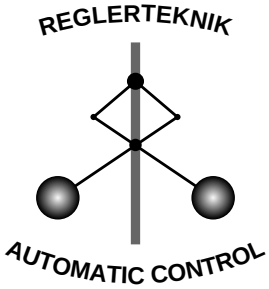
Linköping Studies in Science and Technology
Thesis No. 1258

Statistical Fault Detection with Applications to IMU Disturbances

David Törnqvist



LINKÖPING

Division of Automatic Control
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
http://www.control.isy.liu.se
tornqvist@isy.liu.se

Linköping 2006

This is a Swedish Licentiate's Thesis.
Swedish postgraduate education leads to a Doctor's degree and/or a Licentiate's degree. A Doctor's Degree comprises 160 credits (4 years of full-time studies). A Licentiate's degree comprises 80 credits, of which at least 40 credits constitute a Licentiate's thesis.

Statistical Fault Detection with Applications to IMU Disturbances

Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping
Sweden

ISBN 91-85523-44-5 ISSN 0280-7971 LiU-TEK-LIC-2006:39

Printed by LiU-Tryck, Linköping, Sweden 2006

To Anders, Ann, Lina, Albin and Linnéa

Abstract

This thesis deals with the problem of detecting faults in an environment where the measurements are affected by additive noise. To do this, a residual sensitive to faults is derived and statistical methods are used to distinguish faults from noise. Standard methods for fault detection compare a batch of data with a model of the system using the *generalized likelihood ratio*. Careful treatment of the initial state of the model is quite important, in particular for short batch sizes. One method to handle this is the parity-space method which solves the problem by removing the influence of the initial state using a projection.

In this thesis, the case where prior knowledge about the initial state is available is treated. This can be obtained for example from a *Kalman filter*. Combining the prior estimate with a minimum variance estimate from the data batch results in a smoothed estimate. The influence of the estimated initial state is then removed. It is also shown that removing the influence of the initial state by an estimate from the data batch will result in the parity-space method. To model slowly changing faults, an efficient parameterization using Chebyshev polynomials is given.

The methods described above have been applied to an *Inertial Measurement Unit*, IMU. The IMU usually consists of accelerometers and gyroscopes, but has in this work been extended with a magnetometer. Traditionally, the IMU has been used to estimate position and orientation of airplanes, missiles etc. Recently, the size and cost has decreased making it possible to use IMU:s for applications such as augmented reality and body motion analysis. Since a magnetometer is very sensitive to disturbances from metal, such disturbances have to be detected. Detection of the disturbances makes compensation possible. Another topic covered is the fundamental question of observability for fault inputs. Given a fixed or linearly growing fault, conditions for observability are given.

The measurements from the IMU show that the noise distribution of the sensors can be well approximated with white Gaussian noise. This gives good correspondence between practical and theoretical results when the sensor is kept at rest. The disturbances for the IMU can be approximated using smooth functions with respect to time. Low rank parameterizations can therefore be used to describe the disturbances. The results show that the use of smoothing to obtain the initial state estimate and parameterization of the disturbances improves the detection performance drastically.

Acknowledgments

First of all, I would like to thank my supervisor Professor Fredrik Gustafsson. We have had many fruitful discussions, but I especially admire your way of seeing possibilities in everything and your positive way of thinking. Working hard with a thesis is not always easy, but you know how to give comments that inspires further work. Also my co-supervisor Associate Professor Inger Klein has given me good support, thank you!

I would also like to thank Professor Lennart Ljung for drafting me to the control group and for creating a good atmosphere in the group. Our secretary Ulla Salaneck also deserves my gratitude for always arranging practical matters in a smooth way.

Various parts of this thesis have been proofread by Lic Daniel Axehill, Lic Gustaf Hendeby, Jeroen Hol, Lic Per-Johan Nordlund, Dr Thomas Schön and Lic Johan Sjöberg. Your comments have been invaluable! I especially want to thank Gustaf and Thomas for your many comments and insightful ideas.

