[41] Frank, P. M. and Ding, X. C. (1994). Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. *Automatica*, 30(5):789–804.

[42] Frisk, E. and Nielsen, L. (2006). Robust residual generation for diagnosis including a reference model for residual behavior. *Automatica*, 42(3):437–445.

[43] Gahinet, P. and Apkarian, P. (2011). Decentralized and fixed-structure H_∞ control in MATLAB. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 8205–8210, Orlando, FL, USA.

[44] Gen, M. and Cheng, R. W. (2000). *Genetic algorithms and engineering optimization*, volume 7. John Wiley & Sons.

[45] Ghaoui, L. E. I. and Niculescu, S. I. (2000). *Advances in linear matrix inequality methods in control*, volume 2. SIAM.

[46] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition.

[47] Guo, J. C., Huang, X. H., and Cui, Y. (2009). Design and analysis of robust fault detection filter using LMI tools. *Computers and Biomedical Research Mathematics with Applications*, 57(11):1743–1747.

[48] Hamelin, F., Sauter, D., and Theilliol, D. (1995). Some extensions to the parity space method for FDI using alternated projection subspaces. In *Proceedings of the 34th IEEE Conference on Decision and Control (CDC)*, volume 3, pages 2407–2412.

[49] Haupt, R. L. and Haupt, S. E. (2004). *Practical genetic algorithms*. John Wiley & Sons.

[50] Henrion, D. and Lasserre, J. B. (2004). Solving nonconvex optimization problems. *Control Systems, IEEE*, 24(3):72–83.

[51] Henry, D. and Zolghadri, A. (2005). Design of fault diagnosis filters: A multi-objective approach. *Journal of the Franklin Institute*, 342(4):421–446.

[52] Henry, D. and Zolghadri, A. (2006). Norm-based design of robust fdi schemes for uncertain systems under feedback control: comparison of two approaches. *Control engineering practice*, 14(9):1081–1097.

[53] Henry, D., Zolghadri, A., Castang, F., and Monsion, M. (2002). A multi-objective filtering approach for fault diagnosis with guaranteed sensitivity performance. In *Proceedings of the 15th IFAC World Congress*.

[54] Hou, M. and Patton, T. J. (1996). An LMI approach to H_-/H_∞ fault detection observers. In *Proceedings of the UKACC International Conference on Control*, volume 1, pages 305–310.

[55] Izadi, I., Chen, T. W., and Zhao, Q. (2005). Norm invariant discretization for sampled-data fault detection. *Automatica*, 41(9):1633–1637.

[56] Izadi, I., Zhao, Q., and Chen, T. (2006). An H_∞ approach to fast rate fault detection for multirate sampled-data systems. *Journal of process control*, 16(6):651–658.

[57] Jaimoukha, I. M., Li, Z. H., and Vasilios, P. (2006). A matrix factorization solution to the fault detection problem. *Automatica*, 42(11):1907 – 1912.

[58] Khosrowjerdi, M. J., Nikoukhah, R., and Safari-Shad, N. (2004). A mixed H_2/H_∞ approach to simultaneous fault detection and control. *Automatica*, 40(2):261–267.

[59] Khosrowjerdi, M. J., Nikoukhah, R., and Safari-Shad, N. (2005). Fault detection in a mixed H_2/H_∞ setting. *IEEE Transactions on Automatic Control*, 50(7):1063–1068.

[60] Kim, Y. M. and Watkins, J. M. (2007). A new approach for robust and reduced order fault detection filter design. In *Proceedings of the American control conference (ACC)*.

[61] Kowalczuk, Z. and Białaszewski, T. (2004). Genetic algorithms in the multi-objective optimisation of fault detection observers. In *Fault Diagnosis: Models, Artificial Intelligence, Applications*, pages 511–556. Springer.

