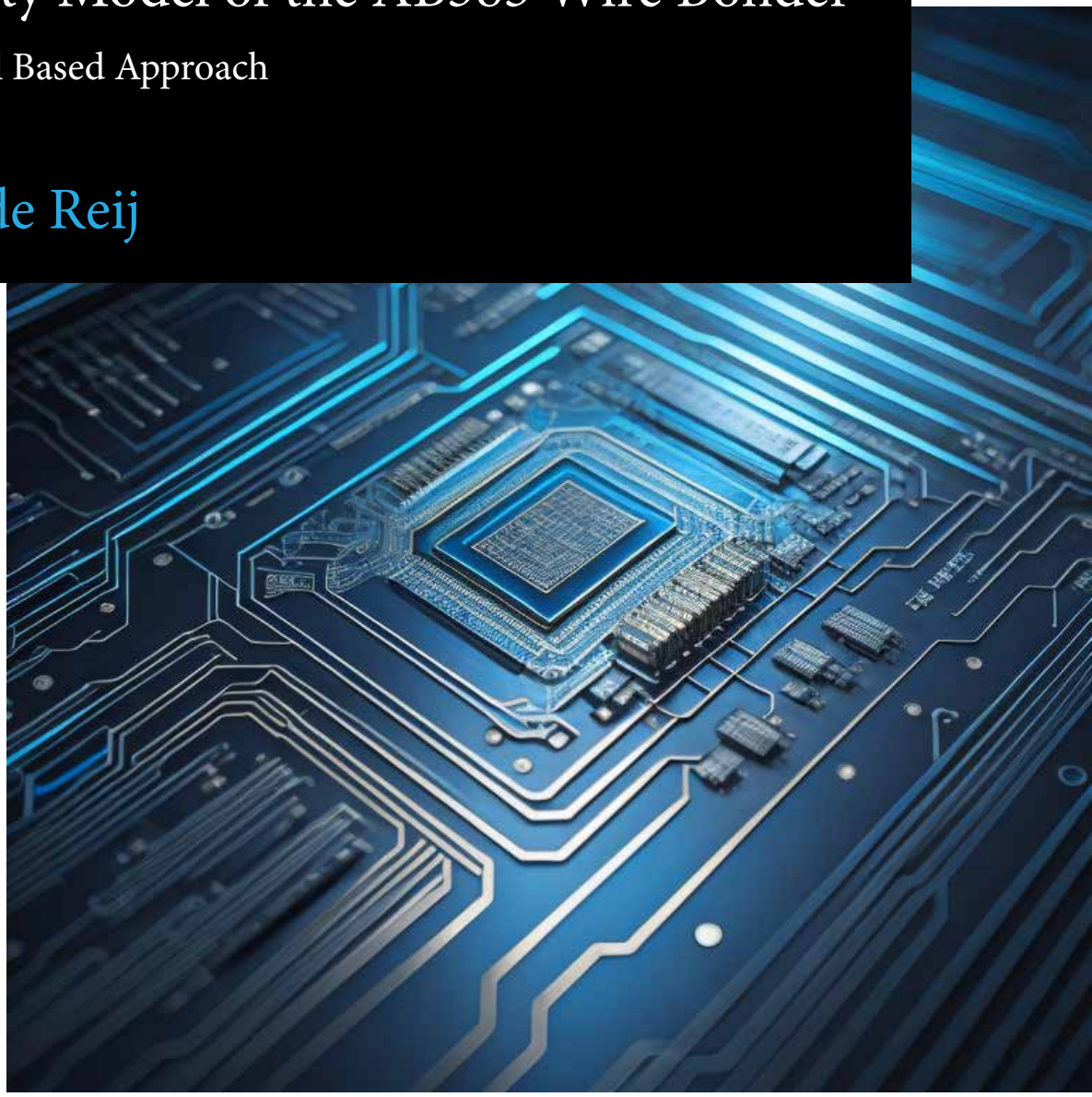
Multivariate Fault Estimation for a High-Fidelity Model of the AB383 Wire Bonder

A Model Based Approach

H.S. de Reij

Master of Science Thesis



TU Delft
Delft Center for Systems and Control

ASMPT
Centre of Competency

TU Delft

Multivariate Fault Estimation for a High-Fidelity Model of the AB383 Wire Bonder

MSc Systems and Control

H. S. de Reij

dr. de Albuquerque Gleizer, G.
ir. Sheikhi, M. A.
dr. Mohajerin Esfahani, P.
Faculty of 3mE
Delft University of Technology

ASMPT

ir. Gerritsen, J.
dr. Kostić, D.
Center of Competency
ASMPT Ltd.

July 20, 2023

Abstract

This research presents the design and implementation of a fault diagnosis filter for a high-fidelity simulation model of the AB383 wire bonder. Fault diagnosis, which consists of detecting, isolating, and estimating faults, enables more effective maintenance strategies and potentially mitigates costly downtime in high-precision motion and positioning systems. When a system deviates from its expected behavior, it can be an indication of the presence of a fault within the system. Faults can occur as either multiplicative faults, arising from deviations in parameters within the system, or additive faults, resulting from external fault signals that impact the system's operation. This study uses a model-based methodology that generates residuals that are subsequently analyzed using a regression method aimed at determining the influence of each fault in the residual signal, thereby facilitating fault estimation. A residual signal represents the difference between the actual system behavior and the expected behavior, mainly serving as an indicator of potential faults within the system. A data-driven threshold design is proposed to determine the detectability of faults. The main contributions of this work include the development of a fault modeling framework for the residual generation method and the application of the fault estimation framework to a linear high-fidelity simulation model affected by both multiplicative and additive faults. The simulation results demonstrate satisfactory performance, that is, accurately detecting and estimating faults. Moreover, the proposed method effectively identifies external disturbances and high levels of noise through power spectral density analysis. The findings highlight the potential of this approach, outperforming alternative methods in terms of accuracy when it comes to fault diagnosis for wire bonder machines. The research findings contribute to the active field of fault detection and estimation for complex systems, offering valuable insights for further studies and potential practical applications.

Keywords **-** Fault Diagnosis, Fault Estimation, Data-Driven Threshold Design, Model-Based Fault Diagnosis, Multiplicative Faults, Additive Faults, Regression

Acknowledgements

I wish to express my deepest gratitude to my daily supervisors, Gabriel Albuquerque Gleizer and Amin Sheikhi, for their invaluable guidance, insight, and unwavering support throughout the progression of my thesis. Their approachable, non-judgmental, and welcoming attitude transformed our meetings into sessions I looked forward to, creating an overall pleasant and intellectually enriching experience. Through their support, I not only enhanced my knowledge about the professional field but also developed a profound understanding of what a truly supportive work environment entails.

I am deeply grateful to Peyman Mohajerin Esfahani, who, despite his demanding schedule, has provided support whenever needed. His results-driven approach has been a source of motivation and inspiration throughout my research journey. His remarkable contributions to the field of fault diagnosis have been instrumental in shaping my research and enhancing my understanding of the discipline.

Additionally, my gratitude extends to Jasper Gerritsen, my supervisor from ASMPT Ltd., for his industry-focused insights, guidance, and critical eye for detail. His practical guidance has assisted me in my understanding of the system of interest and the model he created was essential for this thesis. Special thanks are extended to Dragan Kostić from ASMPT Ltd. for presenting me with the unique opportunity to carry out my thesis in a field that is not in his team's immediate focus.

In conclusion, I wish to convey my appreciation to everyone who has contributed to this project, either directly or indirectly. Your collective input and efforts have made this project possible. I am sincerely grateful for your time, assistance, and for always being there when I needed advice or support.

Contents

Acronyms

I/O Input/Output

DoF Degree of Freedom

DoFs Degrees of Freedom

EoM Equation of Motion

LTI Linear Time-Invariant

MSD Mass-Spring-Damper

DAE Differential-Algebraic Equation

SSM State Space Model

ODE Ordinary Differential Equation

FF Feedforward

FB Feedback

n Nominal

f Faulty

LP Linear Programming

OLS Ordinary Least Squares

PCA Principal Component Analysis

PLS Partial Least Squares

UIO Unknown Input Observer

PSD Power Spectral Density

List of Figures

List of Tables

1 Introduction

This introductory chapter provides an overview of the thesis, which focuses on the development, implementation, and evaluation of a fault diagnosis system for a high-fidelity simulation model of the AB383 wire bonder. The chapter begins by discussing the background and motivation for this research in Section 1.1, highlighting the importance of timely and accurate fault diagnosis in reducing downtime and improving efficiency. Section 1.2 provides a brief overview of fault diagnosis methods for complex systems discussed in the literature. The main research problems to be addressed in the thesis are outlined in Section 1.3. The specific challenges to high-precision motion and positioning systems, such as wire bonding machines, are discussed in Section 1.4, setting the context for the main contributions of this research in Section 1.5. Finally, the organization of the thesis is presented in Section 1.6, providing a roadmap for the subsequent chapters.

1.1 Background and Motivation

The manufacturing industry heavily relies on high-tech systems that operate within high precision and repeatability requirements. Any instance of unplanned downtime in these systems can lead to substantial costs, primarily resulting from factors such as a reduction in product yield or missed delivery deadlines. Hence, it is crucial to minimize such downtime as much as possible. Therefore, timely and accurate fault diagnosis can play an important role in reducing downtime, improving efficiency, and increasing the overall reliability of manufacturing systems.

In recent decades, a wide range of fault diagnosis methods and techniques have been developed and utilized in various fields. A comprehensive literature review on fault diagnosis for high-precision motion and positioning systems can be found in [8]. Fault diagnosis methods have advanced significantly in response to the mounting complexity of modern systems, which often operate across multiple physical domains simultaneously.

ASMPT Ltd. is one of the companies interested in fault diagnosis for their complex motion systems. ASMPT Ltd. specializes in providing hardware and software solutions used in the production of semiconductors and electronics [1]. This thesis focuses on a fault diagnosis system that has been developed and implemented on a high-fidelity simulation model of the AB383 wire bonder, a machine designed and fabricated by ASMPT Ltd. Wire bonders are used in the back-end of the electronics industry, as they create connections between integrated circuits and their packaging [16]. The back-end process in semiconductor manufacturing transforms individual dice into packaged devices ready for integration into electronic systems.

The AB383 wirebonder operates with a high level of precision and accuracy, maintaining submicron accuracy at accelerations exceeding 150 Gs. In these types of systems, several faults can potentially degrade the system's performance. A lack of lubrication can increase friction in one of the motion guides. Over time, the permanent magnets in the actuators can demagnetize. The fasteners, connecting the different components rigidly together, can untighten, leading to a reduction in structural stiffness. External disturbances can be present because the environment of the machine might not properly absorb factory floor vibrations.

Detecting and isolating these faults is challenging because they may only cause minor deviations from the desired trajectory in the measurement data, yet still result in performance that falls outside the necessary specifications. Figure 1 illustrates the comparison between the nominal system and the system with a 20% friction increase in the x-motion guide of the machine, executed during a 4mm closed-loop motion. Despite the presence of the fault, the trajectories remain nearly identical; similar findings for the other possible faults in this research can be found in Appendix A.1. This highlights the challenges in developing a precise fault diagnosis system for the high-precision wire bonder, where even minor deviations can result in performance outside the precision requirements.

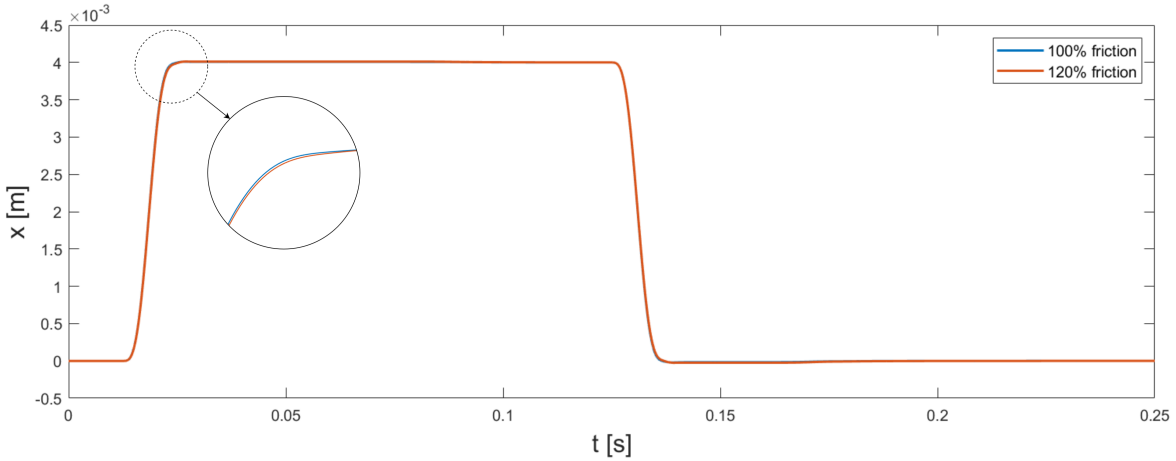


Figure 1: A comparison between the measured x position of the x-motion stage for the nominal system and the same system with increased viscous friction.

1.2 Review of Fault Diagnosis in Literature

This section provides a discussion of some of the most frequently used and well-known fault estimation methodologies for motion systems.

Quantitative knowledge-based fault estimation methods make use of numerical data and

measurements to analyze system behavior [12]. Techniques employed within this framework include machine learning methods like neural networks, which have been successfully applied in estimating actuator faults in wind turbine simulations [20]. Moreover, statistical approaches such as Principal Component Analysis (PCA) have been effectively used to estimate sensor faults in centrifugal chillers [27]. These methods are advantageous when diagnosing faults in complex systems with large volumes of available data. However, they require substantial computational resources and are sensitive to the quality and quantity of data used. Their performance may also deteriorate when maintenance or changes are made to the system [12].

Observer-based fault estimation methodologies, on the other hand, augment the state vector by integrating the faults of interest as additional states. This procedure allows the estimation of the state vector, which in turn, leads to the estimation of both the fault signals and the original system states [13]. A study in [21] demonstrated the effective application of an Unknown Input Observer (UIO) in detecting and estimating unknown actuator faults. Furthermore, the research illustrated in [28] deployed a combination of integral observers, sliding observers, and adaptive observers for the reconstruction of sensor faults in satellite control systems.

Parameter estimation techniques employ statistical methods to approximate unknown parameters within a model or system based on observed data. In [26], various faults (sensor, actuator, and process) were individually implemented in a wind turbine case study and successfully estimated. Despite their effectiveness in fault diagnosis and fault identification, these techniques might require an extensive amount of data and can be sensitive to noise and disturbances within the data [5].

1.3 Research Problems

The main research problem that needs to be addressed is designing a fault estimation system for the high-fidelity simulation model of the AB383 wire bonder. This system is a real-time algorithm that should process measured signals and convert these into an estimate of a range of possible faults related to the actuators, system, and disturbances that might be present in the system simultaneously.

The accurate modeling of faults, such as structural stiffness, friction, and motor faults, is of importance in addressing the main research problem. These faults should be incorporated into the mathematical model of the system, enabling parameter adjustments for performance verification.

1.4 Challenges

Considering the implementation of fault diagnosis methods highlighted in the literature 1.2, a key challenge is the broad spectrum of potential faults, including component failures, actuator faults, and external disturbances, which further complicates the diagnosis process. Effectively managing this diverse range of faults presents a significant challenge that needs to be addressed.

While this work primarily focuses on implementing the fault diagnosis system on a high-fidelity model, it is crucial to consider its potential applicability to a physical machine. This implies that the proposed fault diagnosis system must exhibit a strong balance between detection sensitivity and robustness to disturbances and potential model mismatch.

By addressing these research problems and challenges, the current work aims to develop fault diagnosis techniques for positioning and motion systems, with the ultimate goal of enabling potential implementation on practical applications.

1.5 Main Contributions

This thesis provides several contributions to the field of fault diagnosis for complex systems, specifically concentrating on motion and positioning systems. The insights and methodologies proposed directly address the challenges identified in Section 1.4.

The primary contribution of this thesis is the implementation of a model-based fault estimation framework. This framework, designed for a high-fidelity simulation model of a physical system, takes into account the distinct characteristics and requirements of motion and positioning systems. This proposed method effectively confronts the challenges associated with high precision requirements, intricate system dynamics, and a wide array of fault types. Therefore, it creates a promising foundation for an accurate fault diagnosis system for the AB383 wire bonder.

Furthermore, a notable contribution of this work is the introduction of a novel fault representation and modeling approach within the residual generation method. This approach makes an analysis of the Equations of Motion (EoM) to obtain the necessary parameters within the system matrices which are related to the individual faults in the diagnosis framework. This also includes mathematical transformations that effectively convert parametric faults into external fault signals.

A novel contribution of this research is the introduction of a technique for identifying

and quantifying external disturbances through residual analysis. This technique utilizes power spectral density analysis to determine elevated noise levels or specific frequency disturbances impacting the system. This method is complemented by separating low and high-frequency components, aiding in the real-time estimation of faults and high-frequency disturbances simultaneously.

Lastly, the proposed fault diagnosis methods are validated through comprehensive simulations and experiments using the simulation model of the AB383 wire bonder. The performance of the developed techniques is assessed based on their accuracy, effectively demonstrating the effectiveness of the proposed methods in addressing the challenges and requirements of fault diagnosis for these types of systems.

1.6 Thesis Organization

This thesis is structured into six distinct chapters, each contributing to the progression of our research narrative.

Chapter 2 offers an in-depth introduction to our system of interest, the AB383 wire bonder. The chapter includes a three-dimensional schematic representation of the wire bonder. It also encompasses the linear and nonlinear mathematical representation of this system, drawing upon the mass-spring-damper modeling framework [14]. The potential faults and their implementation in the system are also discussed.

In Chapter 3, the model-based fault diagnosis method is elaborated. This includes the necessary modeling considerations, the residual generation method based on the work in [17], threshold design, the fault estimation method based on the work in [7], and external disturbances estimation.

Chapter 4 is dedicated to presenting the design and implementation details of the research. It encompasses a thorough discussion of the fault detection and estimation architectures, along with the specific configurations employed in simulations. Additionally, the chapter provides a brief explanation of the controller utilized in this study.

In Chapter 5, the research findings are highlighted. The chapter showcases the detectability of various fault types within the system, aided by the designed thresholds. It also outlines the findings from fault estimations and external disturbances estimation. These results are examined and compared to previous work.

Chapter 6 concludes the thesis with a comprehensive discussion of the overall findings,

limitations, and potential directions for future research. The chapter draws conclusions from the body of work and indicates how the study might inform or inspire subsequent research in this field.

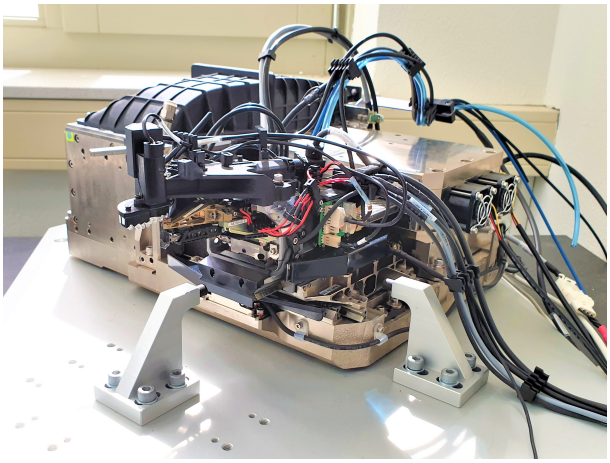
2 System Overview and Analysis

This chapter provides a detailed examination of the system under study, presented in Section 2.1. It features a schematic diagram of the system alongside the linear and nonlinear EoM, which are discussed in Section 2.2. Furthermore, the chapter outlines a fault modeling framework in Section 2.3 accommodating the implementation of specific faults that are considered throughout this research.

2.1 The AB383 Wire Bonder

This study focuses on high-fidelity simulation models of the AB383 wire bonder, a machine designed and manufactured by ASMPT Ltd. [1]. A wire bonder is a precision machine utilized in the back-end process of the electronics and semiconductor industry, creating precise electrical connections with thin wires between integrated circuits (ICs) and their packaging. Figure 2 presents an image of the AB383 wire bonder, including its isolated motion system. These systems are available for experimentation and are located at the Center of Competency in Beuningen.

(a) AB383 wire bonder [9].



(b) Isolated motion system [2].

Figure 2: AB383 wire bonder and the isolated motion system.

2.2 System Modelling

In this section, the available models of the AB383 wire bonder are presented. This includes a schematic representation of the system dynamics, the nonlinear and linear EoM.

2.2.1 Schematic Representation of the System

Figure 3 provides a schematic diagram of the AB383 wire bonder that incorporates both translational and rotational Degrees of Freedom (DoFs) of the motion system, including nominal and parasitic DoFs, as well as prismatic and revolute joints. Nominal DoFs refer to the intended motion paths or directions of the system, such as linear or rotational movements facilitated by prismatic and revolute joints, respectively. In contrast, parasitic DoFs represent unintended motion paths that emerge from the impossibility of achieving perfect rigidity in the connections between subsystems. Table 1 presents the various DoFs and their respective roles in the schematic illustration of the wire bonder shown in Figure 3.

The machine's base frame is connected to the ground through multiple spring-damper connections on roller joints. The base frame serves as the foundation for both the x-motion stage and the y-motion stage. These stages are responsible for movement along the x and y directions, respectively, and their motion is controlled by actuators situated on the base frame. To simulate friction, each employs a damper on a roller joint incorporated into their models. Moreover, to counteract any undesired orthogonal motion (parasitic DoF), they both utilize a combination of a spring and a damper on a roller joint.

In order to monitor the positions of both the x-motion stage and the y-motion stage, sensors are employed. The sensor for the x-motion stage measures the distance between the base frame and itself, while the sensor for the y-motion stage measures the distance between the x-motion stage and the y-motion stage. The z motion is controlled by a rotational force and measured by a rotational sensor. However, to restrict the range of motion, a torsion spring and damper are implemented. The nature of these connections also allows for rotations in the various components.

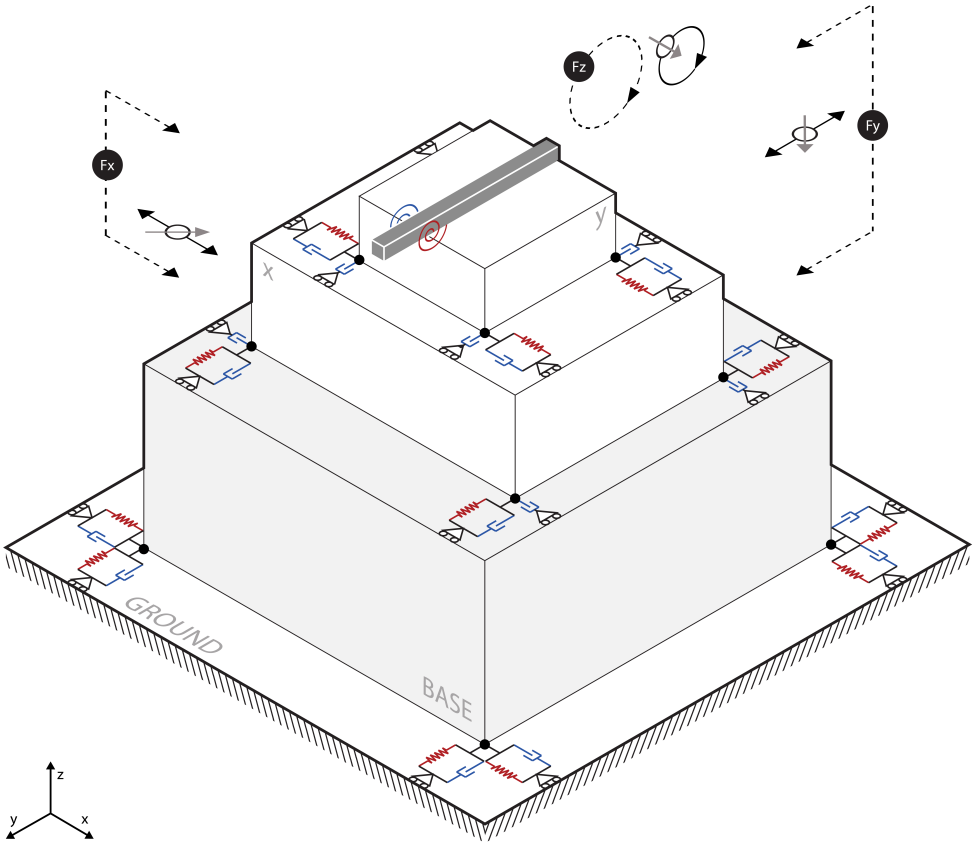


Figure 3: A three-dimensional diagram illustrating the interconnections among various system components that are used in the derivation of the mathematical model.