I would like to thank all the people at the department for creating a good atmosphere and for discussing almost everything during the coffee breaks. There are three people that started at department at the same time as me, Daniel, Gustaf and Johan. You have become really good friends and I really appreciate both our research discussions and spare-time activities. Spending hours discussing audio cables really brings a new dimension into life. :-) Also Henrik, Jeroen and Johanna have helped me to kill some spare time, I enjoy your company.

Our TEX-guru Gustaf Hendeby also deserves appreciation for creating the template this thesis is formatted in and for helping out with LATEX-related questions.

This work was sponsored by Center for Industrial Information Technology, CENIIT, and The Excellence Center in Computer Science and Systems Engineering in Linköping, ECSEL, which are hereby gratefully acknowledged.

My warmest thanks goes to my family for always believing in what I am doing. The support from you is very important to me!

Linköping, May 2006
David Törnqvist

Contents

1

Introduction

NAVIGATION HAS ALWAYS BEEN a challenge to man. Inertial navigation systems, INS, based on Newton's first law, have been developed during the 20th century. The basic components of an INS are gyroscopes and accelerometers for measuring angular rate and linear acceleration. The orientation can be estimated by integrating the gyroscope signals and the position by double-integrating the accelerometer signals. However, the measurements from the sensors contain noise which introduces drift in the estimates. The smaller the noise the smaller the drift. The development until the 1950s made the noise levels low enough to cross the USA with an airplane, relying on an INS. At that time, the sensors were mounted in gimbals (Figure 1.1) which maintained the orientation of the actual sensor. The gimbaled sensors were large and heavy which limited their applicability. During the later part of the 20th century microelectromechanical systems, MEMS, were introduced. This enabled the evolution of small silicon based accelerometers and gyroscopes that could be mounted rigidly onto the moving body, that is, in strapdown configuration. Small units consisting of accelerometers and gyroscopes are referred to as Inertial Measurement Units, IMU.

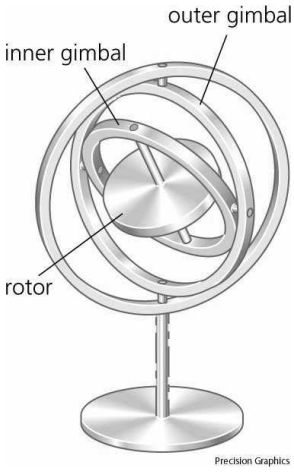


Figure 1.1: *Gimbals*

If the IMU is held still (not accelerated), the orientation can be obtained by integrating the gyroscope signals and the drift can be compensated for by measuring the gravity vector. This approach will still give drift around the gravity vector, which motivates the addition of a magnetometer (electronic compass). Measuring the magnetic field of Earth, a drift-free orientation estimate can be obtained. However, the magnetic field of Earth is very weak and is easily disturbed by magnetic objects. Such disturbances can be detected and compensated for by

using measurements from other sensors. This type of detection problems are treated in this thesis.

The small size and low power consumption of the MEMS-sensors have enabled the use of IMU:s in new areas. An IMU with added magnetometers is used to monitor human motion in Roetenberg (2006). Placing the sensors on different parts of the body, movements can be analyzed. Medical researchers can for example use this information to analyze gait patterns and how the movements are affected by different disorders. Since the analyzed persons are likely to move near metallic objects, detection of magnetic disturbances is essential.

Another application is to estimate position and orientation of a camera. Research about this topic is performed in the EU-founded project Markerless real-time Tracking for Augmented Reality Image Synthesis, MATRIS (MATRIS, Schön, 2006, Hol, 2005). The position and orientation of the camera are important when the image is to be augmented with virtual objects, a practice known as augmented reality. This can for example be directions for how to assemble a car (see Figure 1.2) or virtual marker lines in a sporting event. The position and orientation are today estimated by fusion of image and IMU data. If reliable disturbance detection is used for a magnetometer, the accuracy of the estimate would be improved by the addition of this sensor.