[62] Kowalczuk, Z. and Białaszewski, T. (2014). Gender approach to multi-objective optimization of detection systems with pre-selection of criteria. In *Intelligent Systems in Technical and Medical Diagnostics*, pages 161–174. Springer.

[63] Kowalczuk, Z., Suchomski, P., and Białaszewski, T. (1999). Evolutionary multi-objective pareto optimisation of diagnostic state observers. *International Journal of Applied Mathematics and Computer Science*, 9:689–709.

[64] Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817.

[65] Lasserre, J. B. (2002). An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM Journal on Optimization*, 12(3):756–769.

[66] Li, X. B. (2009). *Fault detection filter design for linear systems*. PhD thesis, Louisiana State University, USA.

[67] Li, X. J. and Yang, G. H. (2009). LMI-based H_-/H_∞ observer design in low frequency domain with application in fault detection. In *Proceedings of the American Control Conference (ACC)*, pages 4316–4321.

[68] Li, Z. H., Jaimoukha, I. M., and Mazars, E. (2006). A reference model based robust H_∞ filtering approach to fault detection in uncertain systems. In *Proceedings of the Fault Detection, Supervision and Safety of Technical Processes*, volume 6, pages 1062–1067.

[69] Liu, J., Wang, J. L., and Yang, G. H. (2005). An LMI approach to minimum sensitivity analysis with application to fault detection. *Automatica*, 41(11):1995–2004.

[70] Liu, N. (2008). *Optimal Robust Fault Detection*. PhD thesis, Louisiana State University, US, 2005.

[71] Liu, N. and Zhou, K. M. (2007). Optimal solutions to multi-objective robust fault detection problems. In *Proceedings of the 46th IEEE Conference on Decision and Control (CDC)*, pages 981–988.

[72] Marler, R. T. and Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395.

[73] Marseglia, G. R., Scott, J. K., Magni, L., Braatz, R. D., and Raimondo, D. M. (2014). A hybrid stochastic-deterministic optimization for active fault diagnosis using scenario optimization. In *Proceedings of the 19th IFAC World Congress*.

[74] Mazars, E., Jaimoukha, I. M., and Li, Z. H. (2008). Computation of a reference model for robust fault detection and isolation residual generation. *Journal of Control Science and Engineering*.

[75] Mazars, E., Li, Z. H., and Jaimoukha, I. M. (2006). A QMI approach to the robust fault detection and isolation problem. In *Proceedings of the 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*.

[76] Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press.

[77] Nguang, S. K., Zhang, P., and Ding, S. X. (2007). Parity relation based fault estimation for nonlinear systems: An LMI approach. *International Journal of Automation and Computing*, 4(2):164–168.

[78] Niemann, H. (2005). Fault tolerant control based on active fault diagnosis. In *Proceedings of the American Control Conference (ACC)*, pages 2224–2229.

[79] Niemann, H. (2006). A setup for active fault diagnosis. *IEEE Transactions on Automatic Control*, 51(9):1572–1578.

[80] Niemann, H. and Poulsen, N. K. (2014). Active fault detection in MIMO systems. In *Proceedings of the American Control Conference (ACC)*, pages 1975–1980.

[81] Niemann, H. and Stoustrup, J. (2005). An architecture for fault tolerant controllers. *International Journal of Control*, 78(14):1091–1110.

[82] Niemann, H., Stoustrup, J., and Poulsen, N. K. (2014). Controller modification applied for active fault detection. In *Proceedings of the American Control Conference (ACC)*, pages 1963–1968.

[83] Nikoukhah, R. and Campbell, S. L. (2002). Active failure detection: Auxiliary signal design and on-line detection. In *Proceedings of the IEEE Mediterranean Conference on Control and Automation*.

[84] Nobrega, E. G., Abdalla, M. O., and Grigoriadis, K. M. (2000). LMI-based filter design for fault detection and isolation. In *Proceedings of the 39th IEEE Conference on Decision and Control (CDC)*, volume 5, pages 4329–4334, Sydney, Australia.