Table 1: Description of DoFs of the wire bonder model [14].

DoF	Description
u_1, u_2	Parasitic translational DoFs between fixed world and chassis + base + x-guide
r_1	Parasitic rotational DoF between fixed world and chassis + base + x-guide
X	Nominal translational DoF between x-guide and x-stage
u_3	Parasitic translational DoF between x-guide and x-stage
r_2	Parasitic rotational DoF between x-guide and x-stage
u_4	Parasitic translational DoF between y-guide and y-stage
Y	Nominal translational DoF between y-guide and y-stage
r_3	Parasitic rotational DoF between y-guide and y-stage
Z	Nominal rotational DoF between y-stage and z-stage

2.2.2 Nonlinear Equations of Motion

A detailed analysis of the nonlinear EoM, provided by ASMPT [14], is outside of the scope of this research, but the model used, as illustrated in Eq. 1, incorporates various aspects of the dynamics of the system. The term q represents the generalized coordinates,

which describe the spatial positions of the joints within the system as illustrated in the DoFs listed in Table 1. These coordinates' respective derivatives, denoted by $\dot{q}$, represent the velocity of these DoFs, while $\ddot{q}$ symbolizes the acceleration. The inertia matrix also referred to as the kinetic energy matrix, is represented as $M(q) \in \mathbb{R}^{n_q \times n_q}$. This term captures how changes in the joint configuration influence the kinetic energy of the system. The Coriolis and centrifugal forces are modeled in $V(q, \dot{q}) \in \mathbb{R}^{n_q}$, which depend on the joints' velocity and position. $G(q) \in \mathbb{R}^{n_q}$ refers to the gravity force vector, or more accurately, the potential energy gradient. This force, oriented downward, is dependent on the spatial configuration of the joints. Finally, $\Gamma \in \mathbb{R}^{n_q \times n_u}$ represents the generalized forces which may include actuator forces or other externally exerted forces on the system. The variable u denotes the input to the system.

$$M(q)\ddot{q} + V(q, \dot{q}) + G(q) = \Gamma u \tag{1}$$

2.2.3 Linearized Equations of Motion

The linear EoM is obtained by linearizing the nonlinear EoM in Eq. 1. The rotational parasitic DoFs, as shown in Table 1, are linearized around their zero-valued equilibrium point, the static friction in the motion guides is neglected and the nominal rotational DoF, Z, is around also linearized around its equilibrium. This linearized model is restructured as seen in Eq. 2 and Eq. 3, which is based on the work in [2]. Distinguishing the model used in this research from previous work is the integration of all motion stages, instead of decoupled systems [14].

The linear EoM are organized as shown in Eq. 2 [2]. This equation introduces a damping $D \in \mathbb{R}^{n_q \times n_q}$ and stiffness matrix $K \in \mathbb{R}^{n_q \times n_q}$.

$$M\ddot{q} + D\dot{q} + Kq = \Gamma u \tag{2}$$

The measured position, denoted by p, is described by Eq. 3, where the matrix $b \in \mathbb{R}^{n_q \times n_y}$ characterizes the relationship between the generalized coordinates q and the measured position p.

$$p = b^\top q \tag{3}$$

The matrices corresponding to this MSD framework can be found in Appendix A.2.

2.3 Fault Scenario Analysis

This section addresses potential faults that may compromise the performance and precision of motion and positioning systems, e.g. the AB383 wire bonder. Methodologies for modeling individual faults associated with motion guides, fasteners, actuators, and external disturbances are presented.

2.3.1 Motion Guides

Motion guides are used in motion systems to provide precise and controlled movement along a specific path. However, increased friction within these guides can cause faulty system behavior. This can be the result of, e.g. insufficient lubrication in the motion guides. The friction in this work is modeled with both static and viscous friction. However, in this study, only deviations in the viscous friction for both x- and y-motion stages are considered.

Viscous friction in prismatic and revolute joints can be modeled with the generalized damping coefficient associated with a corresponding motion guide in the damping matrix D in Eq. 2. The fault corresponding to the viscous friction can be modeled by adjusting the corresponding damping parameters of the nominal translational DoFs shown in Table 1. Specifically, the derivation of the faulty damping parameter d_i^f is shown in Eq. 4, where the nominal damping parameter d_i^n is related through the coefficient ζ_i. This coefficient ranges from 1.00 to a maximum value of 1.20, allowing for a controlled increase in friction within the motion guides for simulation. By incorporating these tuned damping parameters into the damping matrix D of Eq. 2, the model captures the effects of increased friction and enables the analysis of system behavior under these conditions.

$$d_i^f = \zeta_i \, d_i^n \quad \text{for} \quad \zeta_i \in [1.00, 1.20] \quad \forall \, i \in \{X, Y\} \tag{4}$$

2.3.2 Structural Stiffness

Fasteners, which include bolts, nuts, and other mechanical connections, are critical in sustaining the structural stiffness of the system. Over time, these mechanical components can loosen, resulting in a decrease in stiffness. To model faults related to fasteners connecting the motion stages with the other system components, the stiffness coefficients within the stiffness matrix K in Eq. 2 can be adjusted. Specifically, a decrease in structural stiffness in a specific motion stage can be represented by modifying the generalized stiffness coefficients of the parasitic DoFs, both rotational and translational, that connect the corresponding motion stage subsystem to the other subsystems.

For the system of interest, the decrease in structural stiffness related to the x- and y-motion stages can be modeled by deriving faulty stiffness parameters k_i^f and k_j^f as shown in Eq. 5. In this equation, k_i^n and k_j^n represent the nominal stiffness parameters of the parasitic DoFs linked to both the x- and y-motion stages, as outlined in Table 1. The coefficients β_x and β_y range between 0.95 and 1.00, enabling effective adjustment of the

structural stiffness of the corresponding motion stage.

$$\begin{aligned} k_i^f &= \beta_x \, k_i^n \quad \text{for } \beta_x \in [0.95, 1.00] \quad \forall \, i \in \{u_3, r_2\} \\ k_j^f &= \beta_y \, k_j^n \quad \text{for } \beta_y \in [0.95, 1.00] \quad \forall \, j \in \{u_4, r_3\} \end{aligned} \tag{5}$$

2.3.3 Actuators

The system's actuators generate the necessary force to drive both rotational and translational motion, thereby controlling the motion stages. However, these actuators can experience performance degradation over time due to reasons such as the demagnetization of the permanent magnets. This type of fault can be modeled by multiplying a column in Γ, which is specific to the actuator under consideration, with a fault coefficient.

To address a decrease in the available motor force, adjustments can be made to the motor force constants that describe the relationship between the input current and the force generated to move the individual motion stages. The actuator fault can be implemented using Eq. 6, where K_i^f denotes the faulty motor force constant for actuator i, K_i^n is the nominal motor force constant for the respective actuator, and η_i provides the means to tune the available motor force between 0.97 and 1.00.

$$K_i^f = \eta_i \, K_i^n \quad \text{for } \eta_i \in [0.97, 1.00] \; \forall \, i \in \{X, Y, Z\} \tag{6}$$

2.3.4 External Disturbances

The performance of the system can be influenced by external disturbances present in the machine environment. These disturbances can be classified as either unknown or partially known, where certain prior knowledge is available. In the system context, these disturbances can be treated as external signals that exhibit similar dynamics (Γ) as the input to the system.

In the context of this research, there is prior knowledge regarding the frequencies of the external disturbances signal acting on the system. This signal is represented by a multisine input signal with frequencies of 100, 300, and 750 Hz. The mathematical representation of these disturbances is presented in Eq. 7 where the amplitude A is calculated as a proportion r of the maximum force F_{max} that can be exerted on the motion stages by the corresponding actuator. For the x and y-motion stages, the percentage r is set to 1.00%, while for the θ-subsystem, it is set to 0.10%. The variable γ represents the magnitude of the external disturbance signal and it has to be identified as a fault.

$$F_{dist}(t) = \gamma \, A \left(\sin\left(2\,\pi 100 t\right) + \sin\left(2\pi 300\, t\right) + \sin\left(2\pi 750\, t\right) \right) \quad \text{where } \; A = r \, |F_{max}| \tag{7}$$

3 Method

This chapter presents the methodology used in this study, consisting of specific modeling considerations in Section 3.1, the residual generation method for fault diagnosis in Section 3.2, and the estimation of faults, and external disturbances in Section 3.3.

3.1 Modeling Methodology

The residual generation method, discussed in Section 3.2, necessitates particular modeling considerations, which are discussed in this section. These considerations require the transformation of various mathematical models such as a linear MSD equation into a state space model (SSM), which is then further translated into a differential-algebraic equation (DAE). Moreover, a method is introduced that formulates a modeling technique addressing faults and potential disturbances, essential in the deployment of the residual generation filter.

3.1.1 State Space Model

The linear SSM in continuous time is formulated as shown in Eq. 8. The state vector of the SSM can be expressed as $x = [\, q, \dot{q} \,]^T$, where the DoFs presented in Table 1 are employed as the generalized coordinates q and their respective derivatives $\dot{q}$.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_u u(t) \\ y(t) &= Cx(t) + D_u u(t) \end{aligned} \tag{8}$$

The system matrix $A \in \mathbb{R}^{n_x \times n_x}$, the input matrix $B_u \in \mathbb{R}^{n_x \times n_u}$, the output matrix $C \in \mathbb{R}^{n_y \times n_x}$, and the feedthrough matrix $D_u \in \mathbb{R}^{n_y \times n_u}$, associated with a mass-spring-damper model, are defined in Eq. 9. The corresponding MSD matrices are detailed in Section 2.2.3. For clarity, the matrices in the MSD framework are as follows: the stiffness matrix $K \in \mathbb{R}^{n_q \times n_q}$, the damping matrix $D \in \mathbb{R}^{n_q \times n_q}$, the relation between the generalized coordinates and the output matrix $b \in \mathbb{R}^{n_q \times n_y}$, and the generalized forces matrix $\Gamma \in \mathbb{R}^{n_q \times n_u}$.

$$\begin{aligned} A &= \begin{bmatrix} \mathbf{0}_{n_q \times n_q} & I_{n_q \times n_q} \\ -M^{-1}K & -M^{-1}D \end{bmatrix}, \quad B_u = \begin{bmatrix} \mathbf{0}_{n_q \times n_u} \\ M^{-1}\Gamma \end{bmatrix}, \\ C &= \begin{bmatrix} b^T & \mathbf{0}_{n_y \times n_q} \end{bmatrix}, \quad D_u = \begin{bmatrix} \mathbf{0}_{n_y \times n_u} \end{bmatrix}. \end{aligned} \tag{9}$$

3.1.2 Effects of Faults and Disturbances

To consider the potential effects of external disturbances $d(t)$ and fault signals $f(t)$, the SSM is extended as shown in Eq. 10. The extended SSM involves the introduction of an input disturbance matrix $B_d \in \mathbb{R}^{n_x \times n_d}$, a fault input matrix $B_f \in \mathbb{R}^{n_x \times n_f}$, a disturbance

feed-through matrix $D_d \in \mathbb{R}^{n_y \times n_d}$, and a fault feed-through matrix $D_f \in \mathbb{R}^{n_y \times n_f}$. However, the matrices B_d and D_d are excluded from the modeling framework utilized in this study. Additionally, the system's lack of fault feed-through results in the feed-through matrix being represented as $D_f = \mathbf{0}_{n_y \times n_f}$.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_u u(t) + B_d d(t) + B_f f(t) \\ y(t) &= Cx(t) + D_u u(t) + D_d d(t) + D_f f(t) \end{aligned} \tag{10}$$

3.1.3 Modelling of B_f

The design of the fault input matrix B_f depends on the type of fault that needs to be detected. The types of faults considered are external fault signals acting on the system and parametric faults in the system matrices.

External Faults
The modeling of the dynamics of external fault signals acting on the system relies on prior knowledge of these specific signals. In this study, for simplicity, it is assumed that the disturbances have similar dynamics to the input and affect all input channels simultaneously. Therefore, the matrix B_u is used as the B_f matrix to model the external disturbance fault.

Parametric Faults
Parametric faults cannot be directly represented as external signals because their effect is contained within the system matrices. These faults can be associated with actuators, motion guides, or components responsible for preserving structural stiffness.

For actuator faults, the assumption is made that a fault in actuator i affects the system similarly as the corresponding column in matrix B_{u_i}. Resulting in matrix $B_f = B_{u_i}$ as the vector used to model the independent actuator faults.

The faults related to structural stiffness and viscous friction in the motion guides, as outlined in Section 2.3, are simulated by adjusting specific generalized damping coefficients d_i^f and stiffness coefficients k_i^f.

To accurately represent the structural stiffness of the x- and y-motion stages, we modify the generalized stiffness coefficient of the rotational and orthogonal DoFs corresponding to the respective motion stage. In Table 1, these DoFs are denoted by u_3 and r_2 for the structural stiffness fault in the x-motion stage, and by u_4 and r_3 for the structural

stiffness fault in the y-motion stage.

The faults related to viscous friction in the x- and y-motion guides are modeled by adjusting the generalized damping coefficients of the nominal DoFs related to the respective motion stage. As indicated in Table 1, this correlates to DoF X for the viscous friction fault in the x-motion guide, and DoF Y for the viscous friction fault in the y-motion guide.

Each of the aforementioned DoFs also corresponds to specific states, which are presumed to be most affected by the respective faults. This relationship is demonstrated in Eq. 11, where the vector B_f is modeled with a nonzero value for the states impacted by the fault, denoted by i_{DoF_f}, and zero otherwise.

$$B_{f,i} = \begin{cases} 1 & \text{if } i = i_{DoF_f} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Example 1. Parametric Fault Transformation.
Consider a system that incorporates a parametric fault Δ_f as exemplified in the following equation:

$$\dot{x} = (A + \Delta_f A)x = \begin{bmatrix} A_{11} & A_{12} \\ \Delta_f A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

This system can be restructured into a state-dependent fault input vector $B_f(x)$:

$$\dot{x} = Ax + B_f(x) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ (\Delta_f - 1)x_1 \end{bmatrix}.$$

The state-dependent $B_f(x)$ vector can be transformed, with $f = (\Delta_f - 1)x_1$, into a static $B_f = [0, 1]^T$:

$$\dot{x} = Ax + B_f f(x) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (\Delta_f - 1)x_1.$$

The validity of this model relies on x_1 being nonzero, which is a necessary prerequisite. Within the context of dynamical motion systems, states linked to positioning coordinates are conventionally presumed nonzero, being integral components of system dynamics.

3.1.4 Linear Differential-Algebraic Equations

In this study, the residual generation method is applied for DAEs rather than focusing on more specific model classes such as transfer functions, state-space models, and descriptor models. As DAE models encompass all these model classes, the methods proposed in this study are applicable to each of the aforementioned models [11].

The general framework for a DAE for fault diagnosis is shown in Eq. 12 [17]. Signal x captures all unknown signals, including internal states and unknown external inputs (e.g. external disturbances), z contains all known signals, such as control signals and output measurements, and f represents the fault signals. The signals x, z, and f are assumed piece-wise continuous functions. Matrices $H(p)$, $L(p)$, and $F(p)$ represent polynomial matrices where p can represent either a differential operator, a complex variable, distributional derivative operator, a normal derivative operator, or a discrete-time difference operator.

$$H(p)x + L(p)z + F(p)f = 0 \tag{12}$$

Example 2. Generality of the DAE framework.
To demonstrate the generality of the DAE framework, consider the mass-spring-damper system in Figure 4, a second-order system, which can be represented using different models.

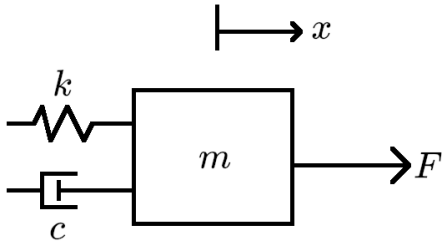


Figure 4: A second-order mass-spring-damper system [23].

- ***Transfer Function***
  The transfer function of this system is given by:

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2 + cs + k}$$

  where m is the mass, c is the damping coefficient, k is the spring constant, $X(s)$ is the Laplace transform of the output $x(t)$, and $F(s)$ is the Laplace transform of the input $F(t)$.

- ***State-Space model***
  The state-space representation of this system can be written as:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix} x + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} u$$
$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} x + 0\,u$$

  where $x = [x_1, x_2]^T$ are the state variables and the force F is the input u.

- ***Descriptor model***
  The descriptor model of the system can be given by:

$$E\dot{x} = Ax + Bu$$
$$y = Cx$$

where

$$E = \begin{bmatrix} 1 & 0 \\ 0 & m \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ -k & -c \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Differential Algebraic Equation Model

Transforming the state-space or descriptor model, the system can be expressed as a DAE by including an algebraic equation that correlates the states with the outputs. If a sensor is implemented to directly measure the position x_1, the system representation becomes:

$$E\dot{x} = Ax + Bu,$$
$$0 = y - Cx.$$

In this DAE formulation, the dynamic behavior of the system is encapsulated by the first equation (a differential equation), while the algebraic constraint, derived from the system or sensor dynamics, is given by the second equation. The use of DAE models provides a general approach to representing systems by accommodating a broader spectrum of dynamics and constraints compared to the aforementioned models.

Assumption 1. Observability.

Following the work outlined in [17], a condition is introduced for the system model matrices, as detailed in Eq. 12. This condition necessitates the absence of linear dependencies when the fault parameter f is zero, thus ensuring the observability of the system state.

This observability prerequisite is fulfilled when the augmented matrix $\begin{bmatrix} H(s) & L(s) \end{bmatrix}$ has full row rank, signifying the maximum number of linearly independent rows. This condition enables the isolation of faults due to the unique definition of all possible observations within the DAE framework if there is no fault present, as described in Eq. 12.

3.1.5 Transform a State Space Model into a Differential-Algebraic Equation

In the context of the method introduced in [17], it is necessary to transform the extended SSM into a DAE with matrices $H(s) \in \mathbb{C}^{(n_x+n_y)\times(n_x+n_d)}$, $L \in \mathbb{R}^{(n_x+n_y)\times(n_y+n_u)}$, and $F \in \mathbb{R}^{(n_x+n_y)\times n_f}$, as depicted in Eq. 13.

$$\underbrace{\begin{bmatrix} -sI + A & B_d \\ C & D_d \end{bmatrix}}_{H(s)} \underbrace{\begin{bmatrix} X \\ d \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} 0 & B_u \\ -I & D_u \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} Y \\ u \end{bmatrix}}_{z} + \underbrace{\begin{bmatrix} B_f \\ D_f \end{bmatrix}}_{F} f = 0 \tag{13}$$

3.2 Residual Generation Filter

The aim of this section is to provide a detailed description of the residual generation and its mathematical foundation. In Section 3.2.1, the theoretical background of the residual generation approach is presented. Followed by Section 3.2.2 which describes the translation of the mathematical theorem into a linear programming (LP) framework. To give more practical insights, an example is given to intuitively understand the mechanism of the residual generation method. Section 3.2.3 discusses the method to establish the thresholds for the residual signals to determine fault detectability.

3.2.1 Theoretical Background

In [17], a linear model is formulated as shown in Eq. 12 where the nominal behavioral set $\mathcal{M}$ is determined as described in Eq. 14. This mathematical representation forms the foundation of the residual generation filter. $N_H(s)$ consists of rows that form an irreducible polynomial basis for the left null space of the matrix $H(s)$, L consists of the input and output signal dynamics as shown in Eq. 13 and signal z consists of the known signals. The set $\mathcal{W}^{n_z}$ is described by the piece-wise continuous function z from $\mathbb{R}_+$ into $\mathbb{R}_{n_z}$ with $n_z = n_u + n_y$.

$$\mathcal{M} = \{z \in \mathcal{W}^{n_z} \mid N_H(s)L\,z = 0\} \tag{14}$$

The transfer function of the residual generation filter, $R(s)$, is shown in Eq. 15. $N(s)$ is chosen as a linear combination of the rows of $N_H(s)$ together with stable dynamics $d(s)$ which is a denominator of sufficient order to ensure the residual filter is both stable and realizable.

$$R(s) = d^{-1}(s)N(s)L. \tag{15}$$

Remark 1. Derivation of $d(s)$.
An LTI system maintains its frequency response characteristics regardless of any variations or modifications in its internal components or parameters. Considering that the system under consideration is an LTI system, the derivation of the denominator $d(s)$ of the residual generation filter, as indicated in Eq. 15, can be determined by evaluating the highest frequencies present in the input signals during typical trajectories. To ensure the observability of faults, $d(s)$ can be designed as a low-pass filter (e.g., a Butterworth filter) of sufficient order and with a cutoff frequency set adequately high to avoid suppressing any fault frequencies.

In order to obtain the residual generation filter, it is necessary to find a suitable $N(s)$ that will be used in generating a residual signal, as shown in Eq. 15. This signal should be insensitive to nominal dynamics, indicating that the residual signal remains unaffected when the system operates within the nominal system behavior $\mathcal{M}$ as defined in Eq. 14, as

shown in Eq. 16a. However, the filter should be sensitive to faults that cause the system to deviate from the nominal behavioral set, as shown in Eq. 16b.

$$N(s)H(s) = 0 \tag{16a}$$

$$N(s)F(s) \neq 0 \tag{16b}$$

Remark 2. Fault Detectibility.
In accordance with the methodology in [10], it is argued that a viable solution $N(s)$ for Eq. 16 can be obtained if and only if:

$$\text{Rank}\left(\begin{bmatrix} H(s) & F(s) \end{bmatrix}\right) > \text{Rank}\left(\begin{bmatrix} H(s) \end{bmatrix}\right)$$

This condition, where $F(s)$ can consist of the dynamics of a single fault or multiple faults, establishes a higher rank for the augmented matrix as compared to $H(s)$.

Example 3. Insight into the Residual Generation Method for Fault Detection.
The general method for residual generation can be represented by the architecture shown in Figure 5. There are a group of signals acting on the LTI system, that is, an external fault signal, an unknown input, and a third signal which corresponds to the input signal in an open-loop system or the reference signal in a closed-loop system. The input and output signals are used to generate a residual signal.

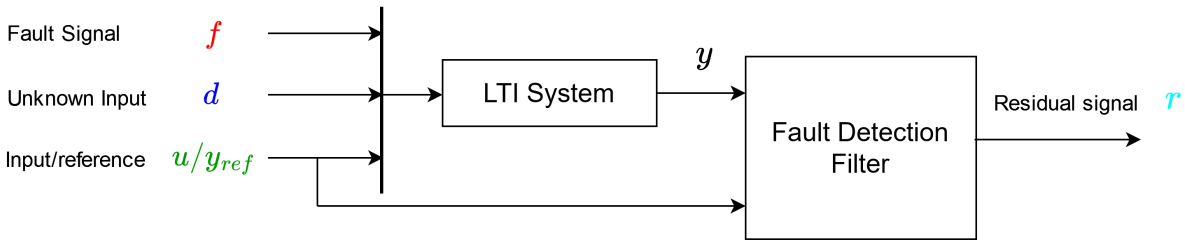


Figure 5: The architecture of the general residual generation filter.

Figure 6 shows the ideal output of the residual generation filter. In the absence of a fault within the system, the residual should approximate zero, meaning the filter must remain insensitive to the nominal dynamics produced by the input/reference signal and any unknown system inputs. However, in the presence of a fault, the filter's output should deviate from zero to give an indication that a fault is present in the system.