Figure 1.2: *Augmented reality is used to give instructions for the assembling of a car. Courtesy of Frauenhofer IGD.*

There are now cell phones equipped with MEMS accelerometers and magnetometers, see for example Figure 1.3. The accelerometers can be used for movement detection. If the cell phone is not moving, there is no need to search for new base stations and power is saved. If both magnetometer and accelerometer are used, the orientation of the phone can be determined. This opens up the possibility to control the menus by tilting the phone or play advanced games. As cell phones today include both camera and high resolution color screen, only the processor limits the ability for augmented reality applications. Imagine walking around the pyramids of Egypt, the phone downloads information about the site since it knows its position. When pointing the camera against the pyramids, virtual work-

ers from the ancient Egypt working with the construction of the pyramids are shown in the screen. A story about how the pyramids were built can also be heard in the headphones.

Figure 1.3: *Cell phone from Samsung with built in accelerometers and magnetometers.*

1.1 Problem Formulation

Using the IMU to estimate orientation, it is essential to detect movement and magnetometer disturbances to not deteriorate the estimate. Thus, the dynamics of the IMU have to be modeled to enable model-based fault detection.

Fault detection methods often use a residual showing the difference between the measurement and the prediction from the model. However, measurements from real-world applications are always contaminated with noise, which will give residuals that are non-zero. Assuming that the noise distribution is known, a fault can be detected if the deviation from the known distribution can be statistically secured. The statistical decision between fault/no fault can be made with a hypothesis test. One common approach to construct residuals for fault detection is the parity-space method (Chow and Willsky, 1984), where a batch of data is collected and compared with the model. However, nothing is assumed to be known about the initial state of the system in the data batch. The parity space approach solves this problem by using a projection that removes the initial state dependency. What if prior information about the initial state is available, can this information be used as well?

Faults are usually correlated in time. Considering the magnetic disturbances in the introduction above, the disturbance (or fault) is caused by someone walking by a magnetic object. Thus, the disturbance will not change as white noise but will have a smooth profile. The parameterization of such faults are discussed in the thesis.

The IMU extended with magnetometer can be used for orientation (attitude) estimation as discussed in the introduction. The orientation estimate is disturbed if the IMU is accelerated or the magnetometer is disturbed by a metallic object. Can these type of disturbances be detected with the linear theory discussed above?

A fundamental question in fault detection is whether a fault can be observed by studying the outputs of the system. Can explicit conditions for observability of faults be derived?

1.2 Contributions

The main contributions of this thesis are:

- The smoothing approach to estimate the initial state of a data batch window, used for generalized likelihood ratio, GLR, tests. This is treated in Chapter 4.
- The minimum variance estimate from a data batch is used to eliminate the initial state for the GLR-test in Chapter 4. The resulting test is shown to be equal to the parity space method.
- The efficient parameterization of incipient faults, presented in Section 2.2.
- The modeling and implementation of the methods above on an IMU and tests with measurement data are described in Chapters 5, 6 and 7.
- Observability tests for linearly growing faults are treated in Chapter 3.

The smoothing approach and the fault parameterization are earlier published in the following conference papers:

> D. Törnqvist, F. Gustafsson, and I. Klein. GLR tests for fault detection over sliding data windows. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 2005.
>
> D. Törnqvist and F. Gustafsson. Eliminating the initial state for the generalized likelihood ratio test. In *Proceedings of IFAC Symposium SAFEPROCESS*, Beijing, China, Aug. 2006. To appear.

Outside the scope of this thesis, the following conference papers are published:

> D. Törnqvist, E. Geijer Lundin, F. Gunnarsson, and F. Gustafsson. Transmission timing - a control approach to distributed uplink scheduling in WCDMA. In *Proceedings of American Control Conference*, Boston, MA, USA, June 2004.
>
> F. Gunnarsson, D. Törnqvist, E. Geijer Lundin, G. Bark, N. Wiberg, and E. Englund. Uplink transmission timing in WCDMA. In *Proceedings of IEEE Vehicular Technology Conference Fall*, Orlando, FL, USA, Oct. 2003.