[85] Nobrega, E. G., Abdalla, M. O., and Grigoriadis, K. M. (2008). Robust fault estimation of uncertain systems using an LMI-based approach. *International Journal of Robust and Nonlinear Control*, 18(18):1657–1680.

[86] Patton, R. J., Korbicz, J., Witczak, M., and Uppal, F. (2005). Combined computational intelligence and analytical methods in fault diagnosis. *Intelligent Control Systems*, 70:349–392.

[87] Polak, E. (2012). *Optimization: Algorithms and Consistent Approximations*. Applied Mathematical Sciences. Springer New York.

[88] Poulsen, N. K. and Niemann, H. (2007). Stochastic change detection based on an active fault diagnosis approach. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, pages 346–351.

[89] Qiu, Z. Y. and Gertler, J. (1993). Robust FDI systems and H_∞-optimization-disturbances and tall fault case. In *Proceedings of the 32nd IEEE Conference on Decision and Control (CDC)*, pages 1710–1715.

[90] Raimondo, D. M., Braatz, R. D., and Scott, J. K. (2013). Active fault diagnosis using moving horizon input design. In *Proceedings of the European Control Conference (ECC)*, pages 3131–3136.

[91] Rambeaux, F., Hamelin, F., and Sauter, D. (1999). Robust residual generation via LMI. In *Proceedings of the 14th world congress of IFAC*, pages 240–246.

[92] Rambeaux, F., Hamelin, F., and Sauter, D. (2001). *Génération et évaluation de résidus pour le Diagnostic de systèmes incertains: approche Fréquentielle*. PhD thesis, L'Université Henri poincaré, Nancy 1.

[93] Rank, M. and Niemann, H. (1999). Norm based design of fault detectors. *International Journal of Control*, 72(9):773–783.

[94] Sauter, D. and Hamelin, F. (1999). Frequency-domain optimization for robust fault detection and isolation in dynamic systems. *IEEE Transactions on Automatic Control*, 44(4):878–882.

[95] Scherer, C., Gahinet, P., and Chilali, M. (1997). Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, 42(7):896–911.

[96] Scott, J. K., Findeisen, R., Braatz, R. D., and Raimondo, D. M. (2013). Design of active inputs for set-based fault diagnosis. In *Proceedings of the American Control Conference (ACC)*, pages 3561–3566.

[97] Simandl, M. and Puncochar, I. (2009). Active fault detection and control: Unified formulation and optimal design. *Automatica*, 45(9):2052–2059.

[98] Simões, A. M. (2009). *Synthèse de compensateurs structurés par l'optimisation non lisse*. PhD thesis, Toulouse University.

[99] Simões, A. M., Apkarian, P., and Noll, D. (2008). Non-smooth progress function algorithm for frequency-shaping control design. *IET Control Theory & Applications*, 2(4):323–336.

[100] Stoorvogel, A. A., Niemann, H. H., Saberi, A., and Sannuti, P. (2002). Optimal fault signal estimation. *Interntional journal of robust and nonlinear control*, 12:697–727.

[101] Tabatabaeipour, S. M. (2013). Active fault detection and isolation of discrete-time linear time-varying systems: a set-membership approach. *International Journal of Systems Science*, pages 1–17.

[102] Tabatabaeipour, S. M., Ravn, A., Izadi-Zamanabadi, R., and Bak, T. (2009a). Active fault diagnosis of linear hybrid systems. In *Proceedings of Safeprocess*, pages 211–216.

[103] Tabatabaeipour, S. M., Ravn, A. P., Izadi-zamabadi, R., and Bak, T. (2009b). Active diagnosis of hybrid systems: A model predictive approach. In *Proceedings of the IEEE International Conference on Control and Automation (ICCA)*, pages 465–470.