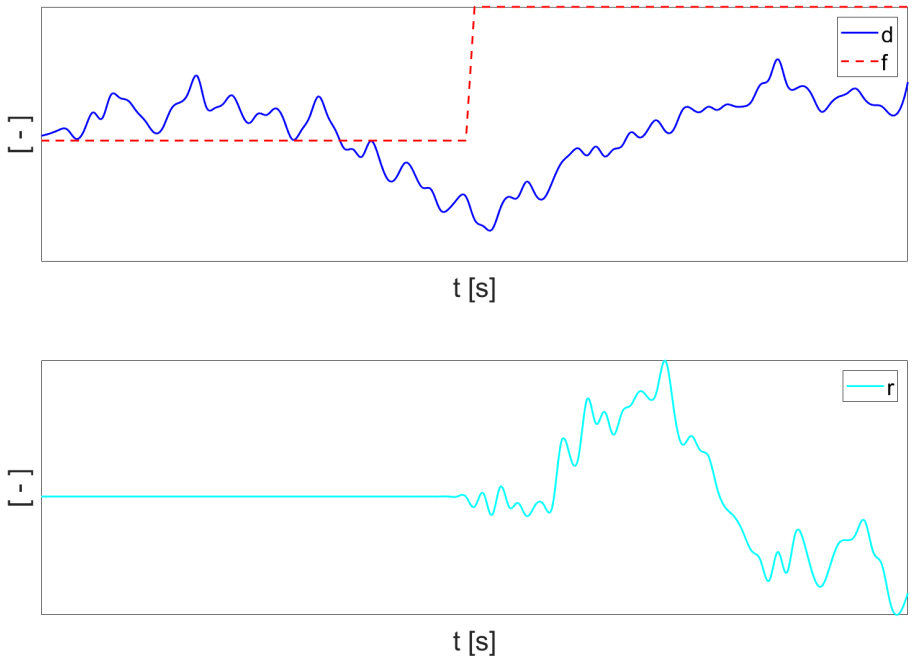


Figure 6: The desired residual filter output: zero-valued in the presence of nonzero unknown inputs, and nonzero-valued when a fault affects the system.

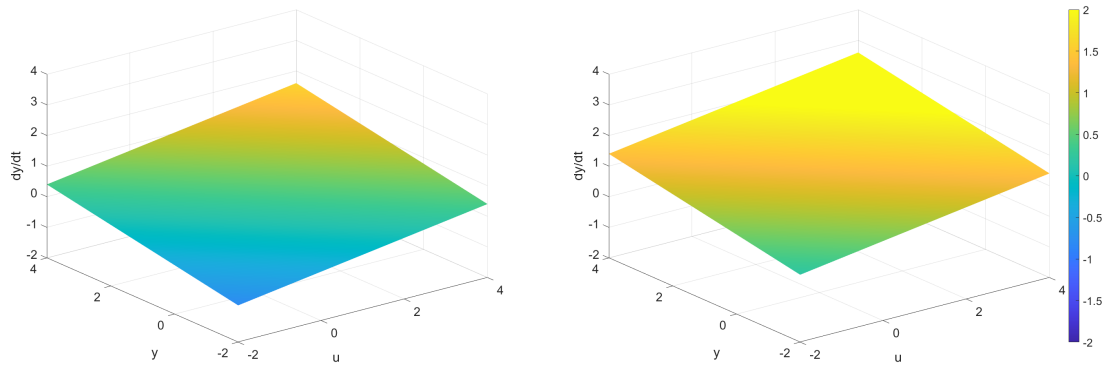
The residual generation method can be mathematically illustrated using a simple ordinary differential equation (ODE) including a possible additive fault f_a*:*

$$\frac{dy}{dt} = 0.2u + 0.2y + f_a$$

When the system operates under normal conditions, $f_a = 0$*, it is anticipated to remain within a defined hyperplane, depicted in Figure 7a. The equation for this hyperplane is defined as:*

$$\frac{dy}{dt} - 0.2u - 0.2y = 0$$

However, the behavior of the system changes in the presence of an additive fault signal, such as $f_a = 1.0$*. Under this circumstance, the behavioral set evolves, and the possible system trajectories are not captured within the initial hyperplane. Instead, these trajectories are found within a new parallel hyperplane shown in Figure 7b.*



(a)

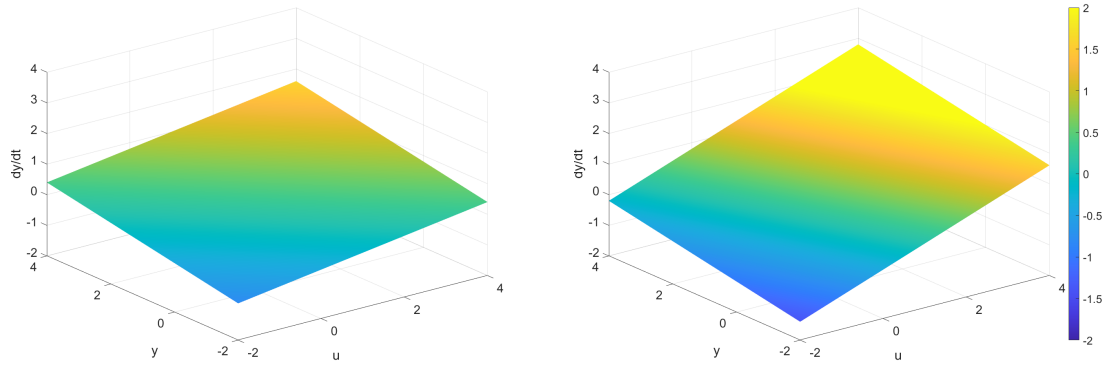
$$\frac{dy}{dt} = 0.2y + 0.2u$$

(b)

$$\frac{dy}{dt} = 0.2y + 0.2u + f_a$$
$$= 0.2y + 0.2u + 1.0$$

Figure 7: The effect of an additive fault $f_a = 1.0$ on the behavioral set $\mathcal{M}$.

Parametric faults f_p can also change the system dynamics and as a result, the possible trajectories are also no longer captured within the initial hyperplane. Instead, they lie within a modified rotated hyperplane as shown in Figure 8b. If the system behavior does not deviate from the axis of rotation, there is a risk this could result in the inability to detect the fault under investigation. However, it is worth noting that the trajectories employed in this study sufficiently excite the system, ensuring that this issue is mitigated.



(a)

$$\frac{dy}{dt} = 0.2y + 0.2u$$

(b)

$$\frac{dy}{dt} = 0.2y + f_p\,0.2u$$
$$= 0.2y + 0.5u$$

Figure 8: The effect of an parametric fault $f_p = 2.5$ on the behavioral set $\mathcal{M}$.

Following the theorem, the residual filter for the given ODE is obtained, which includes an example $d(s)$ that ensures the filter is stable and proper:

$$\begin{aligned} \frac{dy}{dt} - 0.2u - 0.2y &= 0 \\ Ys - 0.2U - 0.2Y &= 0 \\ \frac{Ys - 0.2U - 0.2Y}{\underbrace{s^2 + 2s + 1}_{d(s)}} &= R(s) = 0 \end{aligned}$$

The application of the residual generation method in this work is based on the principle illustrated in this example. Assuming proper operation within the nominal behavioral set, the residual is expected to remain at zero or, in the event of transient behavior triggered by initial conditions, it should eventually converge to zero over time. However, deviations from this set due to faults, should result in a nonzero residual value.

3.2.2 Linear Programming Framework

In [17], a linear programming framework is presented for expressing the polynomial matrices given in Eq. 16. In this framework, if $N(s)$ is a solution to Eq. 16, the polynomial matrices can be expressed as shown in Eq. 17, where d_H, d_F, and d_N correspond to the order of the corresponding polynomial matrices.

$$H(s) := \sum_{i=0}^{d_H} H_i s^i, \quad F(s) := \sum_{i=0}^{d_F} F_i s^i, \quad N(s) := \sum_{i=0}^{d_N} N_i s^i. \tag{17}$$

Then, the conditions stated in Eq. 16 can be written as shown in Eq. 18,

$$\begin{aligned} \bar{N}\bar{H} &= 0, \\ \|\bar{N}\bar{F}\|_\infty &\geq 1 \end{aligned} \tag{18}$$

where the infinite norm is denoted by $\|\cdot\|_\infty$.

It is worth noting that if $\bar{N}$ is a solution to $\bar{N}\bar{H} = 0$, then the negative of $\bar{N}$, denoted as $-\bar{N}$, is also a solution. Consequently, the inequality $\|\bar{N}\bar{F}\|_\infty \geq 1$ can be interpreted as a set of m distinct LP problems, where m corresponds to the number of columns in $\bar{F}$ and is defined as $m = n_f\,(d_F + d_N + 1)$. Here, n_f represents the dimension of the signal f in Eq. 12. The matrices $\bar{N}$, $\bar{H}$, and $\bar{F}$ are provided in Eq. 19 [17].

$$\begin{aligned} \bar{N} &:= \begin{bmatrix} N_0 & N_1 & \cdots & N_{d_N} \end{bmatrix}, \\ \bar{H} &:= \begin{bmatrix} H_0 & H_1 & \cdots & H_{d_H} & 0 & \cdots & 0 \\ 0 & H_0 & H_1 & \cdots & H_{d_H} & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & H_0 & H_1 & \cdots & H_{d_H} \end{bmatrix}, \end{aligned} \tag{19}$$

$$\bar{F} := \begin{bmatrix} F_0 & F_1 & \cdots & F_{d_F} & 0 & \cdots & 0 \\ 0 & F_0 & F_1 & \cdots & F_{d_F} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & F_0 & F_1 & \cdots & F_{d_F} \end{bmatrix},$$

$$\bar{L} := \begin{bmatrix} L_0 & L_1 & \cdots & L_{d_L} & 0 & \cdots & 0 \\ 0 & L_0 & L_1 & \cdots & L_{d_L} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & L_0 & L_1 & \cdots & L_{d_L} \end{bmatrix}$$

Example 4. Generation of the Bar Matrices.

The SSM presented in Example 2 is reconsidered with the inclusion of additive fault dynamics. The new SSM can be expressed as:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix} x + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f, \quad y = \begin{bmatrix} 1 & 0 \end{bmatrix} x$$

The SSM is then transformed into the DAE framework as shown in the method discussed in Section 3.1.5:

$$\underbrace{\begin{bmatrix} -s & 1 \\ -k/m & -s - c/m \\ 1 & 0 \end{bmatrix}}_{H(s)} x + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1/m \\ -1 & 0 \end{bmatrix}}_{L} z + \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{F} f = 0$$

The model is then rewritten in the form of bar matrices, with $d_H = 1$, $d_F = 0$, $d_L = 0$, and $d_N = 2$, as shown in Eq. 19:

$$\bar{H} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -\frac{k}{m} & -\frac{c}{m} & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -\frac{k}{m} & -\frac{c}{m} & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -\frac{k}{m} & -\frac{c}{m} & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \bar{F} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \textit{and} \quad \bar{L} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{m} & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{m} & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{m} \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

Obtain $\bar{N}$ by making use of the method in Eq. 18 and substituting $k = 1$, $b = 1$, and $m = 1$:

$$\bar{N} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

With the corresponding continuous-time residual generation filter:

$$R(s) = \begin{pmatrix} \dfrac{-s^2 - s - 1}{1.333e10^{-9}s^3 + 4e10^{-6}s^2 + 0.003667s + 1} \\ \dfrac{1}{1.333e10^{-9}s^3 + 4e10^{-6}s^2 + 0.003667s + 1} \end{pmatrix}$$

3.2.3 Establishing Thresholds for Residual Signals

To effectively monitor the residual signal $r(t)$ and determine the detectability of potential faults in the system, it is essential to set suitable thresholds. This approach is similar to limit checking, where the minimum and maximum allowable values for the residual signal, denoted as $r_{\min}$ and $r_{\max}$, are established [6]. By comparing the residual signal $r(t)$ with these thresholds, deviations from the acceptable range can be identified. The equation representing this comparison is provided in Eq. 20.

$$r_{\min} < r(t) < r_{\max} \tag{20}$$

To calculate the thresholds $r_{\min}$ and $r_{\max}$, the maximum and minimum values of the residual signal $r(t)$ over a specified number of trajectories N is obtained. The equations for determining $r_{\min}$ and $r_{\max}$, including an additional detectability margin α, are shown in Eq. 21.

$$\begin{aligned} r_{\max} &= \alpha \max_{i=1}^{N} \max_{t} r_i(t) \quad \text{for} \quad i = 1, 2, \ldots, N \\ r_{\min} &= \alpha \min_{i=1}^{N} \min_{t} r_i(t) \quad \text{for} \quad i = 1, 2, \ldots, N \end{aligned} \tag{21}$$

Given that the only unknown disturbance affecting the system is a Band-Limited Gaussian measurement noise, one approach to estimating the parameter α involves fitting a normal distribution to a set of residuals generated from N trajectories. This method allows for the determination of a suitable α based on the characteristics of the observed residuals.

A more detailed design method of the detectability margin α is beyond the scope of this research, but it is noteworthy to mention its functionality. This margin could be obtained after deployment or repair as a calibration step of the system of interest. The inclusion of a detectability margin acts as a buffer, providing an additional layer of robustness to the fault detection mechanism. This detectability margin accommodates potential disturbances or variations in normal operations without triggering a false fault indication. Thus, the optimal selection of a detectability margin is a balance between accurate fault detection and avoiding unnecessary alerts.

3.3 Fault Estimation

The following section outlines the methodological approach implemented for estimating faults as discussed in Section 3.3.1, for estimation of external disturbances referenced in Section 3.3.2, and a method for the segmentation of high and low-frequency components within the residual signal as elaborated in Section 3.3.3.

3.3.1 General Fault Estimation Framework

Figure 9 illustrates the fault estimation framework, an adaptation based on the work in [25]. This enhanced framework integrates additional elements into the fundamental fault detection system, as represented in Figure 5. The signal ϕ comprises multiple sub-signals, each generated by the pre-filter $\Phi(s)$, which are used to assess the impact of individual fault components within the residual signal r, produced by the residual generation filter R(s). In the fault estimation block, a regression method is employed to estimate the magnitude of each specific fault present in the system. Furthermore, the framework introduces the integration of a new filter $\mathrm{R_{dist}(s)}$ in combination with a high-pass filter HP(s). These filters are used in combination with power spectral density analysis, enabling the estimation of elevated levels of noise and unknown external disturbances, represented as $\hat{d}$.

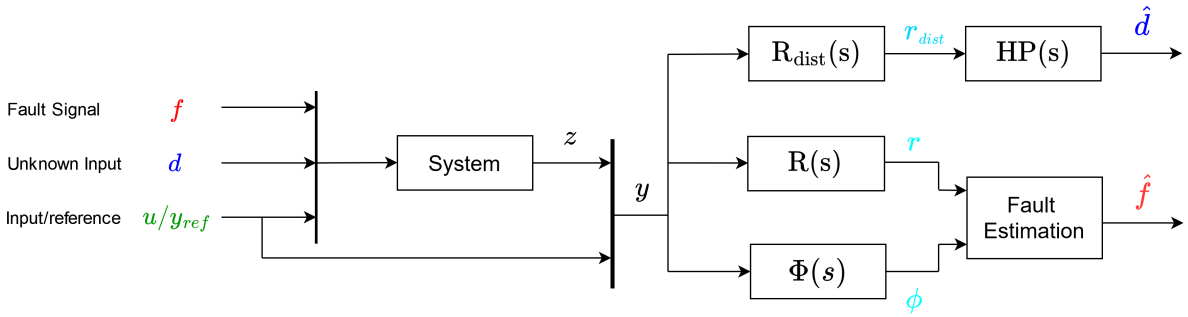


Figure 9: The architecture used for fault and disturbance estimation.

Generation of Fault Mode Components

The set of faults F and the corresponding fault magnitudes C used for fault estimation are presented in Eq. 22. It is important that the fault magnitudes c_i are sufficiently large to ensure detectability through the residual generation method.

$$\begin{aligned} F &:= \{f_1, f_2, \cdots, f_{n_f}\} \\ C &:= \{c_1, c_2, \cdots, c_{n_f}\} \end{aligned} \tag{22}$$

For each fault mode i in the set F, the corresponding system matrices are derived. The difference between these matrices and the nominal system matrices are scaled by the corresponding fault magnitude c_i from C. This process generates delta system matrices that capture the difference between the nominal system and the system under fault mode i.

$$\begin{aligned} \Delta A_i &= \frac{A_i - A_{nom}}{c_i} \quad \Delta B_i = \frac{B_i - B_{nom}}{c_i} \\ \Delta C_i &= \frac{C_i - C_{nom}}{c_i} \quad \Delta D_i = \frac{D_i - D_{nom}}{c_i} \end{aligned} \tag{23}$$

Generation of Prefilter Numerator

It is necessary to derive the $\bar{H}_i$ and $\bar{L}_i$ matrices for each faulty system, characterized by the set of state-space model (SSM) matrices as presented in Eq. 23. To obtain the pre-filter Φ, Eq. 24 can be used, resulting in a residual sub-signal for each corresponding fault [7].

$$\Phi_i = \bar{N}(\bar{L}_i - (\bar{H}_i(\bar{H}^T\bar{H})^{-1})\bar{H}^T\bar{L}) \tag{24}$$

Regression method

The regression framework is shown in Eq. 25 where n_f denotes the number of faults and T represent the time horizon. The time horizon must be sufficiently large to effectively capture each residual sub-signal, thus allowing for accurate detection and estimation of faults. Furthermore, the sampling frequency should be adequate to guarantee that the highest frequency components present in the faulty residual signal are also identified.

$$\underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_T \end{bmatrix}}_{r} = \underbrace{\begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,n_f} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,n_f} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{T,1} & \phi_{T,2} & \cdots & \phi_{T,n_f} \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \vdots \\ \hat{f}_{n_f} \end{bmatrix}}_{\hat{f}} \tag{25}$$

To determine the fault estimates $\hat{f}$ an Ordinary Least Squares (OLS) estimator is implemented as shown in Eq. 26 [15].

$$\hat{f} = (G^TG)^{-1}G^Tr \tag{26}$$

3.3.2 Estimating External Disturbances

In order to detect the presence of high-frequency faults, such as external disturbances or elevated levels of measurement noise within the system, power spectral density (PSD) analysis can be used. PSD is a mathematical framework used to understand how the power of a signal is distributed across different frequencies. It can provide valuable insights into the dominant frequencies and allows the analysis of the overall frequency components present in the signal under examination.

The continuous-time residual signal, denoted as $r(t)$, can be transformed using the Fourier transform [18]. This transformation is represented in Eq. 27.

$$\hat{r}(f) \triangleq \int_{-\infty}^{\infty} e^{-i2\pi ft} r(t)dt \tag{27}$$

The Power Spectral Density of a signal $r(t)$, represented as $S_{rr}(f)$, is defined as the square of the magnitude of its Fourier transform, $\hat{r}(f)$. This is expressed in Eq. 28 [22].

$$S_{rr}(f) \triangleq |\hat{r}(f)|^2 \tag{28}$$

In this study, thresholds are computed for both the peak values and the total frequency spectrum. Peak value thresholds contribute to the detection of specific frequencies of interest, while total frequency spectrum thresholds offer a possible understanding of the measurement noise in the system through integration over the entire PSD signal [3].

Threshold for Peak Values

The peak value threshold, denoted by R_{peak}, is derived by computing the PSD of numerous residual signals associated with typical machine trajectories N. Each signal is generated by the residual generation filter for the nominal system and corresponds to individual trajectories. Threshold R_{peak} includes a detectability margin, denoted as β_{peak}, incorporated to account for variations and uncertainties in the system. The PSD of the residual signal, $R(f)$, can be continuously analyzed to identify dominant frequency components surpassing the pre-set threshold R_{peak}. This concept is mathematically represented in Eq. 29. When the PSD value at a specific frequency exceeds the R_{peak} threshold, it suggests the presence of a predominant frequency within the system, potentially due to a periodic external disturbance signal.

$$\begin{aligned} R_{\text{peak}} &= \max_{i=1}^{N}(\max_{f}(|R_{r,i}(f)|)) + \beta_{peak} \\ R(f) &> R_{\text{peak}} \end{aligned} \tag{29}$$

Threshold for Total Noise

The total frequency spectrum threshold, symbolized as R_{total}, is established by integrating the PSD of residual signals over a specified frequency range, denoted as $[f_{\min}, f_{\max}]$, across a similar set of trajectories N as used for the derivation of R_{peak}. This threshold also incorporates a detectability margin β_{total} for similar reasoning as previously discussed. The integrated value signifies the total frequency spectrum, which is subsequently compared to the pre-set threshold, R_{total}. This concept is mathematically defined in Eq. 30. If the total frequency spectrum exceeds the R_{total} threshold, it may suggest an elevation in system noise compared to nominal noise levels.

$$\begin{aligned} R_{\text{total}} &= \max_{i=1}^{N} \int_{f_{\min}}^{f_{\max}} R_{r,i}(f)df + \beta_{total} \\ \int_{f_{\min}}^{f_{\max}} R_r(f)df &> R_{\text{total}} \end{aligned} \tag{30}$$

By implementing these thresholds for the PSD, the system can effectively monitor high-frequency faults and identify dominant frequencies, which could aid in improving the reliability of the fault estimation system.

3.3.3 Segmentation of High- and Low-Frequency Signal Components

The fault estimation framework shown in Figure 9 incorporates two separate residual generation filters.

Filter $R_{dist}(s)$ aims to detect the existence of external disturbances and elevated noise. The denominator $d(s)$ in this high-frequency oriented residual generation filter should be designed such that it does not attenuate high-frequency components in the signal. This can be achieved using the Butterworth filter denominator, characterized by a sufficiently high cutoff frequency, as discussed in Section 3.2. Furthermore, a high pass filter can be employed to attenuate as many low-frequency components generated by faults as possible. If the noise and disturbances operate at frequencies higher than those produced by the residual generation filter for all other faults, segregation of external disturbances from other faults is feasible.

Filter $R(s)$ is dedicated to detect the parametric faults and suppress high-frequency noise and disturbances. The design approach for the denominator $d(s)$ follows a similar method, where it should have a cutoff frequency equal to or higher than the highest frequency signal produced by all other faults in their respective residual. The maximum frequency present in a signal can be determined by executing a Fourier analysis of the signal, as depicted in Eq. 27.

4 Design and Implementation

This chapter outlines the architectural design and implementation procedures of the study. It is divided into the following parts: Section 4.1 details the fault detection system, outlining its architecture and the specific configurations used during simulations. Section 4.2 goes into depth about the fault estimation system, highlighting its architecture and the settings used in the simulations. The controller implemented in this study is discussed in Section 4.3.

4.1 Fault Detection

This section examines the architecture of the fault detection system employed to assess fault detectability for both linear and nonlinear EoM in this study, as shown in Figure 10. Depending on the implementation of the feedback control element, the reference signal y_{ref} is used to generate the feedback control input signal u_{fb}. The feedforward element is represented by u_{ff}. The external disturbances act as an additive fault f_a, while faults associated with the motion guides, structural stiffness, and the actuator are denoted as the parametric fault signal f_p. To accurately model sensor precision, Band-Limited Gaussian white noise is incorporated as measurement noise w. The augmented input signal u and output signal y are denoted as signal z, which undergoes a transformation into a residual signal r using the residual generation filter discussed in Section 3.2.

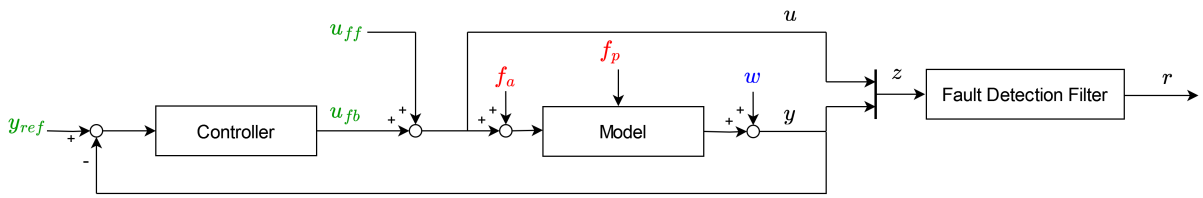


Figure 10: Schematic diagram of the residual generation filter for both the linear and nonlinear model.