1.3 Outline

The thesis starts with the description of used models, estimation and detection theory in Chapter 2. Conditions for the observability of faults are discussed in Chapter 3. Chapter 4 derives the GLR tests for fault detection, an example on a DC-motor is also given. The dynamic model for an IMU is derived in Chapter 5. Measurement data from the IMU is presented and analyzed in Chapter 6. The GLR tests derived in Chapter 4 are then applied to the measurements in Chapter 7. Conclusions and thoughts about further studies are given in Chapter 8.

2

Estimation and Detection Theory

THE STATE OF A SYSTEM refers to the information needed to get complete knowledge about the behavior of the system at the moment. A common problem is to estimate the state of the system from measurements. As an example of a system, consider the economy of a country. Then we are able to measure company profits, salaries, stock values etc to get knowledge about the economy of the country. There are also inputs such as the interest rate given by the central bank and taxes decided by the state. If a model of the economy is available, the state of the economy can be estimated and the future economy can be predicted.

Another important task is change detection. In the economy example above, we might be interested in monitoring the stock market. Is someone trading on information that is not publicly available. If so, this would be a crime and should be investigated. With a proper model of the stock market and enough measurements, abnormal behavior could be detected using change detection methods.

2.1 System Descriptions

It is often of interest to describe the behavior of a system mathematically. If the behavior of a system is known and it is known what affects it, the future behavior can be predicted or simulated. It is often convenient to see what will happen with a system without really trying it in practice. For instance, if you are working at a nuclear power plant and consider to move the control rods, you might be interested if this only will lead to an increase in the power production or if it also will cause a melt down of the plant. A system description is also useful for control of a system. For example the controlling computer of an autonomous airplane has to know how the airplane reacts to movements of the rudders. The computer needs a mathematical model, whereas a human pilot has a mental model built on years of experience.

The behavior of most physical systems can be mathematically modeled using differential equations. These equations can then be represented using different standard structures. This section will primarily focus on state space models, also extended to the case where data batches are treated.

2.1.1 State Space Model

A fairly general description of a physical system is provided by the nonlinear state space model

$$\dot{x}(t) = f\big(x(t), u(t), f(t), v(t)\big), \tag{2.1a}$$
$$y(t) = h\big(x(t), u(t), f(t), e(t)\big), \tag{2.1b}$$

where x is the state of the system, u is the input signal, f is a fault, v and e are noises and y is the measured signal. The functions $f(\cdot)$ and $h(\cdot)$ are in general nonlinear functions. This work primarily focuses on the class of models where these are linear functions, so called linear systems. Linear systems are in general easier to handle and it therefore exist more powerful results for them, see for instance Rugh (1996) and Kailath et al. (2000). To handle nonlinear systems with linear theory, the nonlinear models are often approximated with a linear model.

Linear Model

A time-varying linear state space model can be written as

$$\dot{x}(t) = A(t)x(t) + B^u(t)u(t) + B^f(t)f(t) + B^v(t)v(t), \tag{2.2a}$$
$$y(t) = C(t)x(t) + D^u(t)u(t) + D^f(t)f(t) + e(t). \tag{2.2b}$$

Discrete-time Linear Model

A computer can only handle discrete-time data, so measurements will be sampled in time. Thus, the system model has to be sampled in order to describe the measurements. The procedure of sampling systems is described in Rugh (1996). In the case of a piecewise constant time-invariant linear system with piecewise constant input signal, the matrices of the sampled system can be computed as

$$F_t = e^{A(t)T}, \tag{2.3a}$$
$$G_t^s = \int_0^T e^{A(t)\tau}\, d\tau B^s(t), \tag{2.3b}$$

where $s \in \{u, f, v\}$. For time discrete-time systems, t will be used denote the sample rather than time. The discrete-time system can then be written as

$$x_{t+1} = F_t x_t + G_t^u u_t + G_t^f f_t + G_t^v v_t, \tag{2.4a}$$
$$y_t = H_t x_t + H_t^u u_t + H_t^f f_t + e_t. \tag{2.4b}$$