[104] Tao, F. and Zhao, Q. (2005). Fault detection observer design with unknown inputs. In *Proceedings of the IEEE Conference on Control Applications,.*, pages 1275–1280.

[105] Varrier, S., Koenig, D., and Martinez, J. (2012a). Robust fault detection for vehicle lateral dynamics. In *Proceedings of the 51st IEEE Annual Conference on Decision and Control (CDC)*, pages 4366–4371, Maui, Hawaii.

[106] Varrier, S., Koenig, D., and Martinez, J. (2014). Robust fault detection for uncertain unknown inputs LPV system. *Control Engineering Practice*, 22:125–134.

[107] Varrier, S., Koenig, D., and Martinez, J. J. (2012b). A parity space-based fault detection on lpv systems: Approach for vehicle lateral dynamics control system. In *Proceedings of the Fault Detection, Supervision and Safety of Technical Processes*, volume 8, pages 1191–1196.

[108] Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S. N. (2003a). A review of process fault detection and diagnosis: Part ii: Qualitative models and search strategies. *Computers & Chemical Engineering*, 27(3):313–326.

[109] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., and Yin, K. W. (2003b). A review of process fault detection and diagnosis: Part iii: Process history based methods. *Computers & chemical engineering*, 27(3):327–346.

[110] Venkatasubramanian, V., Rengaswamy, R., Yin, K. W., and Kavuri, S. N. (2003c). A review of process fault detection and diagnosis: Part i: Quantitative model-based methods. *Computers & chemical engineering*, 27(3):293–311.

[111] Wahrburg, A. and Adamy, J. (2013). LMI-based design of robust fault isolation filters for linear systems. In *Proceedings of the European Control Conference (ECC)*, pages 943–949.

[112] Wang, H., Lam, J., Ding, S. X., and Zhong, M. (2005). Iterative linear matrix inequality algorithms for fault detection with unknown inputs. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 219(2):161–172.

[113] Wang, H. and Yang, G. H. (2007). Fault detection observer design in low frequency domain. In *Proceedings of the IEEE International Conference on Control Applications*, pages 976–981.

[114] Wang, H. and Yang, G. H. (2008). A finite frequency domain approach to fault detection observer design for linear continuous-time systems. *Asian Journal of Control*, 10(5):559–568.

[115] Wang, H. B., Wang, J. L., and Lam, J. (2007a). Worst-case fault detection observer design: Optimization approach. *Journal of Optimization Theory and Applications*, 132(3):475–491.

[116] Wang, J. L., Yang, G. H., and Liu, J. (2007b). An LMI approach to H_- index and mixed H_-/H_∞ fault detection observer design. *Automatica*, 43(9):1656–1665.

[117] Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12(6):601–611.

[118] Wuennenberg, J. (1990). *Observer-based fault detection in dynamic systems*. PhD thesis, University of Duisburg.

[119] Yang, J. W., Hamelin, F., and Sauter, D. (2013). Mixed H_-/H_∞ fault detection observer design for multi model systems via nonsmooth optimization approach. In *Proceedings of the Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 164–171.

[120] Ye, H., Ding, S. X., and Wang, G. Z. (2002). Integrated design of fault detection systems in time-frequency domain. *IEEE Transactions on Automatic Control*, 47(2):384–390.

[121] Ye, H., Wang, G. Z., and Ding, S. X. (2004). A new parity space approach for fault detection based on stationary wavelet transform. *IEEE Transactions on Automatic Control*, 49(2):281–287.

[122] Ye, H., Zhang, P., Ding, S. X., and Wang, G. Z. (2000). A time-frequency domain fault detection approach based on parity relation and wavelet transform. In *Proceedings of the 39th IEEE Conference on Decision and Control (CDC)*, volume 4, pages 4156–4161.

[123] Zhang, P. and Ding, S. X. (2006). Multiobjective design of robust fault detection systems. In *Proceedings of the 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, volume 6, pages 1378–1383.