4.1.1 Fault Detection Simulation Parameters

In all simulation scenarios, successful isolation of all faults from the nominal dynamics was accomplished using a filter order, denoted d_N, of seven. The sampling frequency was set at Fs = 8e3 Hz, and the simulations were conducted over a duration of T_{sim} = 2.5 seconds. To mimic real-world conditions, band-limited white noise was incorporated with an amplitude corresponding to a noise power of 5e-20 dB/Hz. The specific controller used in this study is further explained in Section 4.3.

4.1.2 Designing the Denominator of the Residual Generation Filter

The denominator, denoted as $d(s)$, of the residual generation filter $R(s)$ was chosen to be a seventh-order Butterworth low-pass filter with a cut-off frequency of 50 Hz. This choice was based on the order of the numerator, ensuring the realizability of the residual generation filter. The low-pass filter, referred to as LP(s), is represented by Eq. 31, and its Bode magnitude response can be found in Appendix A.3.

$$\text{LP(s)} = \frac{3.02e17}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \tag{31}$$

4.1.3 Residual Generation Filter

The residual generation filter employed in the simulations of the linear model is defined by Eq.32. The denominator of the filter is determined using Eq.31, while the numerator is designed following the approach detailed in Section 3.2.

$$\text{R(s)} = \begin{pmatrix} \frac{-3.464e-07s^7 + 0.07468s^6 + 77.66s^5 + 1.547e07s^4 + 3.761e09s^3 + 5.525e12s^2 + 1.837e14s}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \\ \frac{-0.01483s^6 - 47.13s^5 - 7.812e06s^4 - 1.16e10s^3 - 7.205e12s^2 - 7.783e15s}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \\ \frac{-554.4s^5 - 4.341e07s^4 - 2.124e10s^3 - 3.488e13s^2 - 8.258e14s - 3.02e17}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \\ \frac{2.035e-53s^6 + 2.148e-07s^5 - 0.04644s^4 - 29.11s^3 - 4.586e06s^2 - 9.459e08s - 1.584e12}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \\ \frac{5.819e-52s^6 + 9.146e-12s^5 + 0.008674s^4 + 1838s^3 + 1.46e08s^2 + 6.751e10s + 1.128e14}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \\ \frac{8.838e-52s^6 + 2.245e-46s^5 + 0.04913s^4 + 3.334e06s^3 + 2.61e11s^2 + 1.228e14s + 2.051e17}{s^7 + 1412s^6 + 9.966e05s^5 + 4.524e08s^4 + 1.421e11s^3 + 3.09e13s^2 + 4.32e15s + 3.02e17} \end{pmatrix} \tag{32}$$

4.2 Fault Estimation

This section details the architecture and settings for the fault estimation framework used during the simulations.

4.2.1 Estimation Architecture

Notably, it incorporates the pre-filter $\Phi(s)$ to generate multiple residual sub-signals ϕ. These sub-signals, in combination with the residual r, are transformed into fault estimates $\hat{f}$. Additionally, the disturbance residual generation filter R_{dist} is utilized to produce the disturbance residual r_{dist}. Lastly, the high-pass filter HP(s) is applied to convert r_{dist} into the estimated disturbance signal $\hat{d}$ by making use of spectral density analysis. The architecture of this approach is illustrated in Figure 11.

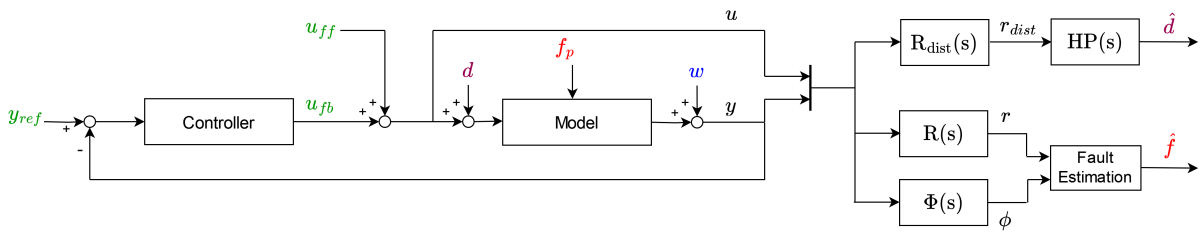


Figure 11: Schematic representation of the implemented fault estimation architecture for the linear simulation model.

4.2.2 Simulation Settings for Fault Estimation

In the simulation scenarios employed for fault estimation, the filter order d_N is set consistently at seven, reflecting the conditions applied during the fault detection simulations. Other parameters, such as the sampling frequency, total simulation time, and noise factors, remain unchanged from those established in the fault detection scenarios. Additionally, the controller from the previous scenarios is repurposed. The matrix G in Equation 26 includes the residual sub-signals ϕ generated by the pre-filter $\Phi(s)$. The sampling rate for the residual sub-signals ϕ is set to 0.0005 seconds to accurately represent the underlying data and a regression horizon, denoted by T, of 60000 samples is chosen, providing a comprehensive analysis of the sub-signals.

4.2.3 Disturbance Residual Generation Filter

The disturbance residual generation filter is formulated as shown in Equation 33. This filter in combination with a high-pass filter aims to isolate the disturbance effect from the original signals to enhance the accuracy of the estimation process.

$$
\mathrm{R}_{dist}(s) = \begin{pmatrix}
\frac{-3.464e-07s^7 + 0.07468s^6 + 77.66s^5 + 1.547e07s^4 + 3.761e09s^3 + 5.525e12s^2 + 1.837e14s}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25} \\
\frac{-0.01483s^6 - 47.13s^5 - 7.812e06s^4 - 1.16e10s^3 - 7.205e12s^2 - 7.783e15s}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25} \\
\frac{-554.4s^5 - 4.341e07s^4 - 2.124e10s^3 - 3.488e13s^2 - 8.258e14s - 3.02e17}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25} \\
\frac{2.035e-53s^6 + 2.148e-07s^5 - 0.04644s^4 - 29.11s^3 - 4.586e06s^2 - 9.459e08s - 1.584e12}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25} \\
\frac{5.819e-52s^6 + 9.146e-12s^5 + 0.008674s^4 + 1838s^3 + 1.46e08s^2 + 6.751e10s + 1.128e14}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25} \\
\frac{8.838e-52s^6 + 2.245e-46s^5 + 0.04913s^4 + 3.334e06s^3 + 2.61e11s^2 + 1.228e14s + 2.051e17}{s^7 + 2.118e04s^6 + 2.242e08s^5 + 1.527e12s^4 + 7.196e15s^3 + 2.347e19s^2 + 4.921e22s + 5.16e25}
\end{pmatrix} \tag{33}
$$

4.2.4 High-Pass Filter Description

The high-pass filter HP(s) integrated into the disturbance estimation simulations is presented in Equation 34. This filter is a third-order Butterworth high-pass filter, with its cutoff frequency configured at 100 Hz, which corresponds to the lowest frequency present in the disturbance signal. The Bode magnitude response of the high-pass filter is provided in Appendix A.4.

$$\text{HP(s)} = \frac{s^3}{s^3 + 1257s^2 + 7.896e05s + 2.481e08} \tag{34}$$

4.3 Controller Design

The controller design for the simulations was facilitated by ASMPT [14]. It should be noted that the detailed design process of the controller falls beyond the scope of this work. The primary focus of this section is to provide an overview of the control scheme's application and the manner in which it regulates system behavior.

4.3.1 Feedforward Control

Feedforward control serves as a predictive control strategy, aiming to improve system performance by compensating for the predictable influences in advance. The control signals for each actuator are generated according to Eq. 35. These control signals are generated by incorporating the predicted trajectory's derivative elements, namely velocity, acceleration, and snap. These derivatives are multiplied by specified gains and added together to create the feedforward control signals. Thus, the feedforward control essentially anticipates the system's input needed based on the desired trajectory.

$$\begin{aligned}
C_x^{ff}(t) &= 2.9926e-09\,\dddot{x}_{ref}(t) + 0.0551\,\ddot{x}_{ref}(t) + 1.9117\,\dot{x}_{ref}(t) \\
C_y^{ff}(t) &= 3.6429e-10\,\dddot{y}_{ref}(t) + 0.0713\,\ddot{y}_{ref}(t) + 2.7063\,\dot{y}_{ref}(t) \\
C_z^{ff}(t) &= 1.3069e-09\,\dddot{z}_{ref}(t) + 0.0032\,\ddot{z}_{ref}(t) - 0.0028\,\dot{z}_{ref}(t)
\end{aligned} \tag{35}$$

4.3.2 Feedback Control

Our research utilizes a feedback control strategy that comprises three distinct controllers. Each controller generates an input signal for the corresponding actuator to minimize the error between the measured and reference signals for each nominal DoF as defined in Table 1. Essentially, the feedback controller continually modifies the control inputs based on the observed error, which is the difference between the desired (reference) trajectory and the actual trajectory. This adjustment process allows the system to respond effectively to disturbances or changes in system dynamics.

$$
\begin{aligned}
C_x^{fb}(s) &= \frac{8.384e05s^2 - 1.644e06s + 8.06e05}{s^2 - 1.119s + 0.1185} \\
C_y^{fb}(s) &= \frac{6.69e04s^2 - 1.316e05s + 6.47e04}{s^2 - 1.828s + 0.8277} \\
C_z^{fb}(s) &= \frac{1.223e04s - 1.192e04}{s - 0.7891}
\end{aligned}
\tag{36}
$$

By effectively combining the predictive capabilities of feedforward control with the responsive capabilities of feedback control, the overall system performance can be significantly enhanced, and the system's behavior can be effectively regulated, even in the face of disturbances or changes in system dynamics. Figure 12 presents the control architecture implemented in all simulations throughout this study. Here, x_{enx}, y_{enc}, and z_{enc} signify the respective measurements of each motion stage.

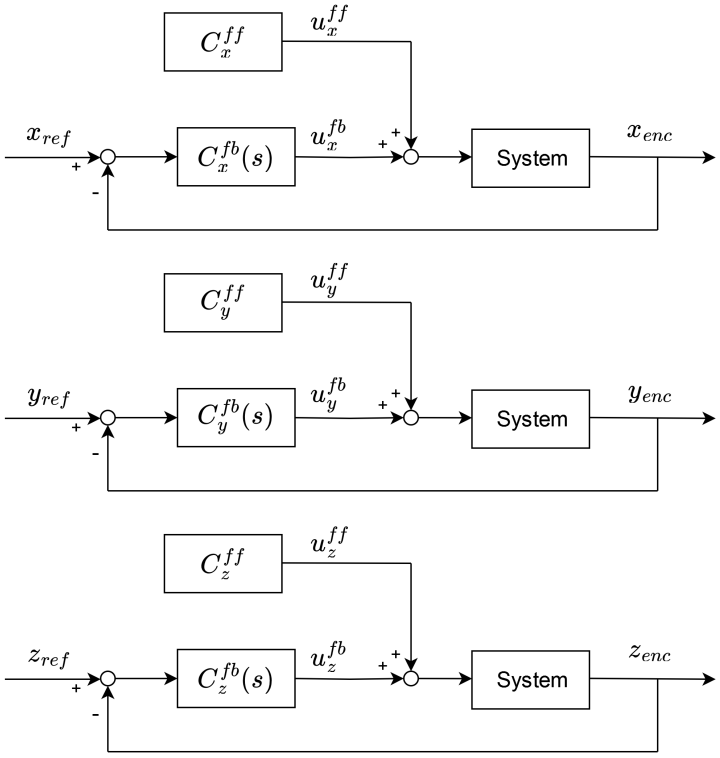


Figure 12: The control architecture implemented in both linear and nonlinear simulations.

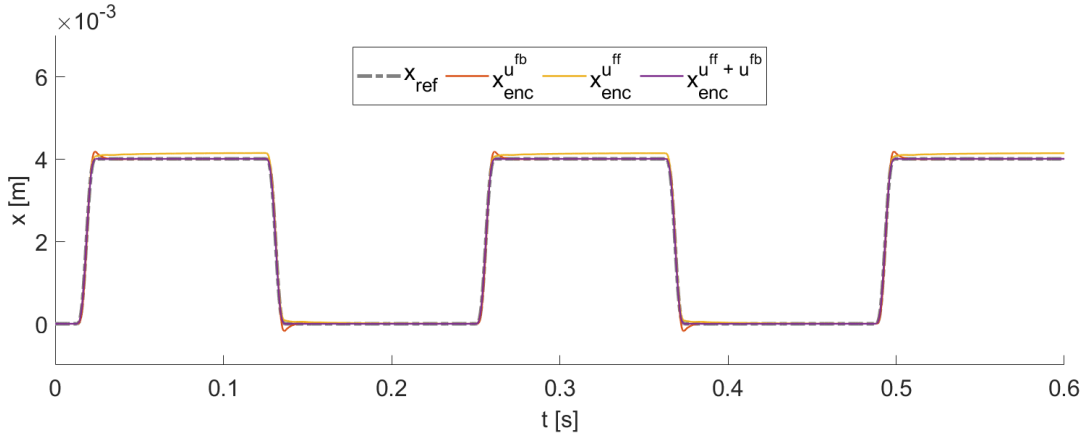
5 Results

This chapter provides a thorough review of our findings. In Section 5.1, the experimental reference and input signals utilized for the evaluation of the fault diagnosis filter are described. In Section 5.2, the analysis of generated thresholds is presented. Section 5.3 explores the detectability of faults for the linear EoM, with various control strategies under consideration. Furthermore, Section 5.4 investigates fault detectability when both feedback and feedforward control are applied to the nonlinear EoM. Lastly, the results of our fault estimation, complete with an estimation of noise levels and the assessment of external disturbances, are illustrated in Sections 5.5 and 5.6. The chapter is concluded in Section 5.7 with a detailed analysis of the obtained results.

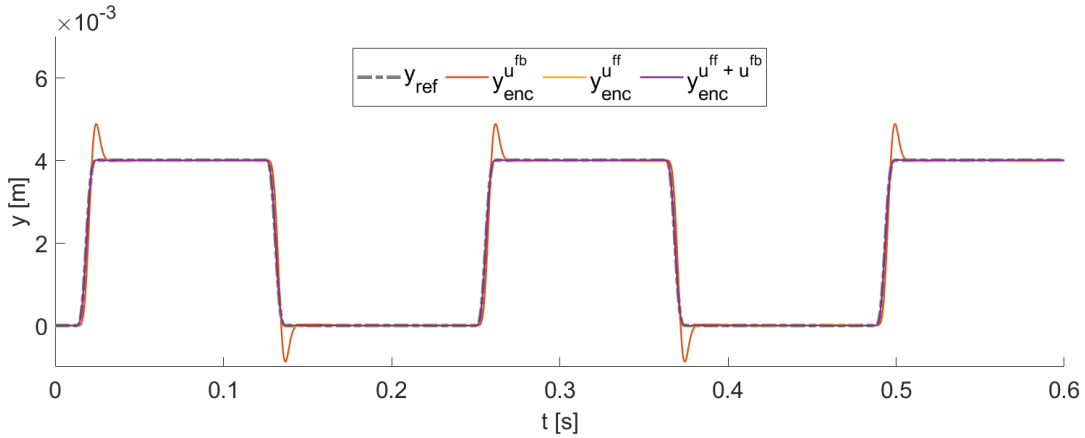
5.1 Reference and Input Signals for Experimentation

To evaluate the effectiveness of the proposed fault diagnosis filter, a series of simulations are conducted. Specifically, a reference trajectory is designed to incorporate typical movements observed in the operation of the AB383 wire bonder, as depicted in Figure 13. This trajectory consists of three distinct reference trajectories, covering translational movements in the x and y direction and a rotational movement to obtain the z motion. This provides a comprehensive representation of the system's operational behavior. Different control strategies are implemented to track the reference signal to enable performance evaluation of the fault diagnosis filter under different control approaches consisting of both feedforward and feedback control elements discussed in Section 4.3.

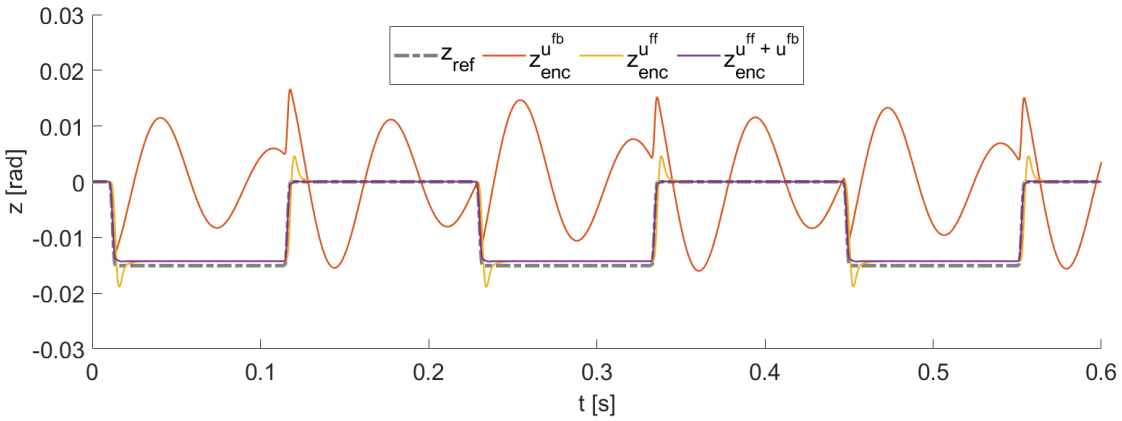
Figure 14 provides a visual representation of the input signals that are used to follow the reference signals of the first movement, as depicted in Figure 13. These input signals correspond to the three spatial dimensions: x-motion, y-motion, and z-motion. The results demonstrate the system exhibits high accuracy and conformity to the desired reference signal when a combination of feedback and feedforward control is implemented. This observation highlights the importance of testing the fault diagnosis filter under the presence of both control strategies to achieve optimal performance in system behavior.



(a) A comparison of x-motion measurements under various control strategies during the process of reference tracking.

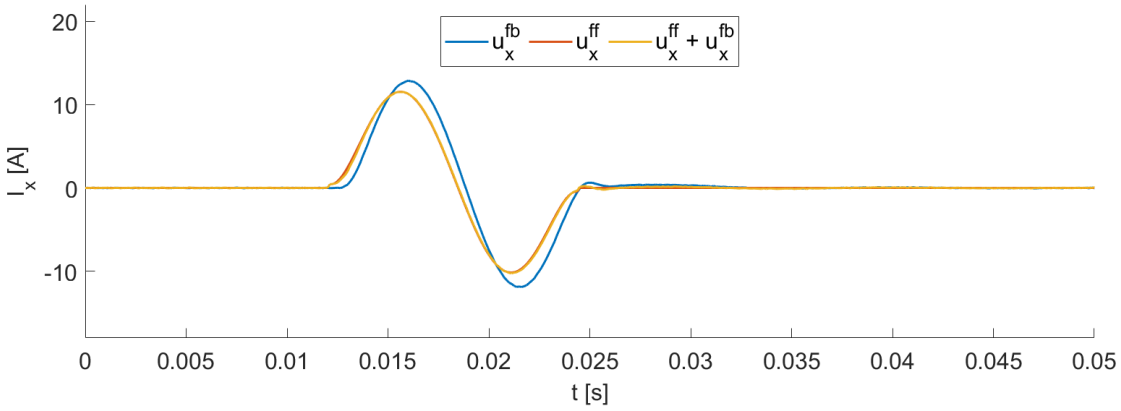


(b) A comparison of y-motion measurements under various control strategies during the process of reference tracking.

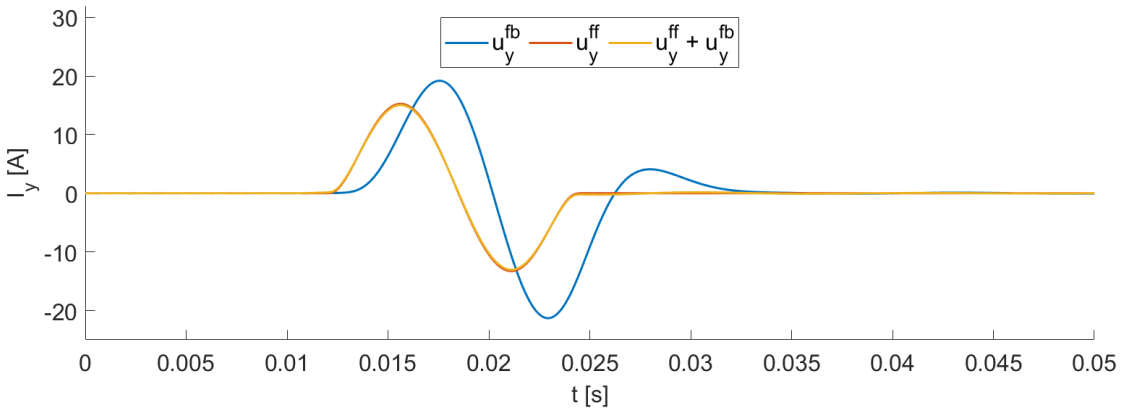


(c) A comparison of z-motion measurements under various control strategies during the process of reference tracking.

Figure 13: The reference signals used to analyze the different control strategies on all spatial measurements during the process of reference tracking.

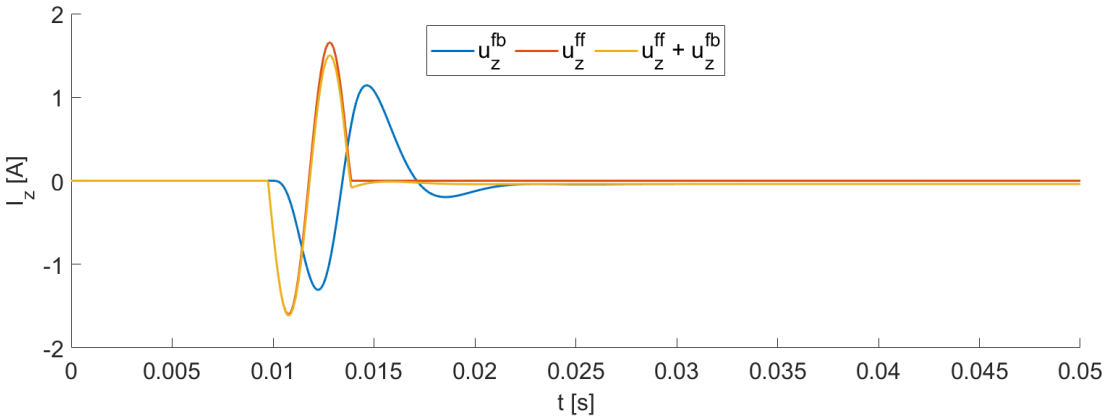


(a) The input signal corresponding to the x-actuator while tracking the reference signal in Figure 13a



(b) The input signal corresponding to the y-actuator while tracking the reference signal in Figure 13b

s



(c) The input signal corresponding to the z-actuator while tracking the reference signal in Figure 13c

Figure 14: Input signals applied to the actuators during the first 0.05 s of the reference signal tracking process, as depicted in Figure 13.

5.2 Threshold Design and Detectability Margin

In the context of fault detectability, the initial procedure involves the design of thresholds for the residual signal. This process, outlined in Section 3.2.3, makes use of the maximum r_{max} and minimum r_{min} values of the residual signal. These values are determined for various trajectories during the normal, fault-free operation of the system, consisting of the setpoints shown in Figure 13. These trajectories include movements in a single direction, represented as r_i where i belongs to the set of spatial dimensions x, y, z. Furthermore, they encompass combinations of movements, denoted by r_{xy} and r_{xyz}. Figure 15 provides a graphical representation of the designed thresholds using the aforementioned method, including a detectability margin α.