2.1.2 Autoregressive Models

A simple model for describing temporal dependencies in a signal is the autoregressive model, AR-model. It describes how present values of the signal depends on previous values as

$$y_t = -a_1 y_{t-1} - \cdots - a_n y_{t-n} + e_t. \tag{2.5}$$

The n:th order AR-model in (2.5) can also be described using a state space model. Writing it in observability canonical form, see for instance Kailath (1980), the equivalent state space model becomes

$$x_{t+1} = \begin{pmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{pmatrix} x_t + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ e_t \end{pmatrix}, \tag{2.6a}$$

$$y_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} x_t. \tag{2.6b}$$

2.1.3 Batched Systems

It is sometimes convenient to describe the behavior of a system over a certain time window. The system can then be described in batched form, directly derived from the linear state-space form. The batch form is often used in fault detection and diagnosis, see Chow and Willsky (1984), Gertler (1998), Gustafsson (2001). Stack L signal values to define the signal vectors like $\mathbb{Y} = \left(y_{t-L+1}^T, \ldots, y_t^T\right)^T$, for all signals. If the initial state in a time window of length L is known, the outputs in that window can be computed as

$$\mathbb{Y} = \mathcal{O}_t x_{t-L+1} + \bar{H}_t^u \mathbb{U} + \bar{H}_t^f \mathbb{F} + \bar{H}_t^v \mathbb{V} + \mathbb{E}, \tag{2.7}$$

with the extended observability matrix

$$\mathcal{O}_t = \begin{pmatrix} H_{t-L+1} \\ H_{t-L+2} F_{t-L+1} \\ \vdots \\ H_t \prod_{k=t-1}^{t-L+1} F_k \end{pmatrix}, \tag{2.8}$$

which in the time invariant case will have the following form

$$\mathcal{O} = \begin{pmatrix} H \\ HF \\ \vdots \\ HF^{L-1} \end{pmatrix}. \tag{2.9}$$

The matrices determining how the input signals affect the system is described by

$$\bar{H}_t^s = \begin{pmatrix} H_{t-L+1}^s & 0 & \cdots & 0 \\ H_{t-L+2} G_{t-L+1}^s & H_{t-L+2}^s & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ H_t \prod_{k=t-1}^{t-L+2} F_k G_{t-L+1}^s & \cdots & H_t G_{t-1}^s & H_t^s \end{pmatrix}, \tag{2.10}$$

and in case of a time invariant system, they will have the form of a Toeplitz matrix

$$\bar{H}^s = \begin{pmatrix} H^s & 0 & \cdots & 0 \\ HG^s & H^s & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ HF^{L-2}G^s & \cdots & HG^s & H^s \end{pmatrix}. \tag{2.11}$$

It is often convenient to describe a model where the influence of the input is removed. Therefore define

$$\mathbb{Z} \triangleq \mathbb{Y} - \bar{H}^u\mathbb{U} = \mathcal{O}x_{t-L+1} + \bar{H}^f\mathbb{F} + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{2.12}$$

The batched system can also be described as a function of the last state in the window instead of the first one. This requires that the matrix F is invertible. In this form, all the outputs can be computed as

$$\mathbb{Y} = \Gamma x_t + \tilde{H}^u\mathbb{U} + \tilde{H}^f\mathbb{F} + \tilde{H}^v\mathbb{V} + \mathbb{E}, \tag{2.13}$$

where

$$\Gamma = \begin{pmatrix} HF^{-(L-1)} \\ \vdots \\ HF^{-1} \\ H \end{pmatrix} \tag{2.14}$$

and

$$\tilde{H}^s = \begin{pmatrix} -HF^{-1}G^s & \cdots & -HF^{-(L-1)}G^s & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -HF^{-1}G^s & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} + \begin{pmatrix} H^s & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H^s \end{pmatrix}. \tag{2.15}$$