[124] Zhang, P., Ye, H., Ding, S. X., Wang, G. Z., and Zhou, D. H. (2006). On the relationship between parity space and H_2 approaches to fault detection. *Systems & control letters*, 55(2):94–100.

[125] Zhong, M. Y., Ding, S. X., Lam, J., and Wang, H. B. (2003). An LMI approach to design robust fault detection filter for uncertain LTI systems. *Automatica*, 39(3):543–550.

[126] Zhou, J. (2010). A new robust filter design for robust fault detection in an aerocraft's system. In *Proceedings of the International Conference on Logistics Systems and Intelligent Management*, volume 1, pages 446–450.

[127] Zhou, K. M. and Doyle, J. C. (1998). *Essentials of robust control*, volume 104. Prentice hall Upper Saddle River, NJ.

Résumé

Cette thèse consiste à utiliser des méthodes d'optimisation non lisse à des fins de diagnostic de défauts.

Dans un premier temps, afin de surmonter les inconvénients des méthodes classiques, une approche fondée sur l'optimisation non lisse est présentée. Elle permet de résoudre le problème de détection de défauts dans le pire des cas. La rapidité de la réponse résiduelle peut y être intégrée en tant que contrainte. Le diagnostic des systèmes à commutation est ensuite considéré via un générateur de résidus. Dans le cas d'un modèle connu avec certitude, un filtre de détection de défauts robuste aux perturbations est enfin construit.

Dans la seconde partie de la thèse, une méthode est proposée afin de concevoir un observateur permettant de détecter des défauts dans un cas général (défaut $\mathcal{L}_2$ borné et inconnu) et dans un cas particulier (défaut spécifique). La synthèse est réalisée en considérant les domaines temporels et fréquentiels. Dans le domaine temporel, l'enveloppe inférieure est utilisée afin de régler la rapidité de la réponse alors que l'enveloppe supérieure permet de régler le taux de fausses alarmes.

Une approche active de diagnostic est finalement présentée. Elle consiste à injecter des extra-signaux sur les commandes du système de manière à révéler au mieux la présence de défauts. Les effets des extra-signaux sur les entrées/sorties sont pris en compte tant lors de leur génération que dans la synthèse d'un post-filtre. Deux modèles sont tout d'abord considérés permettant de considérer un fonctionnement normal et anormal du système. Dans le cas de plusieurs défauts, une méthode permettant de les localiser est enfin proposée.

Mots clé: Diagnostic de défauts, Optimisation non lisse, Optimisation multi-objectif, Domaine temporel, Domaine fréquentiel.

Abstract

This thesis considers the application of nonsmooth optimization approach on several FDI problems.

Firstly, to overcome the drawback of classical methods, a nonsmooth optimization approach is proposed to solve a multiobjective fault detection problem in the worst case. An additional constraint of fast transients of residual responses could be added into the design, which could be solved by nonsmooth optimization approach. A framework of designing a unique observer gain and residual weighting matrix is proposed for a system with multiple models. When the exact model is unknown, a new framework of robust fault detection filter and an unchanged threshold are proposed.

Secondly, a method is proposed to design an integrated fault detection observer for general case (unknown $\mathcal{L}_2$ bounded faults and disturbances) and specific case (some specific faults) in frequency and time domain. The lower bound envelope is used to design a fast fault detection observer for the specific faults with a guaranteed ability of fault detection in the worst. By contrast, to decrease false alarms when fault disappearing, a constraint of an upper bound envelope is added into the design.

Thirdly, a new framework of active diagnosis with auxiliary signal is proposed. A criterion of peak amplitude is proposed to evaluate the worst effects from the auxiliary signal on the system. The effects of auxiliary signal on the outputs and control signals are considered in the design. The design is firstly shown with a case of two models, which is then extended to multiple models case.

Keywords: Fault diagnosis, Nonsmooth optimization, Multi-objective optimization, Frequency domain, Time domain.