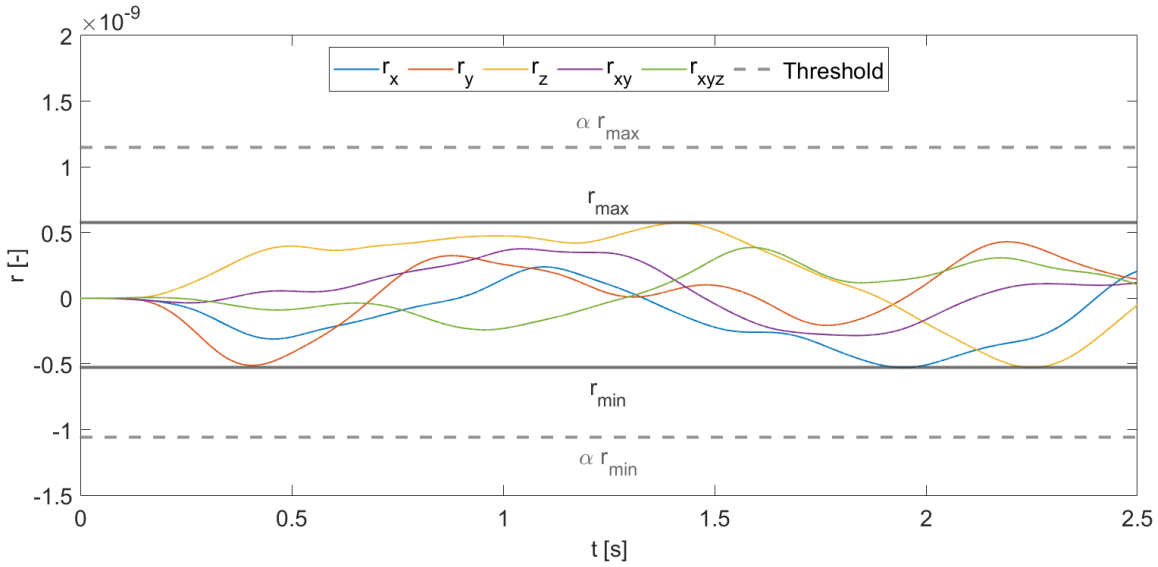


Figure 15: The designed threshold for the residual signal, based on the residual signals for the nominal system during various different trajectories consisting of the reference trajectories illustrated in Figure 13.

The histogram and estimated probability distribution of the residual signals for the nominal system are presented in Figure 16. These findings depict the thresholds applied in the fault detectability examination, set at $\alpha = 2$ to minimize the likelihood of false positives.

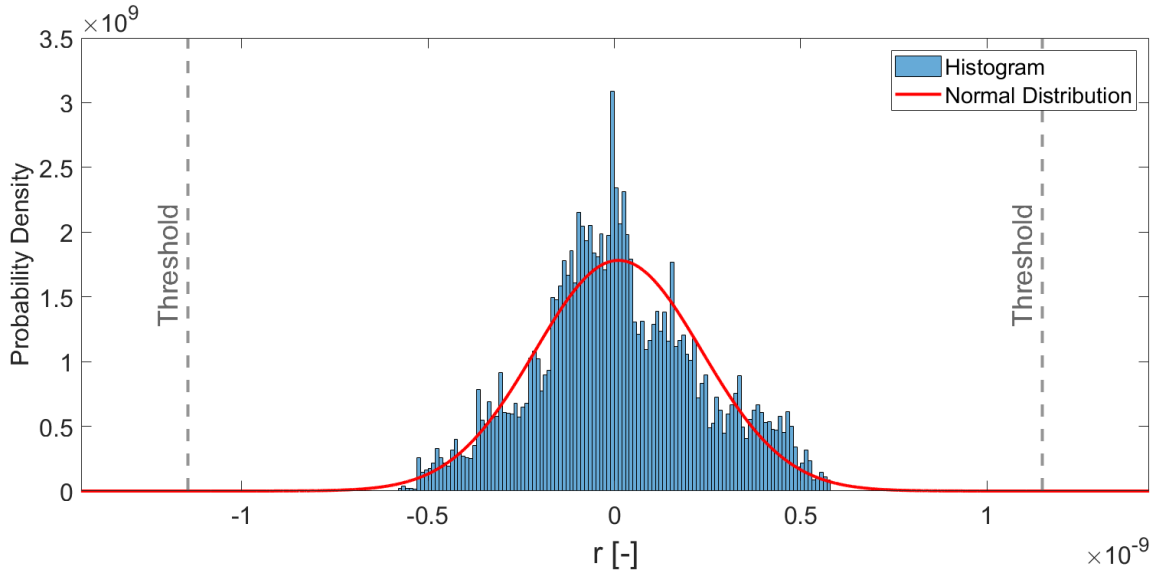


Figure 16: Histogram and estimated normal distribution of the nominal residual signals shown in Figure 15, including the thresholds used in Figures 17.

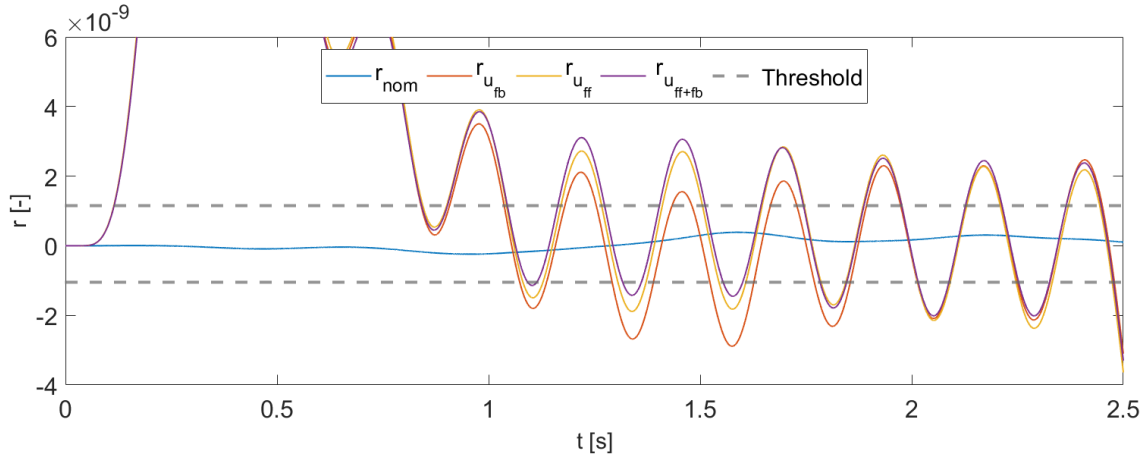
5.3 Fault Detectability for the Linear Model

To assess fault detectability before progressing to fault estimation, the technique elaborated in Section 3.2 is utilized for generating residuals. Each fault is incorporated individually into the system, and its detectability is tested. The thresholds are determined through the procedure discussed in Section 3.2.3. The specifics of the linear model, controller, and simulation configurations, integral to the fault detection process, are detailed comprehensively in Section 4.1.

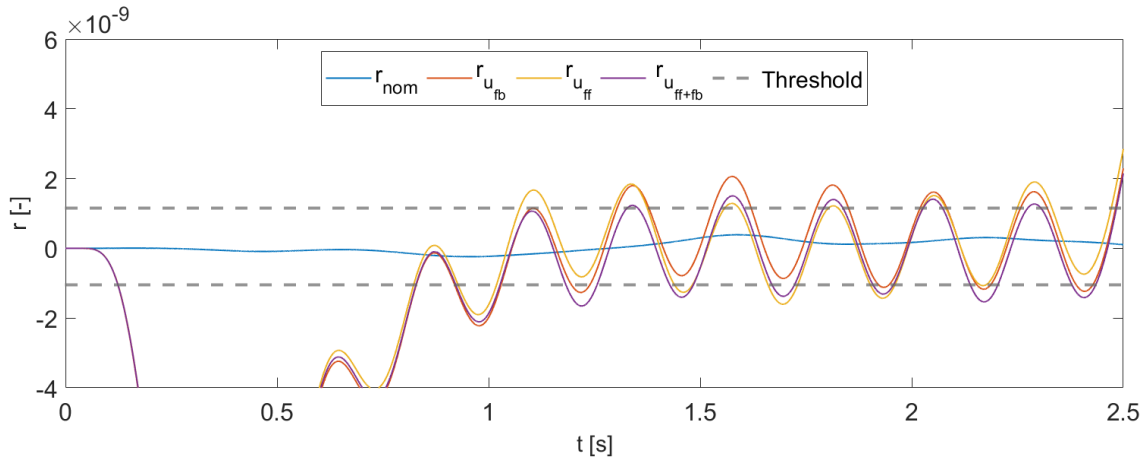
5.3.1 Generated Residuals

To analyze the effect of various control strategies on the produced residuals, three different experimental scenarios were designed. In these scenarios, the system is manipulated using either feedback control, feedforward control, or a combination of both strategies. Figure 17 provides a visual representation of the residuals corresponding to each fault under different control architectures.

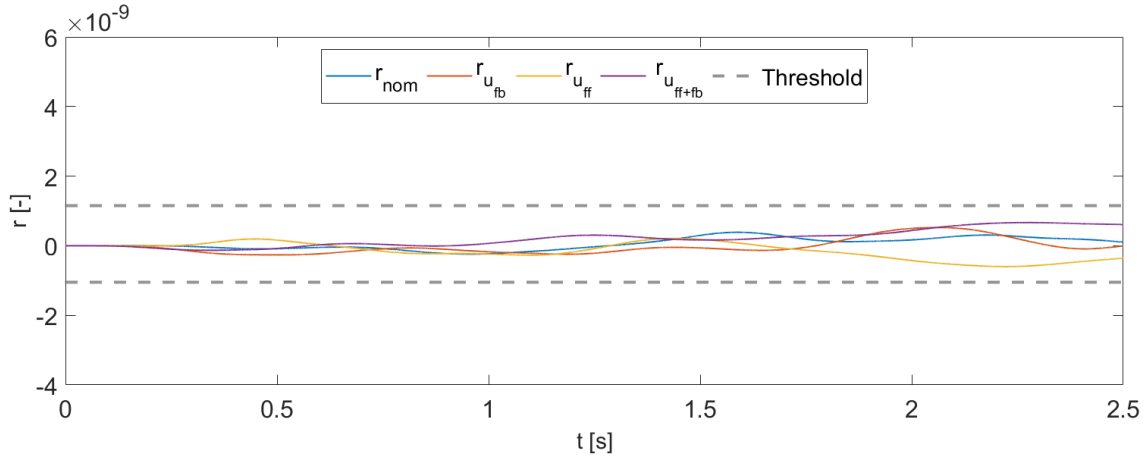
The residuals corresponding to the trajectory in the nominal system are represented as r_{nom}. Residuals generated in the faulty system are represented as follows: under feedback control as $r_{u_{fb}}$, under feedforward control as $r_{u_{ff}}$, and under the combined application of both control methods as $r_{u_{ff+fb}}$.



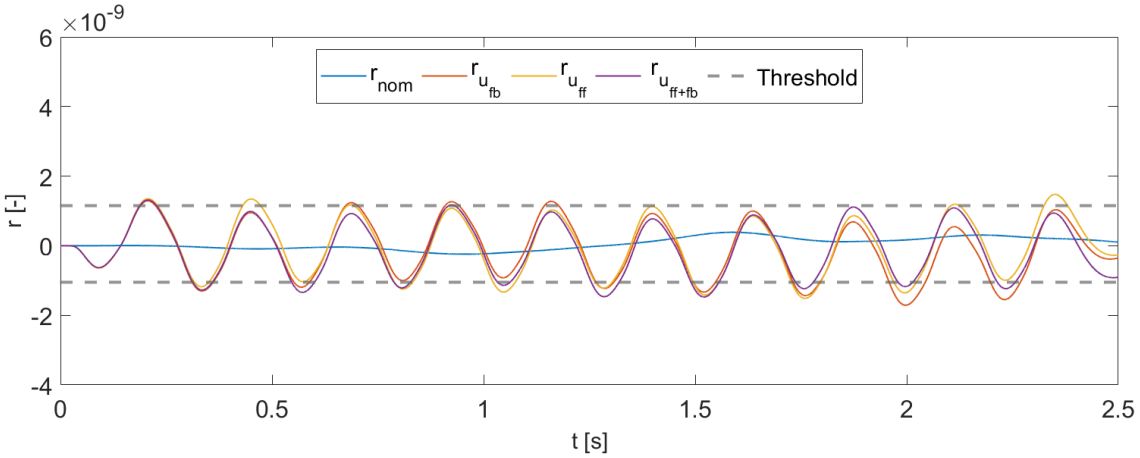
(a) Residual filter output with a 0.5% increase in viscous friction in the x-motion guide. ($\zeta_x = 1.005$)



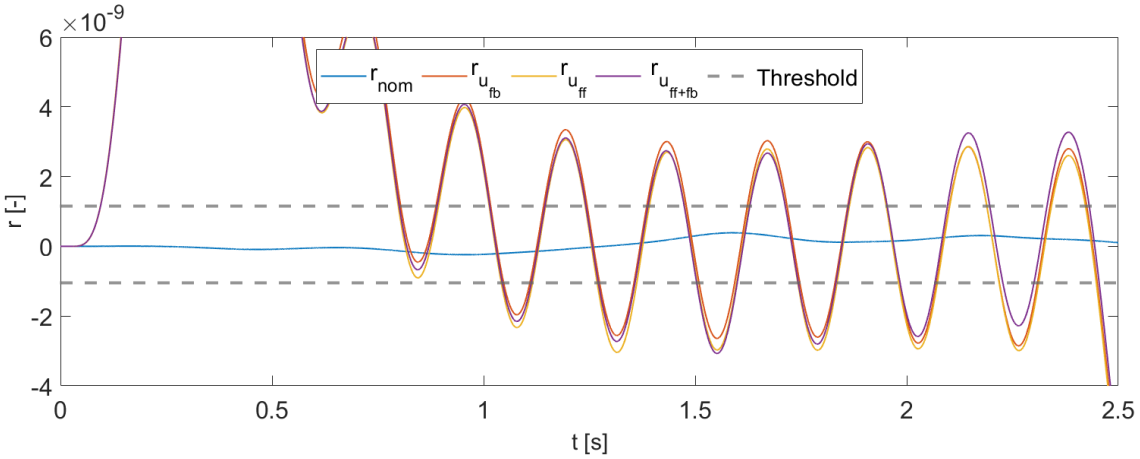
(b) Residual filter output with a 0.1% increase in viscouss friction in the y-motion guide. ($\zeta_y = 1.001$)



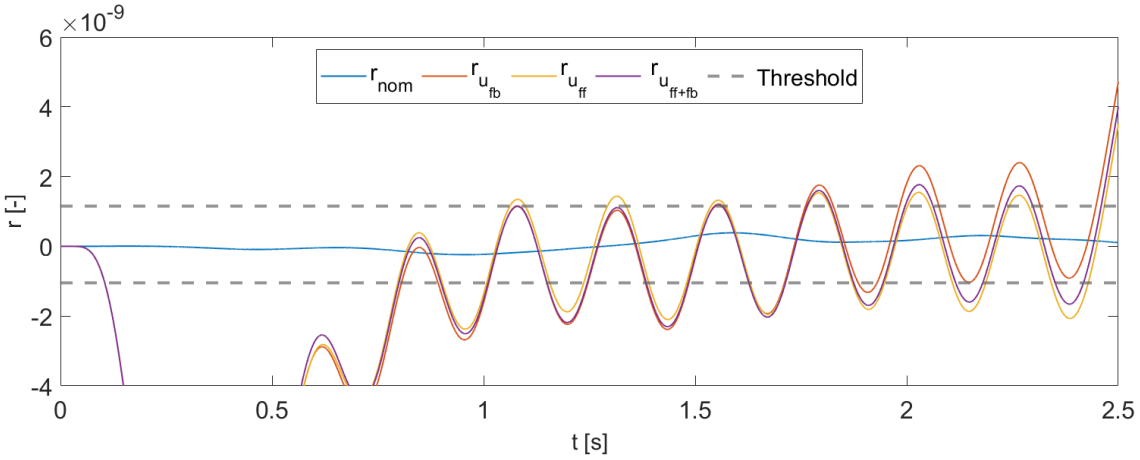
(c) Residual filter output with a 99.9% decrease in structural stiffness in the x-motion stage. ($\beta_x = 0.001$)



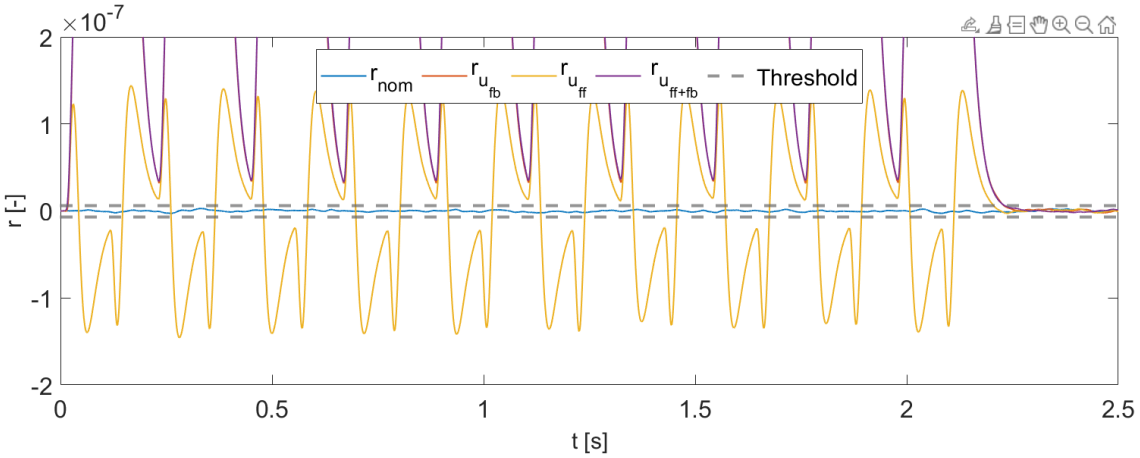
(d) Residual filter output with a 99.9% decrease in structural stiffness in the y-motion stage. ($\beta_y = 0.001$)



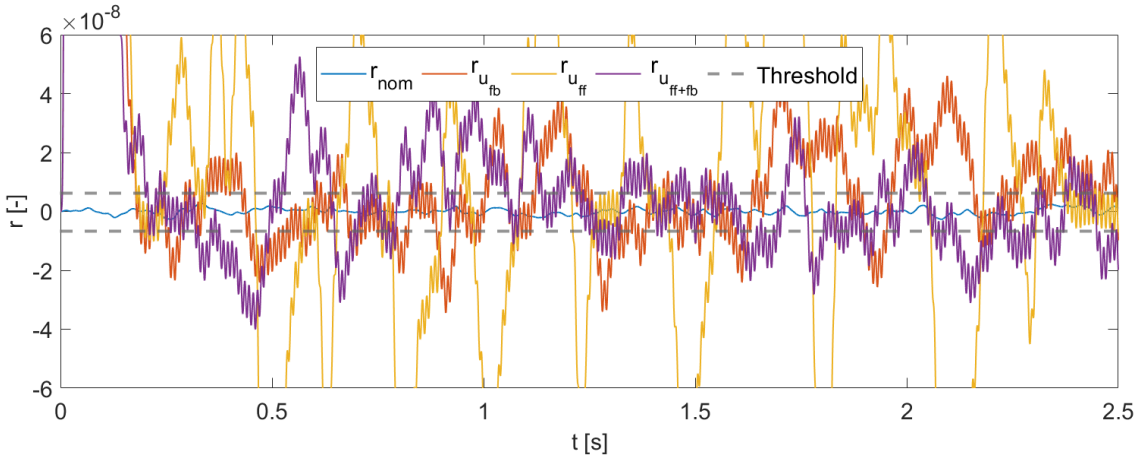
(e) Residual filter output with a 0.5% decrease of actuator force the x-motion. ($\eta_x = 0.995$)



(f) Residual filter output with a 0.1% decrease of actuator force the y-motion. ($\eta_x = 0.999$)



(g) Residual filter output with a 0.01% decrease of actuator force the z-motion. ($\eta_x = 0.9999$)



(h) Residual filter output with 0.1% External disturbances present in the system. ($\gamma = 0.001$)

Figure 17: Residual signals during a simulation with individual faults implemented and various control strategies such as feedforward, feedback, and a combination as input signals shown in Figure 14.

5.3.2 Fault Detection Performance Analysis

Table 2 presents an overview of the detectability of each fault using the residual generation filter for the linear system controlled by a combination of feedback and feedforward control. The fault symbols and their corresponding values are provided, reflecting the effectiveness of the fault detection process.

Fault	Symbol	Value	Δ%
Viscous friction in the x-motion guide	ζ_x	1.005	+0.5
Viscous friction in the y-motion guide	ζ_y	1.001	+0.1
Structural stiffness between the x-motion stage and the base frame in the global y-direction	β_x	*undetectable*	-
Structural stiffness between the y-motion stage and the x-motion stage in the global x-direction	β_y	0.001	-99.9
Actuator for the x-motion	η_x	0.995	-0.5
Actuator for the y-motion	η_y	0.999	-0.1
Actuator for the z-motion	η_z	0.9999	-0.01
External disturbances	γ	0.001	0.1

Table 2: An overview of the fault detectability for each individual fault in the system, controlled by a combined application of feedback and feedforward control, making using the residual generation filter.

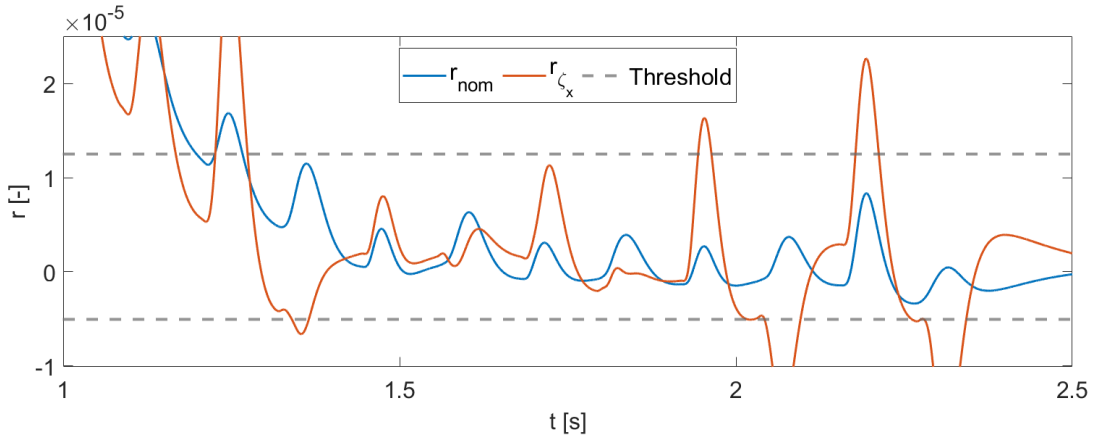
Among the implemented faults, the viscous friction in the x-motion guide and the viscous friction in the y-motion guide are detectable by an increase of 0.5% and 0.1%, respectively. A fault in the structural stiffness related to the x-motion stage, is classified as *undetectable* in Table 2. This implies that the residual generation filter is not sensitive enough to detect this fault reliably. A fault related to the structural stiffness related to the y-motion stage is nearly undetectable, but with a decrease of 99.9%, it is detectable. Regarding the actuators for the x-motion, y-motion, and z-motion, a reduction in motor force of 0.5%, 0.1%, and 0.01%, is detectable. Lastly, the external disturbances in the system have a detectability value of 0.1%, indicating that the residual generation filter is sensitive in detecting such disturbances.