Define, similarly to (2.12), a system with removed input as

$$\tilde{\mathbb{Z}} \triangleq \mathbb{Y} - \tilde{H}^u\mathbb{U} = \Gamma x_t + \tilde{H}^f\mathbb{F} + \tilde{H}^v\mathbb{V} + \mathbb{E}. \tag{2.16}$$

A combination of the batched system models described in (2.7) and (2.13) can also be used. If a state in the middle of the window is known, that is x_{t-n}, the outputs are described by

$$\mathbb{Y} = \begin{pmatrix} \mathcal{O} \\ \Gamma \end{pmatrix} x_{t-n} + \begin{pmatrix} \bar{H}^u \\ \tilde{H}^u \end{pmatrix} \mathbb{U} + \begin{pmatrix} \bar{H}^f \\ \tilde{H}^f \end{pmatrix} \mathbb{F} + \begin{pmatrix} \bar{H}^v \\ \tilde{H}^v \end{pmatrix} \mathbb{V} + \mathbb{E}. \tag{2.17}$$

2.2 Structured Faults

The influence of a fault is often correlated in time and is changing much slower than white noise. It is therefore natural to model the behavior of the fault with as few parameters as possible. In this section, the fault is parameterized using an orthogonal basis generated by the Chebyshev polynomial. The Chebyshev polynomial is an orthogonal polynomial of

a discrete variable, see Abramowitz and Stegun (1965), Rivlin (1974). The polynomial, denoted $\Phi_n(t)$, is orthogonal in the interval $0 \leq t \leq N-1$, where t is an integer. Moreover, the vector of length N

$$\Phi_n^T = \begin{pmatrix} \Phi_n(0) \\ \vdots \\ \Phi_n(N-1) \end{pmatrix}, \tag{2.18}$$

is orthogonal to another vector Φ_m when $n \neq m$. The Chebyshev polynomial is defined by the function

$$\Phi_n(t) = n!\Delta^n \left[\binom{t}{n} \binom{t-N}{n} \right], \tag{2.19}$$

where Δ^n is the difference operator defined as

$$\Delta f(t) = f(t+1) - f(t), \quad \Delta^{n+1} f(t) = \Delta(\Delta^n f(t)). \tag{2.20}$$

The orthonormal vectors Φ_n will be used as basis vectors for the fault. Denoting the parameters with θ_t, the fault vector is modeled as

$$\mathbb{F} = \Phi^T \theta_t, \tag{2.21}$$

where

$$\Phi^T = \begin{pmatrix} \Phi_1^T & \Phi_2^T & \cdots \end{pmatrix} \tag{2.22}$$

The regressors obtained have the form of a constant, linear, quadratic term and so on. This means that a step fault can be described using a one dimensional basis and smooth faults using a low dimensional basis. An example of a parameterized fault is given in Example 2.1.

Example 2.1: Parameterized Fault

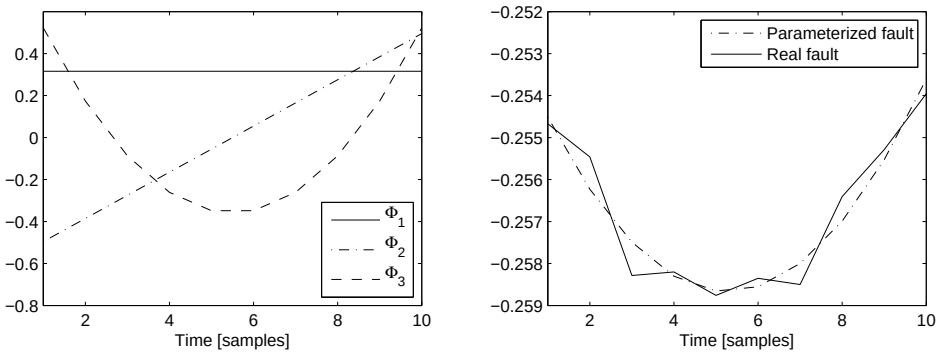
The output from a magnetometer (compass) is monitored. After some time, the magnetic field is disturbed by a magnetic object. This disturbance can be modeled using the Chebyshev polynomial matrix Φ^T. To estimate the model parameters, the following loss function is minimized

$$\hat{\theta}_t = \arg\min_{\theta_t} \|\mathbb{F} - \Phi^T \theta_t\|^2. \tag{2.23}$$