5.4 Fault Detectability for the Nonlinear Model

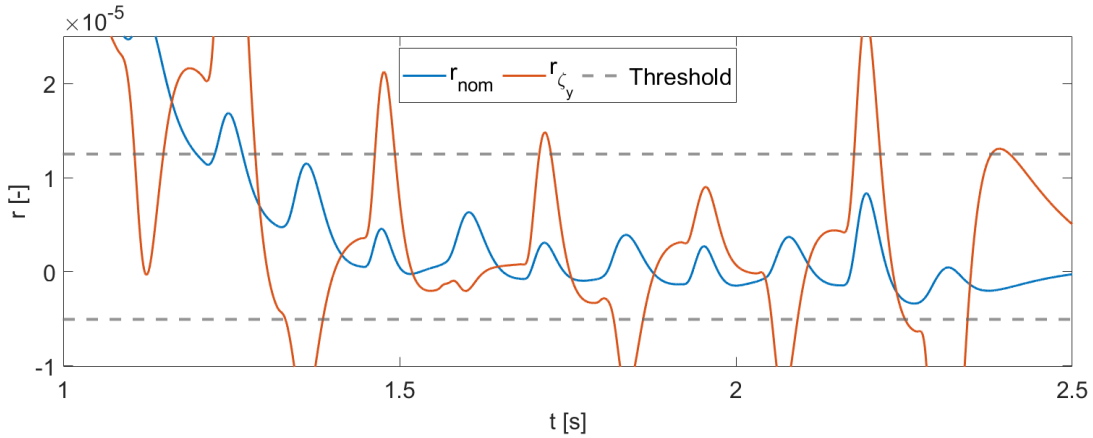
The aim of this section is to explore the detectability of faults within the nonlinear, high-fidelity simulation model, shown in Eq. 1. To achieve this, the same reference signal is applied, composed of the reference trajectories as displayed in Figure 13. A combination of feedback and feedforward control, which is discussed in Section 4.3, is implemented. The simulation configuration remains consistent with that described in Section 5.3.

5.4.1 Generation and Analysis of Residuals

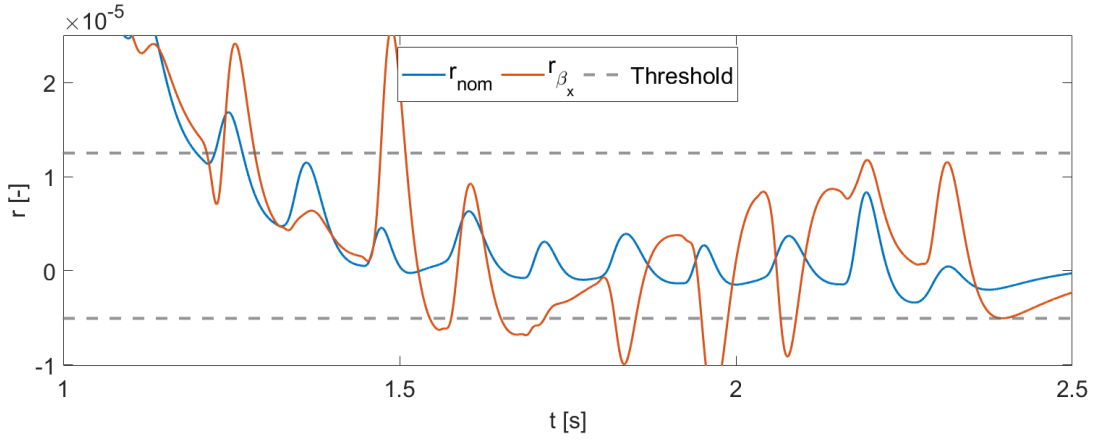
Figure 18 illustrates an analysis of the residuals resulting from the possible faults in the system. The threshold design for this analysis follows the same approach as in the residual analysis in Section 5.2.



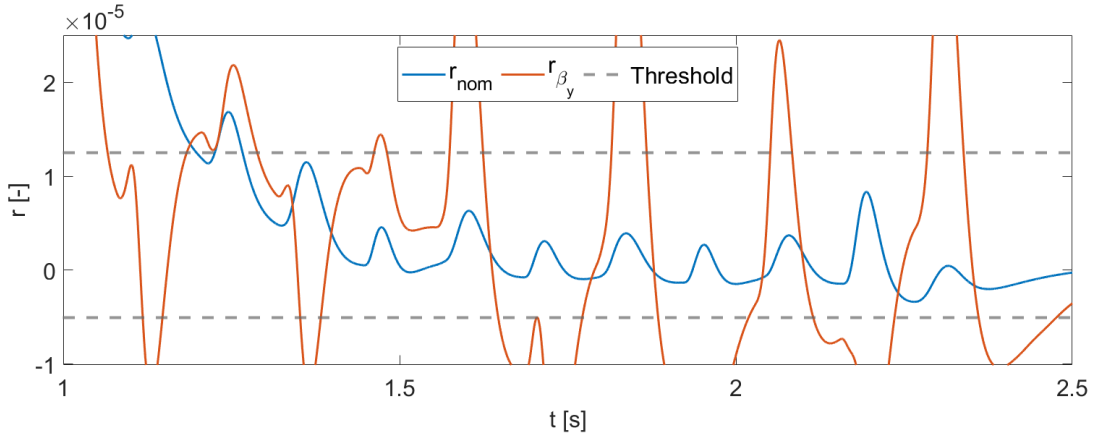
(a) Residual filter output with an 50% increase in friction in the x-motion guide. ($\zeta_x = 1.50$)



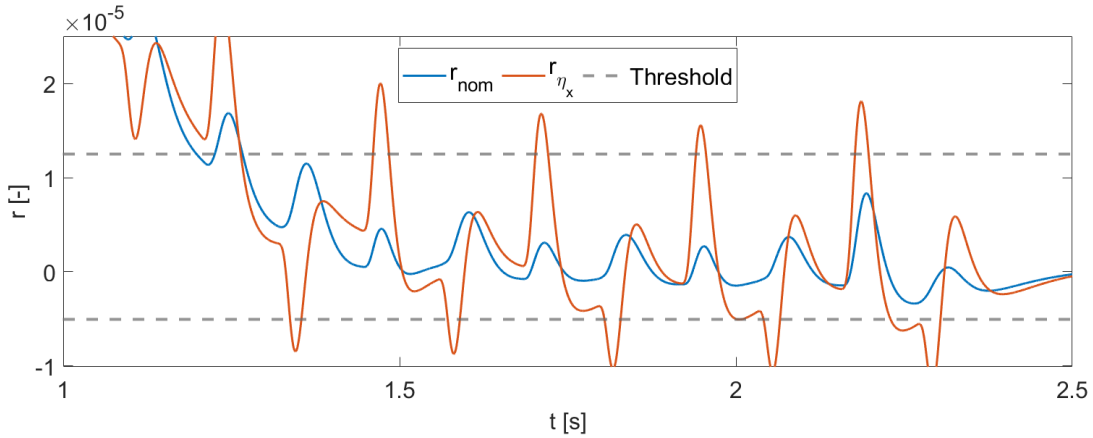
(b) Residual filter output with an 3% increase in friction in the y-motion guide. ($\zeta_y = 1.03$)



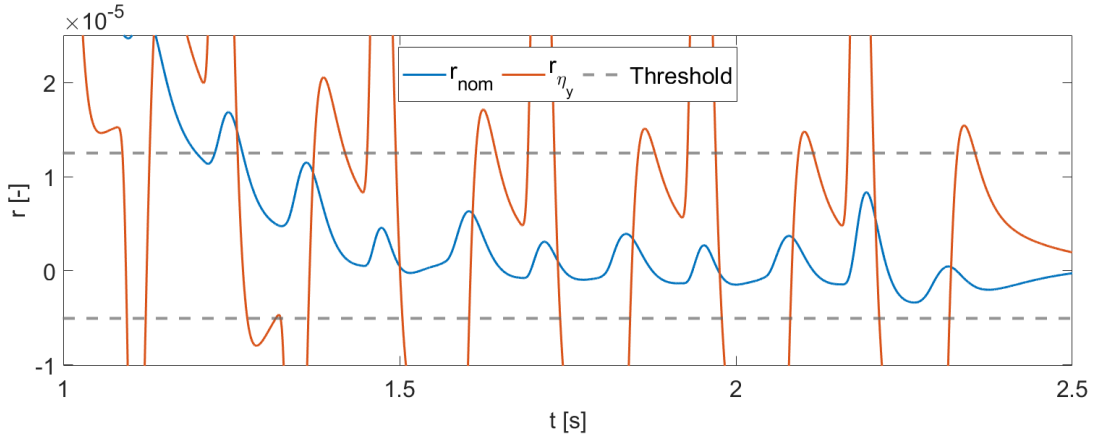
(c) Residual filter output with an 99% decrease in structural stiffness in the x-motion stage. ($\beta_x = 0.01$)



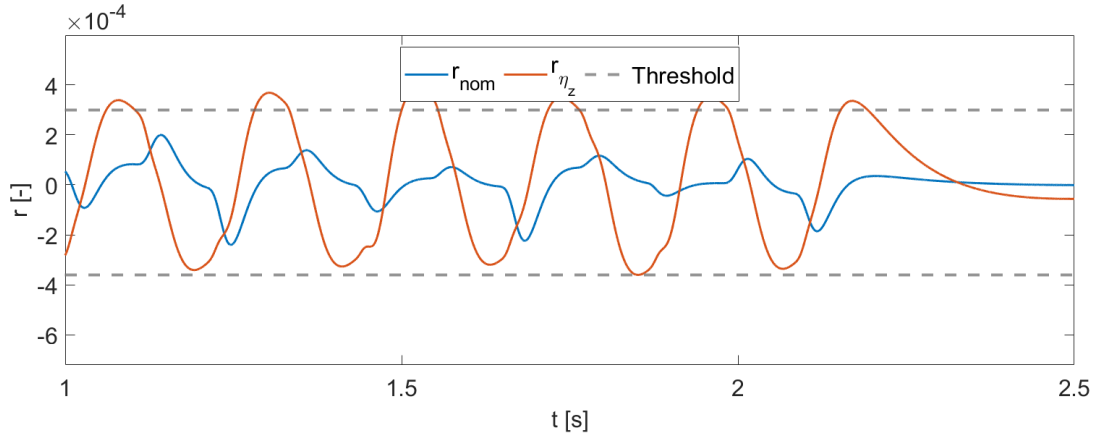
(d) Residual filter output with an 99% decrease in structural stiffness in the y-motion stage. ($\beta_y = 0.01$)



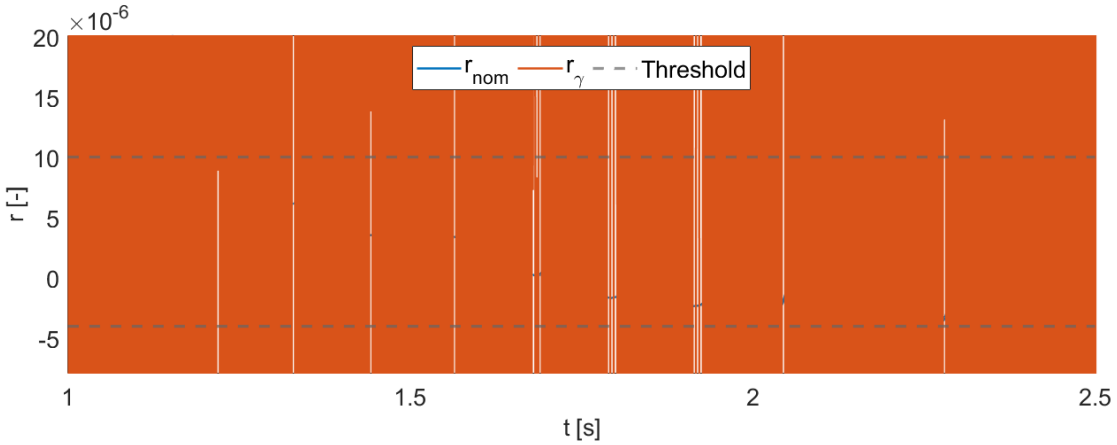
(e) Residual filter output with an 40% decrease of actuator force the x-motion. ($\eta_x = 0.60$)



(f) Residual filter output with an 35% decrease of actuator force the y-motion. ($\eta_y = 0.65$)



(g) Residual filter output with an 20% decrease of actuator force the z-motion. ($\eta_z = 0.80$)



(h) Residual filter output with 1% External disturbances present in the system. ($\gamma = 0.01$)

Figure 18: An overview of the fault detectability for each individual fault in the system, controlled by a combined application of feedback and feedforward control, utilizing the residual generation filter during the simulations with the nonlinear model.

5.4.2 Evaluation of Fault Detection Performance

Following the generation and analysis of residuals, the performance of fault detection is quantified in Table 3. The table provides an overview of the detectability of various faults using the residual generation filter for the nonlinear simulation model.

Fault	Symbol	Value	Δ%
Viscous friction in the x-motion guide	ζ_x	1.50	+50
Viscous friction in the y-motion guide	ζ_y	1.03	+3.0
Structural stiffness between the x-motion stage and the base frame in the global y-direction	β_x	*undetectable*	-
Structural stiffness between the y-motion stage and the x-motion stage in the global x-direction	β_y	*undetectable*	-
Actuator for the x-motion	η_x	0.60	-0.4
Actuator for the y-motion	η_y	0.65	-0.35
Actuator for the z-motion	η_z	0.80	-0.20
External disturbances	γ	0.001	0.1

Table 3: An overview of the fault detectability using the residual generation filter for the nonlinear simulation model for each individual fault implemented in the system.

A comprehensive simulation has been conducted for the nonlinear model using a specific calibration trajectory. The detailed results of this simulation can be found in Appendix A.6.

5.5 Evaluation of Fault Estimation Performance

This section presents a detailed assessment of the fault estimation process which was developed and outlined in Section 3.3. The detectable faults, as specified in Table 2, are incorporated into the system. The fault estimation results are graphically represented in Figure 19. Section 4.2 provides the simulation settings used in the simulations discussed in this section.

For the purpose of fault estimation, both viscous friction faults (ζ_x and ζ_y) and actuator faults (η_x, η_y, and η_z) are taken into account. The implemented values of the incorporated faults are denoted by ζ_i^{real} for $i \in x, y$ and η_i^{real} for $i \in x, y, z$. The approximations of these faults are symbolized by ζ_i^{est} for $i \in x, y$ and η_i^{est} for $i \in x, y, z$. Lastly, the average estimates of the faults are represented by ζ_i^{avg} for $i \in x, y$ and η_i^{avg} for $i \in x, y, z$.

Figure 19 illustrates the estimation of considerable faults through a residual analysis, utilizing the fault estimation framework. The results provide insights into the effectiveness of the fault estimation process for the implemented faults. The estimation errors during the simulation are shown in Figure 20.

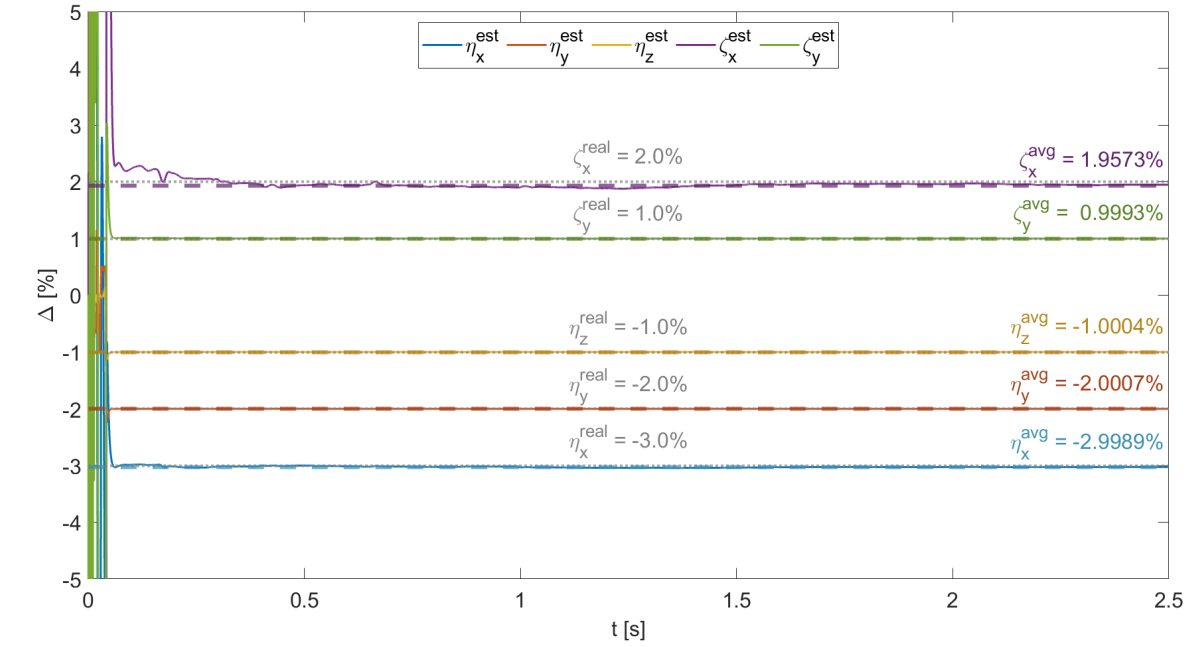


Figure 19: Estimation of considerable faults to evaluate the fault estimation framework discussed in Section 4.2.

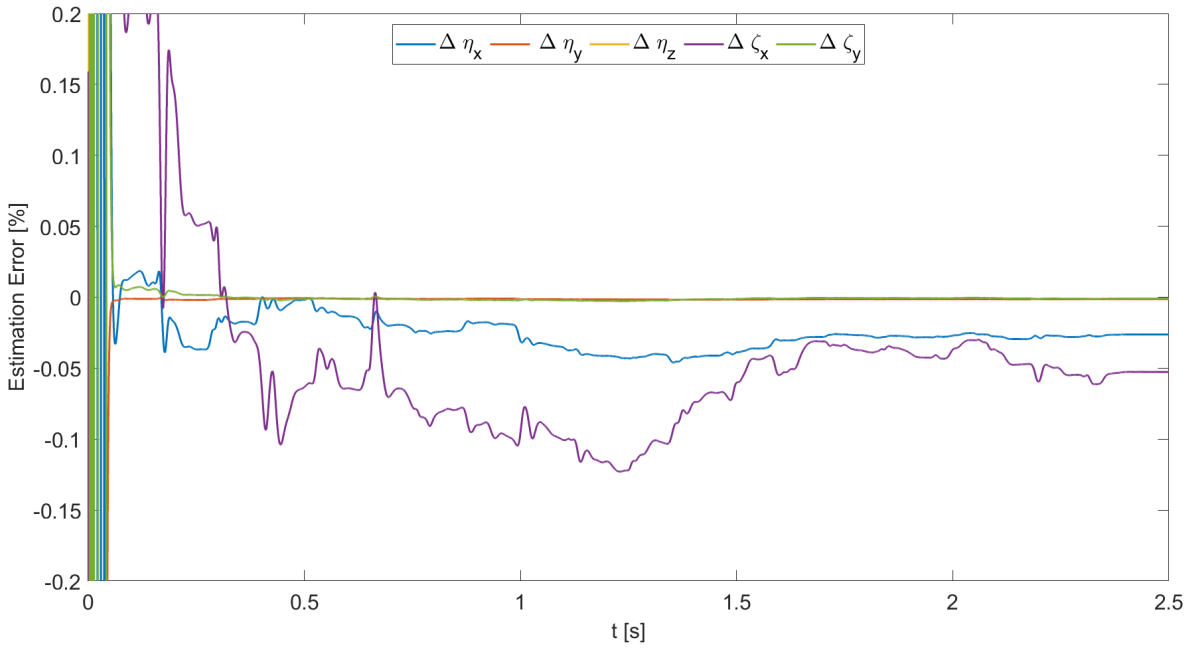


Figure 20: Graphical representation of the estimation errors in the results corresponding to Figure 19.

Finally, the performance of the fault estimation framework is tested with the smallest detectable faults, as shown in Table 2. For clarity in visualization, a minor increase in the

y-actuator fault is implemented. The corresponding fault estimations and the estimation errors are represented in Figures 21 and 22, respectively.

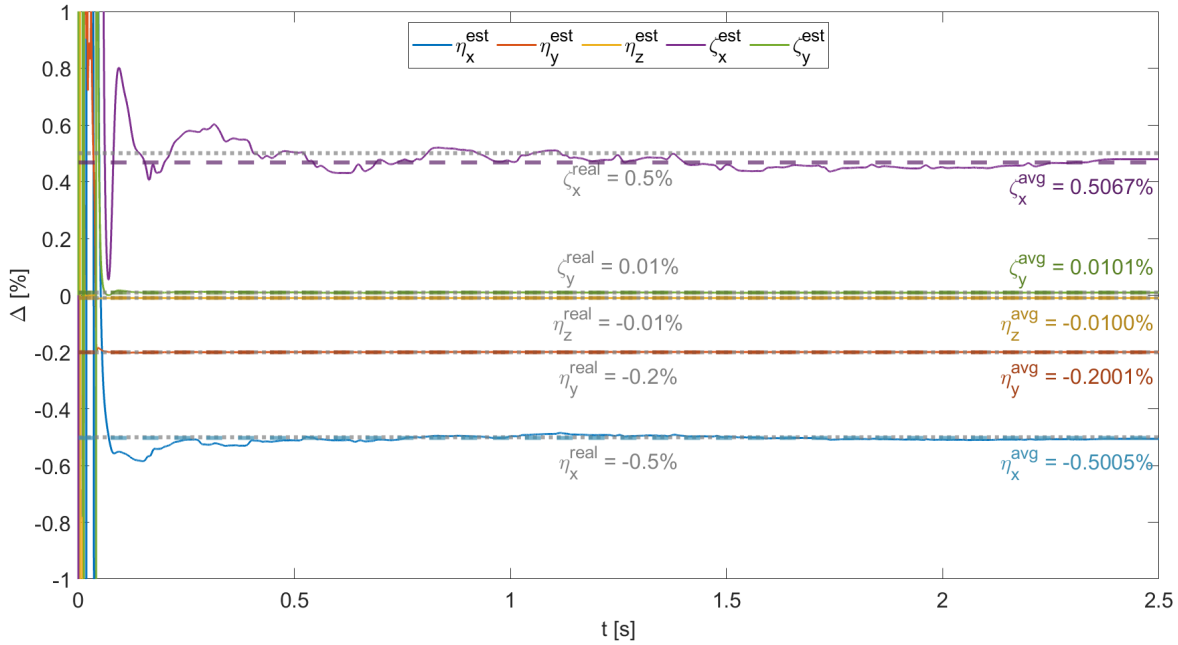


Figure 21: Estimation of smallest detectable faults, as shown in Table 2, based on the method discussed in Section 4.2.

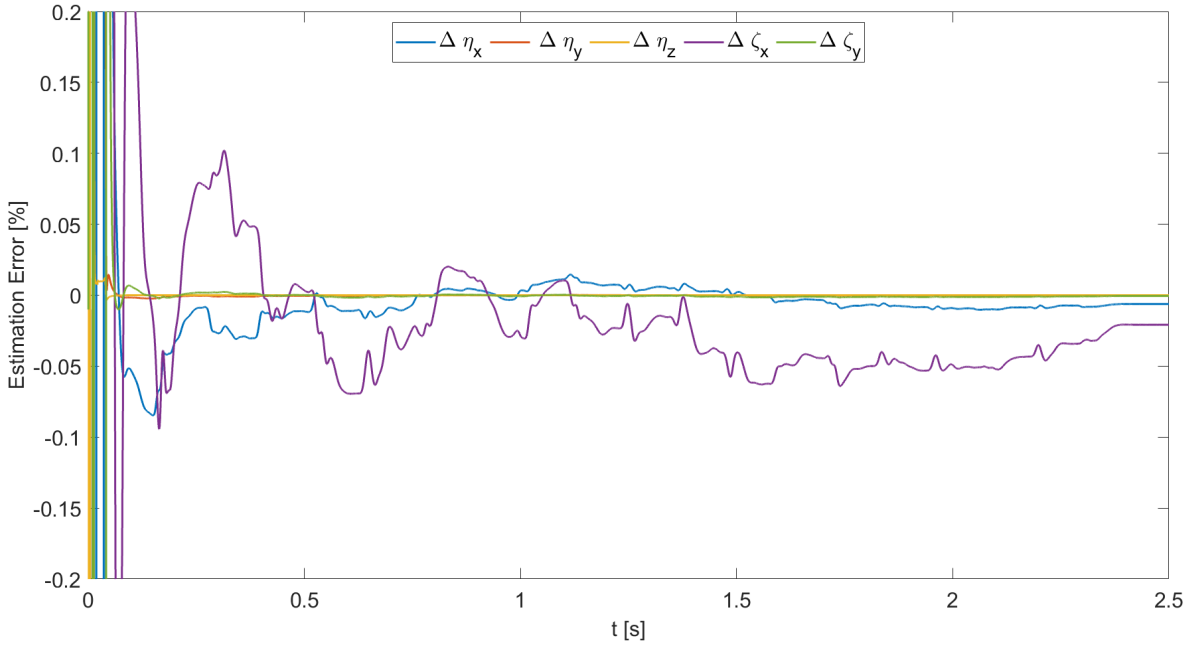


Figure 22: Graphical representation of the estimation errors from the simulation results corresponding to Figure 21.

5.6 Analysis and Estimation of Disturbances

The presence of external disturbances and an amplified level of measurement noise can be identified through PSD analysis performed on the residual signal, as elaborated in Section 3.3. The residual signal is acquired with the same approach implemented in Section 5.3. However, an adjustment in the denominator of the residual filter (Eq. 33) is necessary to ensure that the high-frequency content in the signal remains unattenuated. Furthermore, applying a high-pass filter (Eq. 34) of adequate order attenuates the frequencies generated by the other faults, provided that its cutoff frequency remains lower than the lowest frequency disturbances intended to be detected or estimated.

5.6.1 External Disturbances

Figure 23 presents the PSD plot of the residual signal $r_{dist}(t)$ after undergoing high-pass filtering. The plot visually illustrates the distribution of power across different frequencies in the residual signal. It includes a comparison between the nominal case and a scenario where the system is subject to 1% of external disturbances during the execution of the reference tracking task outlined in Figure 13. The resulting analysis shows peak values at the frequencies of 100, 300, and 750 Hz. These frequencies correlate with the external disturbance signal, as defined in Section 2.3.

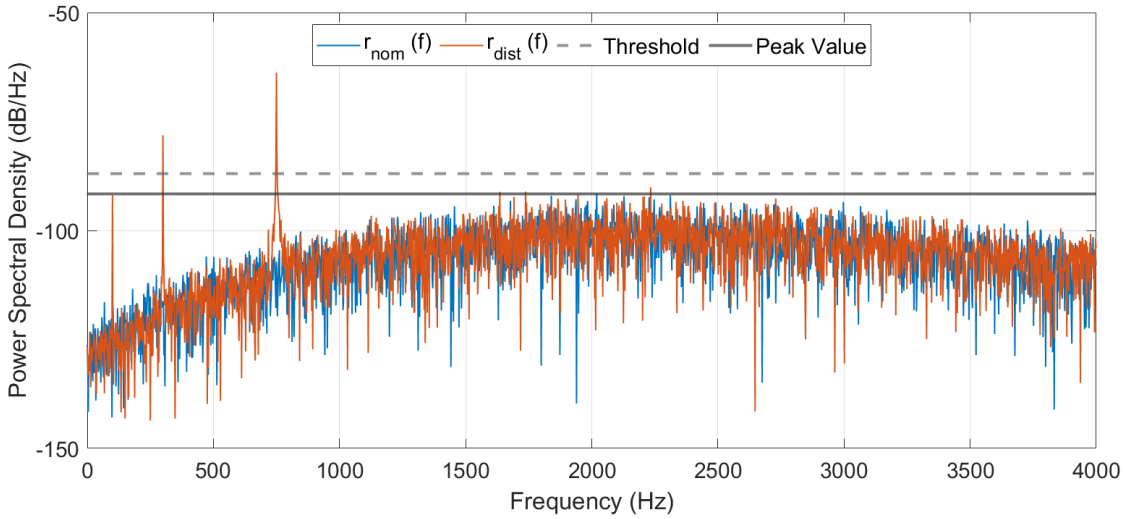


Figure 23: PSD analysis of the residual signal under the influence of 1% external disturbances impacting the system.

5.6.2 Measurement Noise

Measurement noise, represented by Band-Limited Gaussian white noise, has been consistently integrated into all the simulations conducted throughout this study. Figure 24

visualizes the PSD analysis of the residual signal generated with the nominal system, compared with a scenario where the measurement noise is increased by a factor of ten.

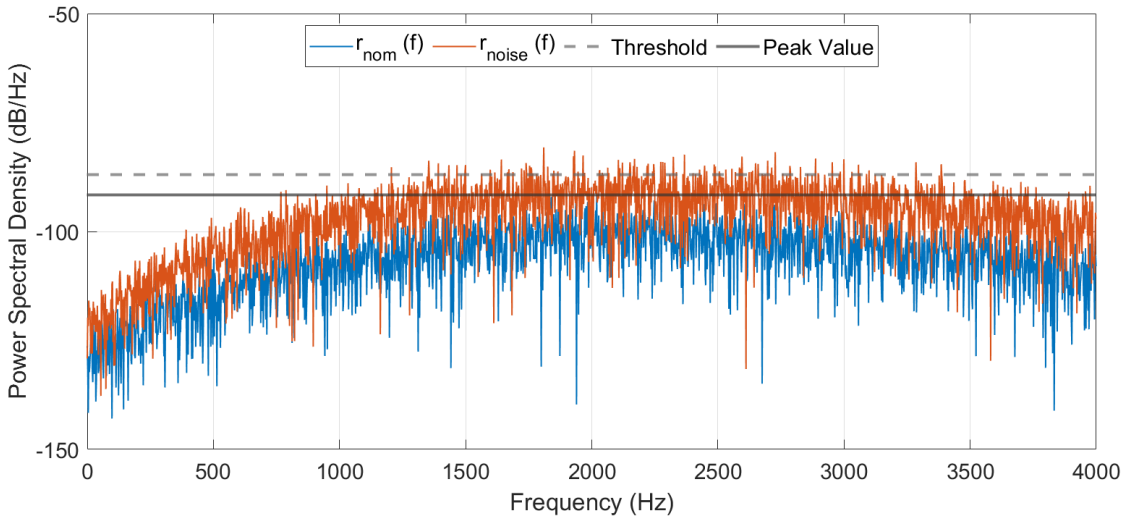


Figure 24: PSD analysis of the residual signal depicting amplified measurement noise.

5.7 Analysis of the Results

This section evaluates the different results obtained during this study. The impact of diverse control strategies on residual generation for fault detection, the detectability of faults, threshold design, and the estimation of faults and disturbances.

The reference signal used during the simulations is emulating realistic wire bonder machine movements. Even though z-motion movements are usually isolated, the system is tested with a variety of isolated and combined trajectories, yielding consistent results across all simulations.

For the z-motion, the feedback control, could be further improved. However, the best performance in terms of reference tracking was observed with a hybrid of feedback and feedforward control and this control structure is used throughout all simulations. The variations in residuals generated by a system under feedback, feedforward, or a combination of both control strategies are negligible which shows that the residual generation is relatively robust against closed-loop control.

For the linear model, as highlighted in Section 2.3, viscous friction can potentially increase by 20%. Nevertheless, the friction was successfully detected with an increase of 0.5% in the x-motion guide and 0.1% in the y-motion guide. The force exerted by the actuators can reduce by up to 3%, and a reduction of 0.5% for the x-actuator, 0.1% for

the y-actuator and 0.01% for the z-actuator is detected. The external disturbances with amplitudes less than 0.1% of the disturbance signal have been successfully detected. Furthermore, the methodology employed for threshold design yielded no false alerts during all the simulations.

The structural stiffness faults are not detectable. Despite the system's observability, comparing the system with structural stiffness faults and the nominal system by Bode analysis (Appendix A.5) reveals a minimal magnitude change across all input and output relations. This may be attributed to the effect of structural stiffness on the parasitic DoFs within the machine, as discussed in Section 2.3. Detection of changes in the parasitic DoFs might be challenging since they are not measured directly, demonstrating its effectiveness.

The research presented in [24] applied a data-driven fault diagnosis approach to the decoupled linear simulation model of the wire bonder shown in Section 2.2.3. This method involved the use of time and frequency features for fault detection, followed by the implementation of a Support Vector Machine (SVM) for fault isolation. All simulations were designed such that the fault magnitude incrementally increased throughout the simulation. The friction in the motion guides experienced an increase that varied between 8% and 300%. The resonance frequency, which is directly linked to the structural stiffness faults, demonstrated a decrease from 900 Hz to 800 Hz. It's important to note that these faults are considered *non-comparable* because they are modeled using the stiffness parameters of the nominal DoFs, not those associated with the parasitic DoFs related to the corresponding motion stage. The motor force in the actuators experiences a reduction that varied from 0.4% to 10%. External disturbances are tested with 0.04-1% of the maximum amplitude. The most significant result for the data-driven method is an accuracy of 98.4% for fault detection and 93% for fault isolation. A comparative analysis of the performance between the method proposed in this study and the data-driven solution from [24] is presented in Table 4.

Fault	Symbol	Residual-Based	Δ %	Data-Driven	Δ %
Friction in the x-motion guide	ζ_x	1.005	+0.5	1.08-3.00	+ 8-300
Friction in the y-motion guide	ζ_y	1.001	+0.1	1.08-3.00	+ 8-300
Structural stiffness between the x-motion stage and the base frame in the global y-direction	β_x	*non-comparable*	-	*non-comparable*	-
Structural stiffness between the y-motion stage and the x-motion stage in the global x-direction	β_y	*non-comparable*	-	*non-comparable*	-
Actuator for the x-motion	η_x	0.995	-0.5	0.996-0.9	- 0.4-10
Actuator for the y-motion	η_y	0.999	-0.1	0.996-0.9	- 0.4-10
Actuator for the z-motion	η_z	0.9999	-0.01	0.996-0.9	- 0.4-10
External disturbances	γ	0.001	0.1	0.0004-0.01	0.04-1

Table 4: A comparison between the performance in fault detection and estimation for the data-driven method implemented in [24] and the direct residual-based methodology.

The fault detectability for the nonlinear EoM, as mentioned in Section 5.4, does not meet the requirements outlined in Section 2.3. Nevertheless, this limitation can be addressed by employing a method that is discussed in Section 6. This method focuses on robustifying the residual signals against model mismatch, thereby enhancing fault detectability [19].

The fault estimation framework for the linear model proposed in this study demonstrates promising results, illustrating successful fault estimation with the proposed methodology. As part of future work, the fault estimation for the nonlinear model can be assessed once the aforementioned robustification method has been implemented.

The presence of external disturbances and increased levels of noise are successfully detected. Although the performance for the detection and estimation of structural stiffness faults lacks comparability with previous work, the precision of detecting and estimating faults related to the actuators, friction, and external disturbances in the high-fidelity linear simulation model is promising.

6 Conclusion

This research presents a comprehensive design and implementation of a fault diagnosis filter for a high-fidelity simulation model of the AB383 wire bonder, a high-precision motion and positioning system. By integrating a model-based methodology with residual generation and regression techniques, the study successfully addresses the detection, isolation, and estimation of faults within the system. Furthermore, the research findings highlight the potential of the power spectral density analysis for residual signals in identifying unknown external disturbances and elevated noise levels.

The reference signal used during the tracking task consisted of several setpoints which are representative of typical realistic movements for the wire bonder machine. The control design, provided by ASMPT Ltd., can be further enhanced to improve the tracking results, but for the performance evaluation of the fault diagnosis filter, the implemented control structure is considered sufficient. The differences in the residuals generated by either feedback, feedforward, or a hybrid control structure are negligible, concluding that the diagnosis performance is relatively independent of the control structure used.

All the parametric faults, except for structural stiffness faults, are detectable to a relatively small deviation in system parameters from the nominal system. Similar performance is obtained for the residual regression methodology utilized in the fault estimation framework. The spectral density analysis of the disturbance residual signal, designed to be sensitive to high-frequency signal components, demonstrates the possibilities in the estimation of elevated levels of noise and unknown external disturbances. In conclusion, the fault diagnosis performance is significantly improved in terms of accuracy with respect to previous work.

The main strength of the model-based methodology is the performance in terms of fault detection and estimation accuracy. The frequency characteristics of the external disturbances are considered known throughout this work, but the spectral density analysis also allows the assessment of unknown disturbances, such as elevated levels of noise or disturbance signals containing specific frequencies. The primary limitation lies in the necessity for system modeling, which might be unavailable in some applications.

To improve fault detectability for the nonlinear EoM, and potentially implement the fault estimation framework to the nonlinear EoM, a residual robustification method can be applied discussed in [19]. This method allows for the creation of an abstract linear model, capable of producing residuals robust with respect to nonlinearities and model mismatch.

This approach has the potential to improve the performance of nonlinear simulations and practical applications involving the physical wire bonder. It offers a viable solution to the challenges associated with achieving an accurate model of complex systems.

Moreover, further development and implementation of the residual generation method in the actual system could establish the fault diagnosis performance for the AB383 wire bonder. This work could include the implementation of various different disturbance signals to verify the diagnosis performance in different machine environments. The real-world implementation can also include the verification of the threshold design strategy.

To solve possible false detection alerts, a possible extension could be the analysis of the integral of the residual signal. Most inherent faults in high-precision motion and positioning systems gradually affect the system while disturbances can change more frequently in the machine environment.

In this study, external disturbances are considered faults whose detection is necessary. Consequently, they are not included in the D_d and B_d matrices within the DAE framework to maintain the observability of the disturbance faults. Nevertheless, it might be possible to employ the D_d and B_d matrices in the design of a filter that remains robust against the effect of disturbance dynamics.

Another possible direction for future research is the design of a fault-tolerant controller. In Figure 25, the framework for fault-tolerant control is provided. This method evaluates the health of the system and re-designs the controller based on this diagnosis to potentially enhance the system's performance.

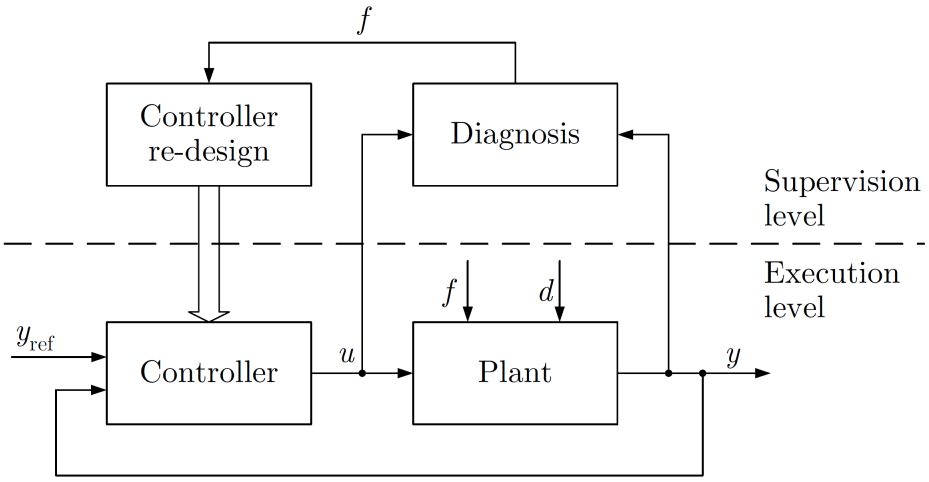


Figure 25: The architecture used in fault-tolerant control [4].

In summary, the main contributions of this study can be summarized as follows: the development of a model-based fault estimation framework for a high-fidelity simulation model, the introduction of a systematic fault representation, and the implementation of a technique for identifying and quantifying external disturbances using power spectral density analysis. This work demonstrates the potential to improve maintenance strategies and reduce downtime for complex systems like the AB383 wire bonder. Thus, setting the stage for further research and practical applications.

A Appendix

A.1 Implementation of Diverse Fault Conditions

To illustrate the effects of various faults, the trajectory from Section 5 is simulated for both the nominal system and each individual faulty system. The open loop architecture, which is employed to create residuals, can be observed in Figure 26. The resulting residuals are displayed in Figure 27. All faulty systems, except for the one with the structural stiffness fault, fail to meet the machine's precision requirement of submicron precision deviations with respect to the nominal system.

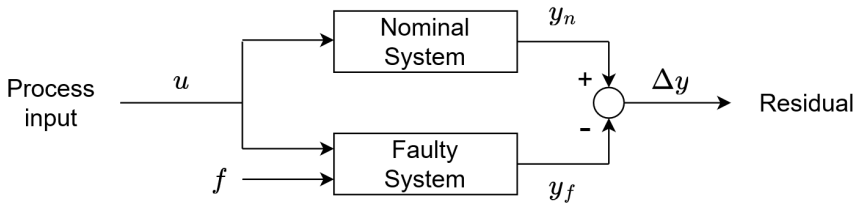
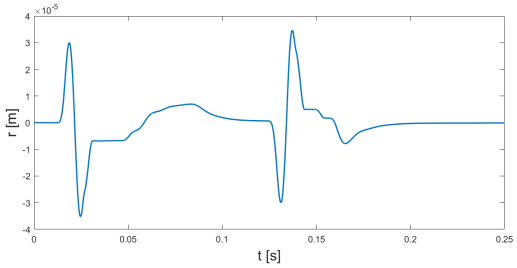
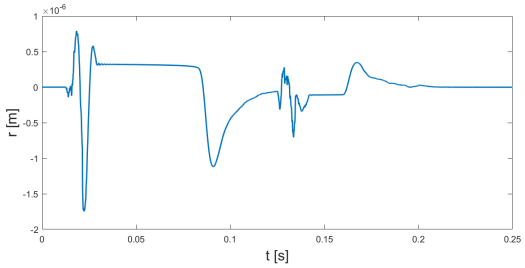


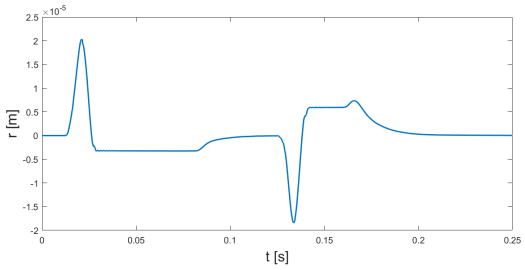
Figure 26: The architecture open loop residual generation.



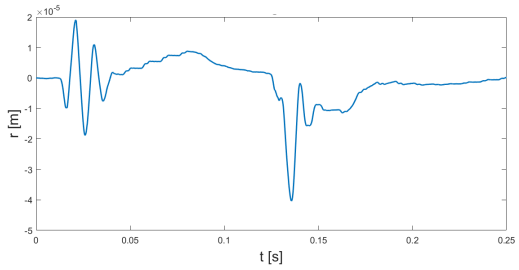
(a) Motor force residual.



(b) Structural stiffness residual.



(c) Friction residual.



(d) External disturbance residual.

Figure 27: Residuals of the difference between the x position during the trajectory in Figure 13 for the nominal behavior and with the different faults implemented.

A.2 MSD Matrices of the Linear EoM

$$
M = \begin{bmatrix}
315.3 & 0 & -6.771 & 3.419 & 0 & 0.01219 & 1.807 & 0 & 0.02761 & 0 \\
0 & 315.3 & -9.093 & 0 & 1.613 & -0.1075 & 0 & 1.807 & -0.009357 & -0.0009811 \\
-6.771 & -9.093 & 29.49 & 0.04354 & -0.1351 & 0.03133 & 0.02761 & -0.04775 & 0.008539 & 2.454e-05 \\
3.419 & 0 & 0.04354 & 3.419 & 0 & 0.01219 & 1.807 & 0 & 0.02761 & 0 \\
0 & 1.613 & -0.1351 & 0 & 1.613 & -0.09809 & 0 & 0 & 0 & 0 \\
0.01219 & -0.1075 & 0.03133 & 0.01219 & -0.09809 & 0.02877 & 0.01104 & -0.009357 & 0.008087 & 3.687e-06 \\
1.807 & 0 & 0.02761 & 1.807 & 0 & 0.01104 & 1.807 & 0 & 0.02761 & 0 \\
0 & 1.807 & -0.04775 & 0 & 0 & -0.009357 & 0 & 1.807 & -0.009357 & -0.0009811 \\
0.02761 & -0.009357 & 0.008539 & 0.02761 & 0 & 0.008087 & 0.02761 & -0.009357 & 0.00834 & 3.687e-06 \\
0 & -0.0009811 & 2.454e-05 & 0 & 0 & 3.687e-06 & 0 & -0.0009811 & 3.687e-06 & 0.0001668
\end{bmatrix}
$$

$$
K = \begin{bmatrix}
1.08e08 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1.08e08 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1.407e07 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2.0e08 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1.479e06 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2.0e08 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.479e06 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.472
\end{bmatrix}
$$

$$
D = \begin{bmatrix}
64000.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 64000.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 8340.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 116.0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1600.0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 12.12 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1600.0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -69.0 & -0.1152 & 0 & 69.0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11.91 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.003145
\end{bmatrix}
$$

$$
b^T = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0.07375 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & -0.00167 & 0 & 1 & 0.03923 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

$$
\Gamma = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
$$

(37)

A.3 Bode Analysis of the Low-Pass Filter

Figure 28 provides a graphical representation of this low-pass filter's response across different frequencies using a Bode plot.

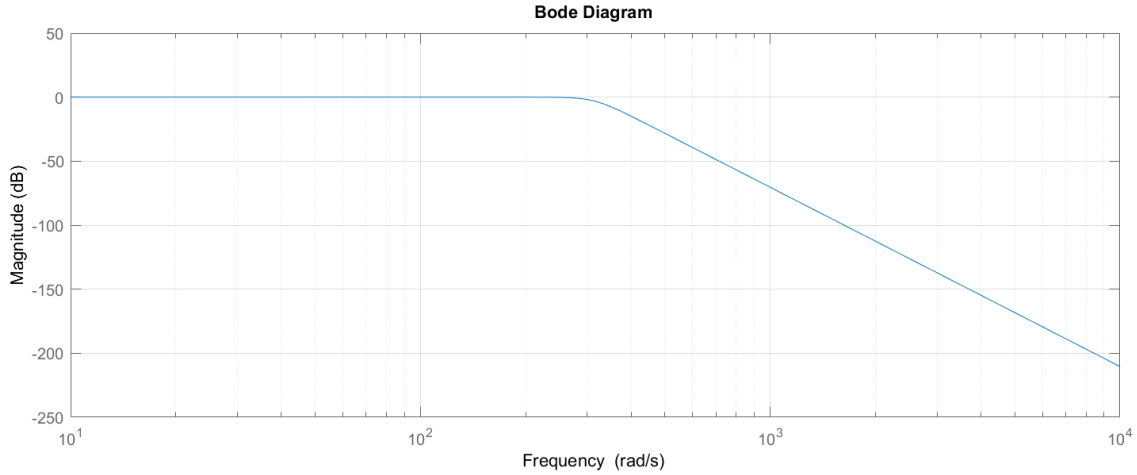


Figure 28: Bode magnitude plot of the low-pass filter, which is used as the foundation for determining the denominator in the residual generation filter.

A.4 Bode Analysis of the High-Pass Filter

The magnitude response of this high-pass filter, depicted using a Bode Magnitude plot, is presented in Figure 29. This graphical representation illustrates how the filter's response varies across different frequencies.

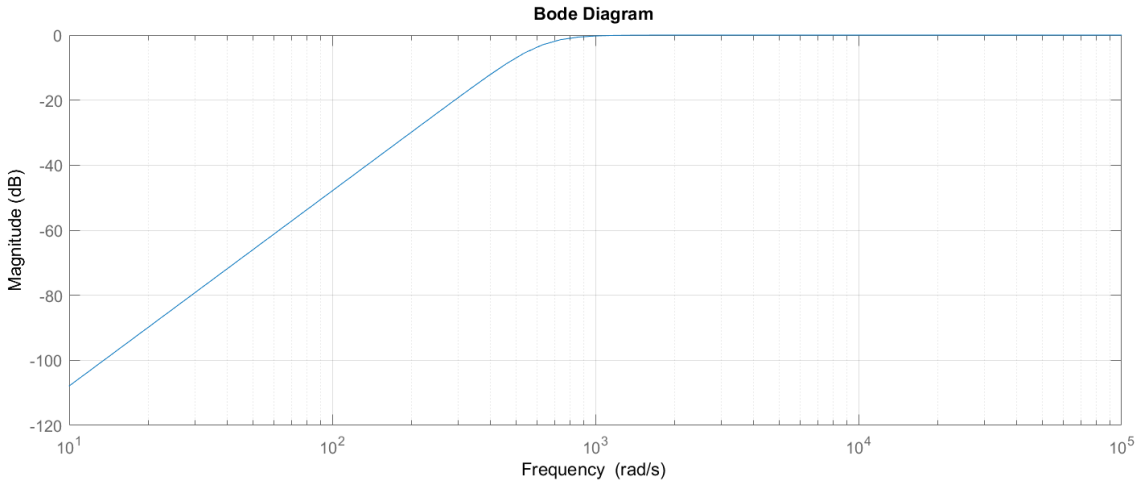


Figure 29: Bode magnitude plot of the implemented high-pass filter in the fault estimation simulations.