For this example, the measurement data comes from a real process given by data set D1 in Section 6.1. The fault is parameterized using three basis functions shown in Figure 2.1a. The parameterized fault is shown in Figure 2.1b.

2.3 Minimum Variance Estimation

A common problem in signal processing is to estimate the state of the system given a system model and measurement data. There are many different approaches to solve this problem. There are both recursive and non-recursive methods which can yield unbiased



(a) *Fault basis* ***(b)*** *Parameterized fault*

Figure 2.1: *The basis functions (a) are used to parameterize the fault from a magnetometer disturbance (b).*

estimates of the state. Some of the methods also produce estimates with minimum variance. For linear estimators, they are then referred to as a *Best Linear Unbiased Estimator*, BLUE. With the batched system representation from Section 2.1.3 it is easy to write a non-recursive minimum variance estimator.

For the estimation here, it is assumed that no fault affects the system, that is $\mathbb{F} = 0$. Denote $\mathbb{N} \triangleq \tilde{H}^v \mathbb{V} + \mathbb{E}$ and assume that $\mathbb{N}$ is white Gaussian noise with $\mathsf{Cov}(\mathbb{N}) = \tilde{S} \succ 0$. The equation system $\Gamma x_t + \mathbb{N} = \tilde{\mathbb{Z}}$ is known as the *general Gauss-Markoff linear model* (Björck, 1996). The parameters in x_t can be estimated with minimum variance by solving the generalized least squares problem

$$\min_{x_t} (\Gamma x_t - \tilde{\mathbb{Z}})^T \tilde{S}^{-1} (\Gamma x_t - \tilde{\mathbb{Z}}), \tag{2.24}$$

see Björck (1996). This can be rewritten as the least squares problem

$$\min_{x_t} \|\tilde{S}^{-1/2} \Gamma x_t - \tilde{S}^{-1/2} \tilde{\mathbb{Z}}\|^2, \tag{2.25}$$

which has the solution

$$\hat{x}_t = (\tilde{S}^{-1/2} \Gamma)^\dagger \tilde{S}^{-1/2} \tilde{\mathbb{Z}}, \tag{2.26}$$

where $\star^\dagger$ denotes the Moore-Penrose pseudo inverse (see Golub and van Loan (1996)). The estimate will be distributed according to

$$\hat{x}_t \sim \mathsf{N}\left(x_t, (\Gamma^T S^{-1} \Gamma)^{-1}\right), \tag{2.27}$$

where N denotes the Gaussian distribution, see Section 2.7.1.

The initial state in the window can also be estimated with the framework described above. Denote the covariance matrix $S = \mathsf{Cov}(\bar{H}^v \mathbb{V} + \mathbb{E})$, then the initial state is estimated by

$$\hat{x}_{t-L+1} = (S^{-1/2} \mathcal{O})^\dagger S^{-1/2} \mathbb{Z} \sim N(x_{t-L+1}, (\mathcal{O}^T S^{-1} \mathcal{O})^{-1}) \tag{2.28}$$

An intermediate state of the window can be estimated using the model in (2.17). This estimate will then be referred to as a *smoothed estimate*, see Kailath et al. (2000).