A.5 Bode Analysis of Structural Stiffness Faults

Figure 30 presents a Bode analysis, illustrating the difference between the nominal system and the system featuring a reduction of structural stiffness by 99%.

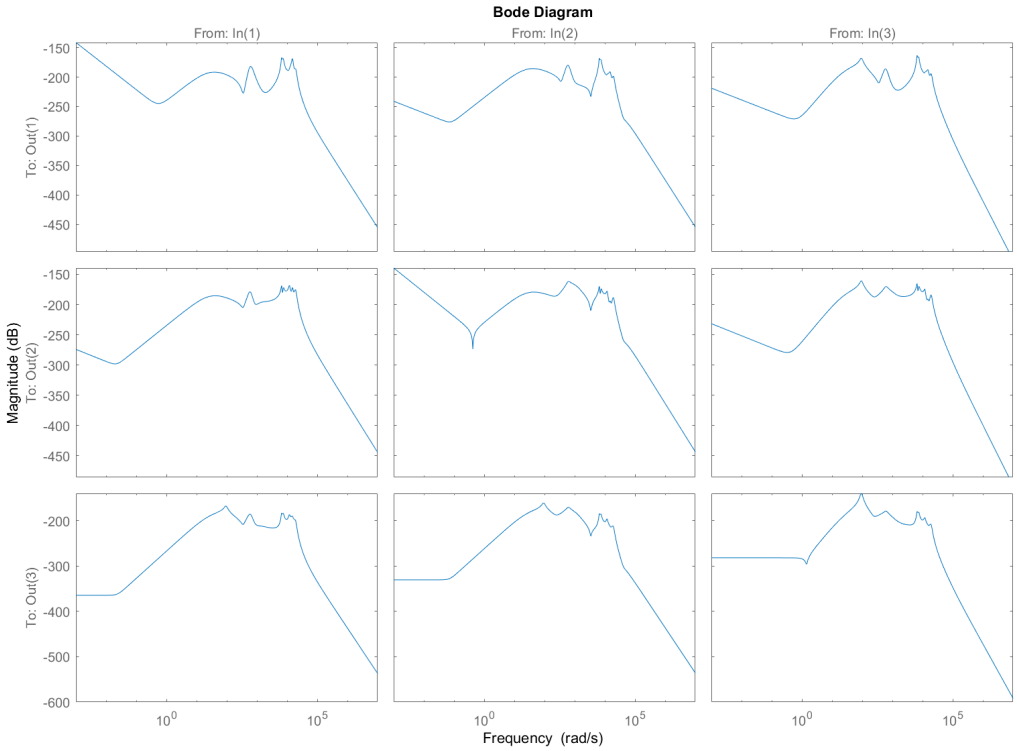


Figure 30: The bode plot of the difference between the nominal system and the system with 99% less structural stiffness.

A.6 Nonlinear Fault Detection with a Detectability Trajectory

The architecture illustrated in Figure 31 was utilized to generate the residual. This included the use of the residual filter and the subtraction of the nominal residual created during the same trajectory. All the other settings are similar to the implementation shown in Section 5.4.

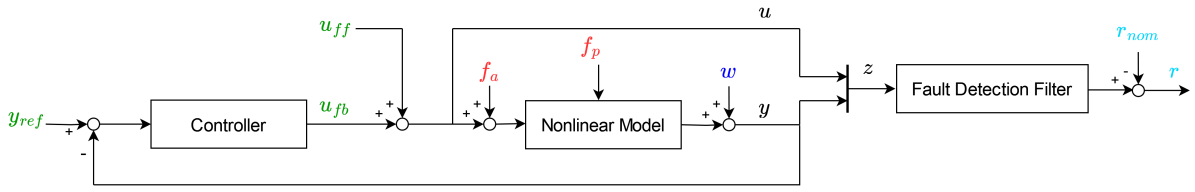
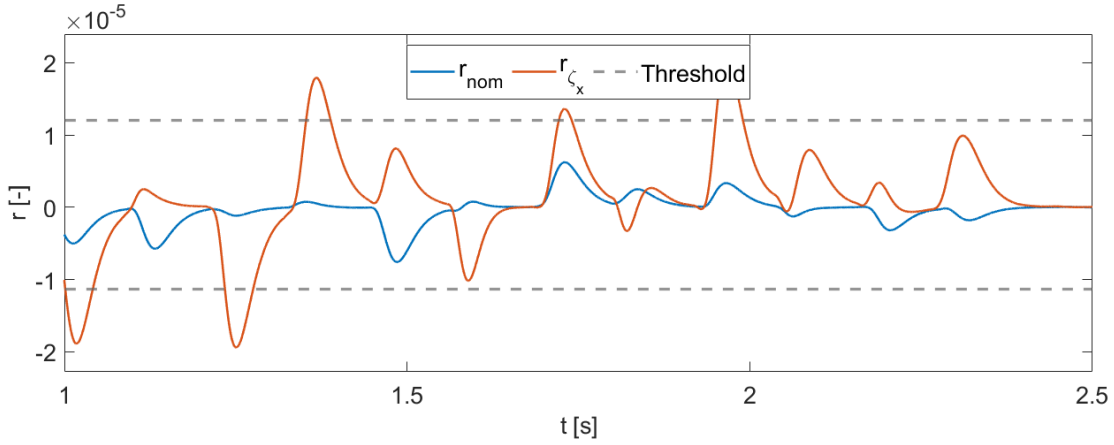
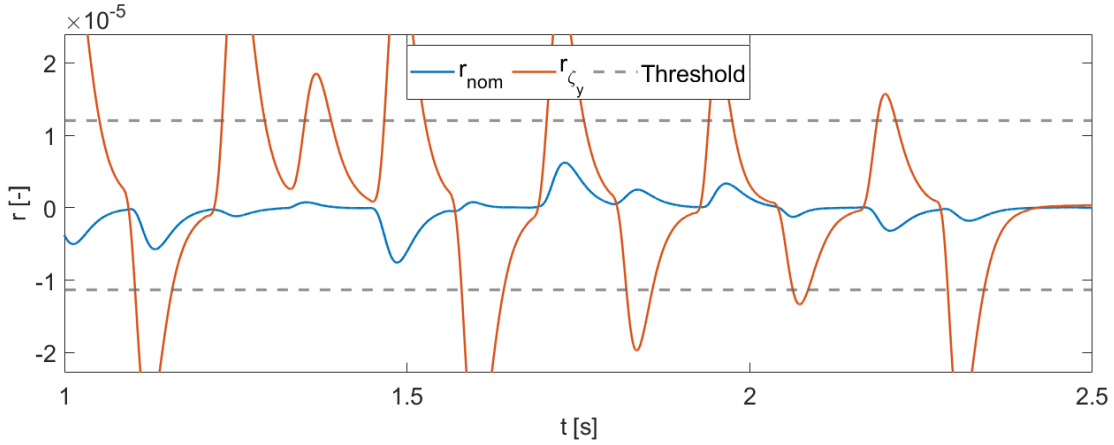


Figure 31: The architecture of the implemented residual filter on the closed loop nonlinear model during the detectability trajectory.

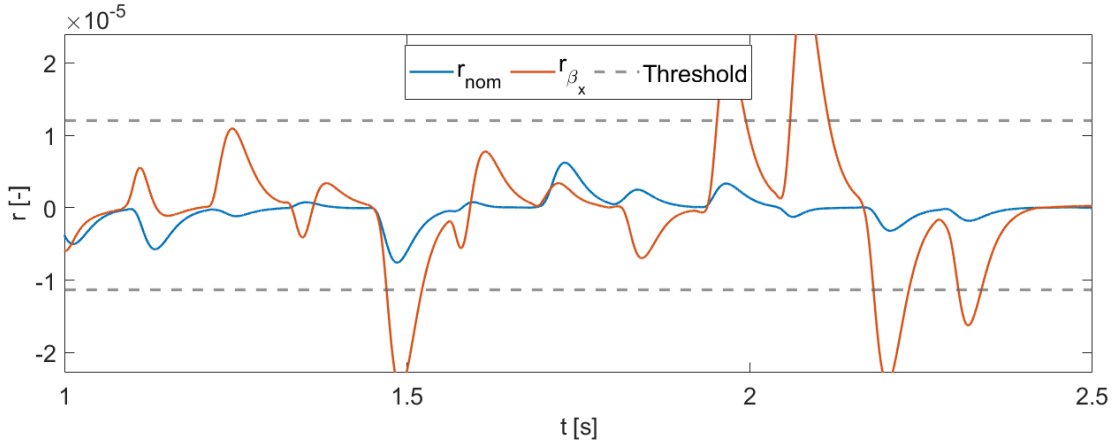
The residual baseline r_{nom} and the thresholds are constructed in a similar method as discussed in Section 5.4.



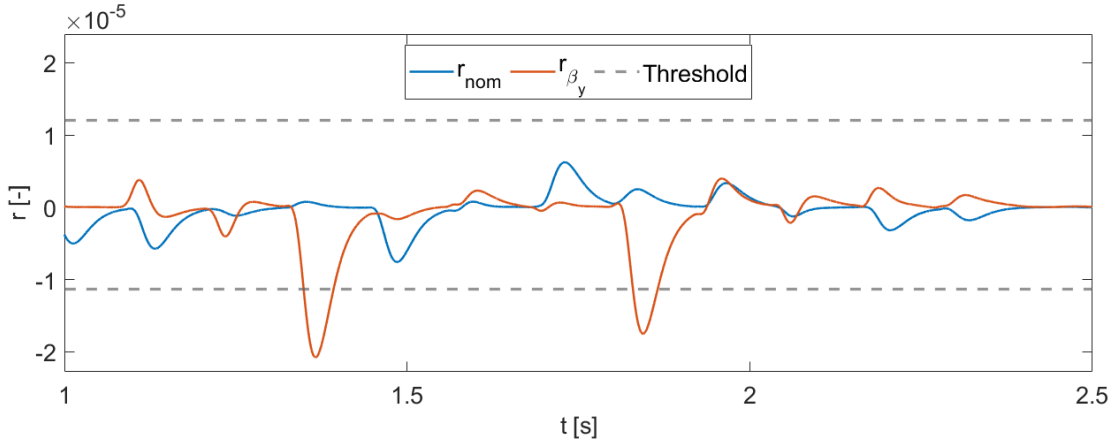
(a) Residual filter output with an 20% increase in friction in the x-motion guide. ($\zeta_x = 1.20$)



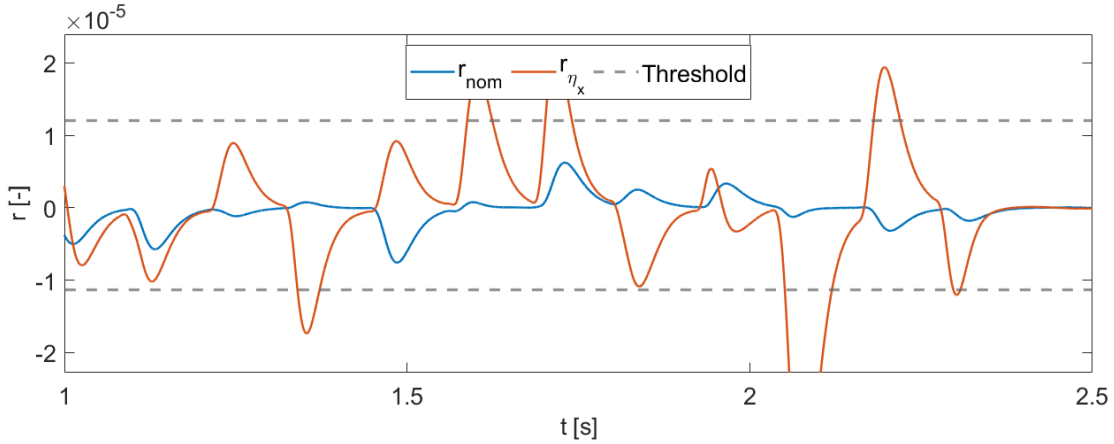
(b) Residual filter output with an 1% increase in friction in the y-motion guide. ($\zeta_y = 1.01$)



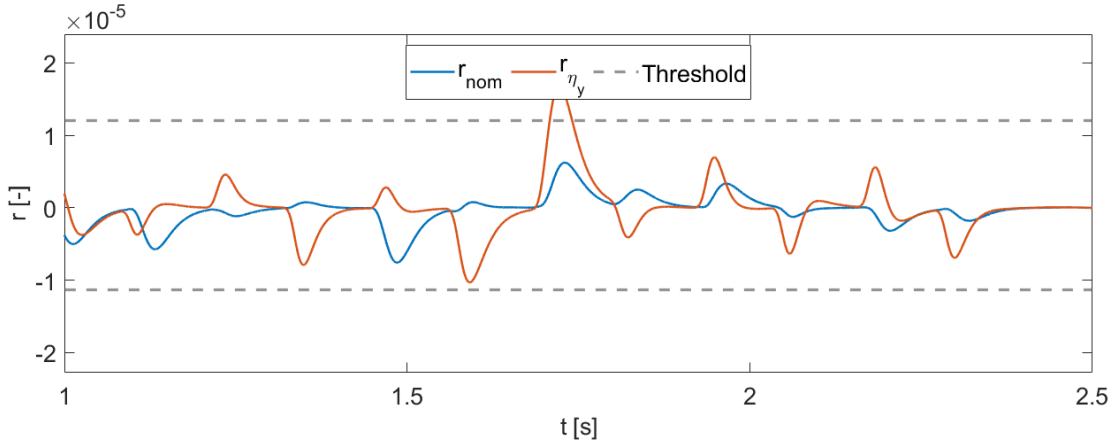
(c) Residual filter output with an 99% decrease in structural stiffness in the x-motion stage. ($\beta_x = 0.01$)



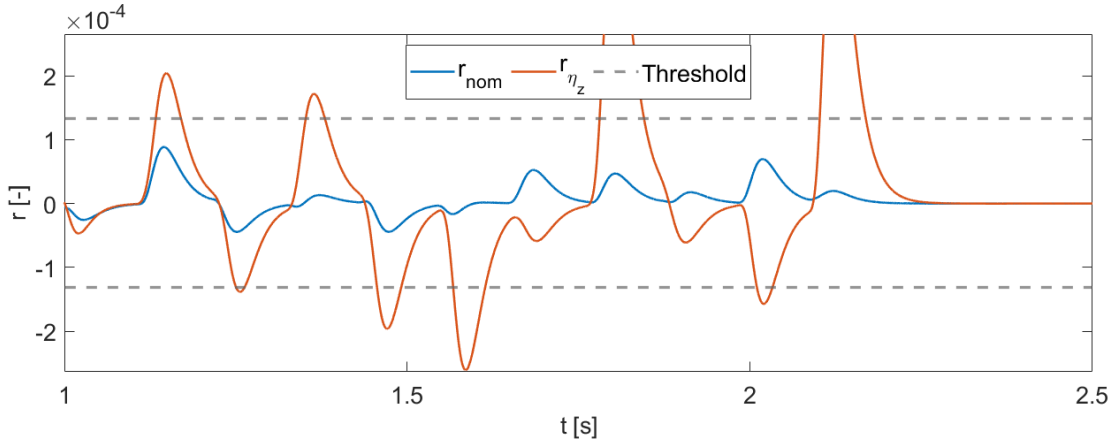
(d) Residual filter output with an 99% decrease in structural stiffness in the y-motion stage. ($\beta_y = 0.01$)



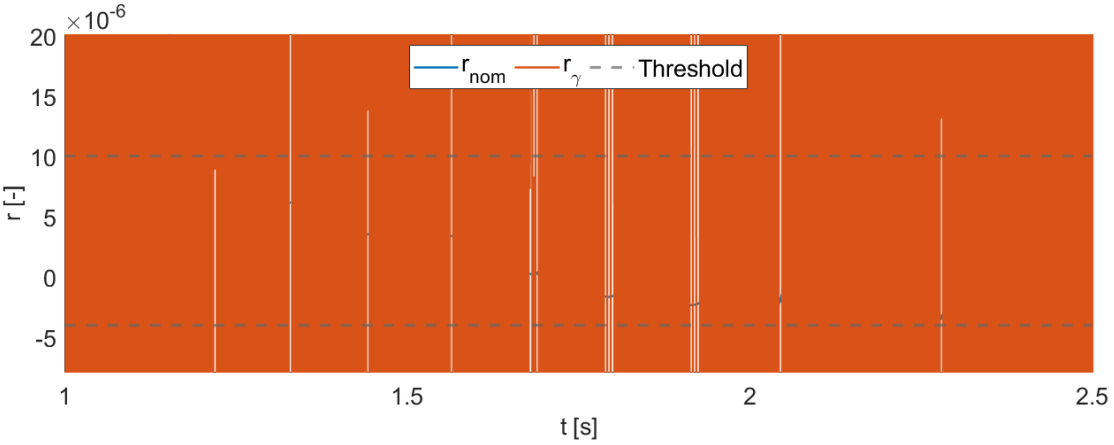
(e) Residual filter output with an 10% decrease of actuator force the x-motion. ($\eta_x = 0.90$)



(f) Residual filter output with an 10% decrease of actuator force the y-motion. ($\eta_y = 0.90$)



(g) Residual filter output with an 2% decrease of actuator force the z-motion. ($\eta_z = 0.98$)



(h) Residual filter output with 1% External disturbances present in the system. ($\gamma = 0.01$)

Figure 32: An overview of the fault detectability using the residual generation filter in combination with a fixed trajectory for the nonlinear simulation model for each individual fault implemented in the system.

Detection Performance

Fault	Symbol	Value	Δ%
Viscous friction in the x-motion guide	ζ_x	1.20	+20
Viscous friction in the y-motion guide	ζ_y	1.01	+1.0
Structural stiffness between the x-motion stage and the base frame in the global y-direction	β_x	0.01	-99
Structural stiffness between the y-motion stage and the x-motion stage in the global x-direction	β_y	0.01	-99
Actuator for the x-motion	η_x	0.90	-0.10
Actuator for the y-motion	η_y	0.98	-0.02
Actuator for the z-motion	η_z	0.98	-0.02
External disturbances	γ	0.001	0.1

Table 5: The fault detectability for the nonlinear simulation model during the detection trajectory for each individual fault implemented in the system.

References

[1] ASM Pacific Technology Ltd. https://www.asmpt.com/. Accessed: May 4, 2023.

[2] S. Beer. Data-driven Axes Decoupling and Motion Control of Bonding Machines. Unpublished internal company document, 2019.

[3] J. S. Bendat and A. G. Piersol. *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons, 2010.

[4] M. Blanke, M. Kinneart, J. Lunze, and M. Staroswiecki. *Diagnosis and fault-tolerant control with 270 figures, 121 examples, 5 application studies, and 36 exercises*. Springer, 2nd edition, 2006.

[5] E. Che Mid and V. Dua. Model-Based Parameter Estimation for Fault Detection Using Multiparametric Programming. *Industrial & Engineering Chemistry Research*, 56(28):8000–8015, 2017.

[6] J. Chen and R. J. Patton. *Robust model-based fault diagnosis for dynamic systems*, volume 3. Springer Science & Business Media, 2012.

[7] G. de Albuquerque Gleizer and P. Esfahani. Active fault estimation of multiplicative faults. In Preparation, 2023.

[8] H. S. de Reij. Fault Diagnosis Methods for High Precision Motion Systems - Literature Review. https://acrobat.adobe.com/id/urn:aaid:sc:EU:0a3af578-66d4-4a0d-a5fd-cb4a7bf34102, 2023.

[9] DIGITAL TWIN Research. ASM digital twin of industrial wire bonder machine. https://www.digital-twin-research.nl/research/use-cases/asm-digital-twin-of-industrial-wire-bonder-machine/. Accessed: June 1, 2023.

[10] P. M. Esfahani and J. Lygeros. A tractable fault detection and isolation approach for nonlinear systems with probabilistic performance. *IEEE Transactions on Automatic Control*, 61(3):633–647, 2015.

[11] E. Frisk and M. Nyberg. Residual generation for fault diagnosis of systems described by general linear differential-algebraic equations. *IFAC Proceedings Volumes*, 35(1):137–142, 2002. 15th IFAC World Congress.

[12] Z. Gao, C. Cecati, and S. Ding. Fault Diagnosis With Knowledge-Based and Hybrid/Active Approaches. *IEEE Transactions on Industrial Electronics*, 62(6):3768–3774, 2015.

[13] Z. Gao, C. Cecati, and S. Ding. Fault Diagnosis With Model-Based and Signal-Based Approaches. *IEEE Transactions on Industrial Electronics*, 62(6):3757–3767, 2015.

[14] J. Gerritsen. A Model with Controller for the AB383 Wire bonder Machine. Unpublished internal company document, 2023.

[15] W. H. Greene. *Econometric Analysis*. Prentice Hall, 5th edition, 2003.

[16] D. Lu and C. Wong. *Materials for advanced packaging*, volume 181. Springer, 2009.

[17] P. Mohajerin Esfahani, M. Vrakopoulou, G. Andersson, and J. Lygeros. A tractable nonlinear fault detection and isolation technique with application to the cyber-physical security of power systems. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 3433–3438, 2012.

[18] A. V. Oppenheim and R. W. Schafer. *Discrete-time signal processing*. Prentice-hall signal processing series. Prentice-Hall, 3rd edition, 2010.

[19] K. Pan, P. Palensky, and P. M. Esfahani. Dynamic Anomaly Detection With High-Fidelity Simulators: A Convex Optimization Approach. *IEEE Transactions on Smart Grid*, 13(2):1500–1515, 2022.

[20] R. Rahimilarki, Z. Gao, A. Zhang, and R. Binns. Robust neural network fault estimation approach for nonlinear dynamic systems with applications to wind turbine systems. *IEEE Transactions on Industrial Informatics*, 15(12):6302–6312, 2019.

[21] D. Rotondo, M. Witczak, V. Puig, F. Nejjari, and M. Pazera. Robust unknown input observer for state and fault estimation in discrete-time takagi–sugeno systems. *International Journal of Systems Science*, 47(14):3409–3424, 2016.

[22] J. Y. Stein. *Digital Signal Processing: A Computer Science Perspective*. Wiley, 2000.

[23] W. Thomson. *Theory of vibration with applications*. CrC Press, 1996.

[24] R. van Bussel. Design and Benchmarking of Generic Machine Health Condition Indicators for a Wirebonder Application. Unpublished internal company document, 2020.

[25] C. van der Ploeg, M. Alirezaei, N. van de Wouw, and P. M. Esfahani. Multiple faults estimation in dynamical systems: Tractable design and performance bounds. *IEEE Transactions on Automatic Control*, 67(9):4916–4923, 2022.

[26] M. Witczak, D. Rotondo, V. Puig, F. Nejjari, and M. Pazera. Fault estimation of wind turbines using combined adaptive and parameter estimation schemes. *International Journal of Adaptive Control and Signal Processing*, 32(4):549–567, 2018.

[27] X. Xu, F. Xiao, and S. Wang. Enhanced chiller sensor fault detection, diagnosis and estimation using wavelet analysis and principal component analysis methods. *Applied Thermal Engineering*, 28(2):226–237, 2008.

[28] J. Zhang, A. K. Swain, and S. K. Nguang. Robust sensor fault estimation scheme for satellite attitude control systems. *Journal of the Franklin Institute*, 350(9):2581–2604, 2013.