2.4 Kalman Filter

The Kalman filter is a recursive filter for linear systems that is a BLUE. The filter was first presented in Kalman (1960) and is the most common estimator used. Many books give a detailed description of the Kalman filter, see for example Kailath et al. (2000), Anderson and Moore (1979). The noise is assumed to be white, Gaussian and the covariance matrices are given by $\mathsf{E}\, v_t v_{t+\tau} = Q_t \delta(\tau)$ and $\mathsf{E}\, e_t e_{t+\tau} = R_t \delta(\tau)$. The Kalman filter recursions are given in Algorithm 2.1. During the recursions, the error covariance for the states are computed as P_t.

Algorithm 2.1 Kalman Filter (KF)

An initial state, $\hat{x}_{0|-1} = x_0$, and an initial error covariance, $P_{0|-1} = P_0$, is given. Then the filter updates are given by

- Time update:

$$\hat{x}_{t+1|t} = F_t \hat{x}_{t|t} + G_t^u u_t, \tag{2.29a}$$
$$P_{t+1|t} = F_t P_{t|t} F_t^T + G_t^v Q_t {G_t^v}^T, \tag{2.29b}$$

- Measurement update:

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}, \tag{2.29c}$$
$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1} - H_t^u u_t), \tag{2.29d}$$
$$P_{t|t} = (I - K_t H_t) P_{t|t-1}. \tag{2.29e}$$

2.5 Extended Kalman Filter

To be able to use the Kalman filter framework also for a nonlinear system, the model has to be linearized. Linearizing the system around the estimated trajectory and then applying the Kalman filter gives the Extended Kalman Filter, EKF. Given the nonlinear system

$$x_{t+1} = f(x_t, u_t) + g(x_t, u_t) v_t, \tag{2.30a}$$
$$y_t = h(x_t, u_t) + e_t, \tag{2.30b}$$

the linear system is a first order Taylor-expansion around the latest state estimate,

$$f(x_t) \approx f(\hat{x}_{t|t}, u_t) + F_t (x_t - \hat{x}_{t|t}), \tag{2.31a}$$
$$h(x_t) \approx h(\hat{x}_{t|t-1}, u_t) + H_t (x_t - \hat{x}_{t|t-1}), \tag{2.31b}$$
$$g(x_t) \approx g(\hat{x}_{t|t}, u_t) \triangleq G_t^v, \tag{2.31c}$$

where

$$F_t \triangleq \left.\frac{\partial f(x,u)}{\partial x}\right|_{(x,u)=(\hat{x}_{t|t},u_t)} \quad \text{and} \quad H_t \triangleq \left.\frac{\partial h(x,u)}{\partial x}\right|_{(x,u)=(\hat{x}_{t|t-1},u_t)}.$$

Using this linearization, the state trajectory can be estimated by the EKF in Algorithm 2.2. Observe that optimality is not guaranteed by the EKF, but it usually performs good in practice.

Algorithm 2.2 Extended Kalman Filter (EKF), see (Kailath et al., 2000, p. 340)

An initial state, $\hat{x}_{0|-1} = x_0$, and an initial error covariance, $P_{0|-1} = P_0$, is given. Then the filter updates are given by

- Time update:

$$\hat{x}_{t+1|t} = f(\hat{x}_t, u_t) \tag{2.32a}$$
$$P_{t+1|t} = F_t P_{t|t} F_t^T + G_t^v Q_t G_t^{vT} \tag{2.32b}$$

- Measurement update:

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1} \tag{2.32c}$$
$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t\big(y_t - h(\hat{x}_{t|t-1}, u_t)\big) \tag{2.32d}$$
$$P_{t|t} = (I - K_t H_t) P_{t|t-1} \tag{2.32e}$$

2.6 Detection Theory

A common problem is to detect whether a change of some type has occurred. One important and common application is to detect a signal in noise. This is done everyday and is essential in for example cell phones and other wireless devices. Other applications are for example the area of fault detection. In an airplane, it can be a question of detecting if a rudder gives the normal response, otherwise we have a fault which might be a leakage in the hydraulic system controlling the rudder.

The detection problem is often formulated in a way that it is a question of deciding between a number of hypotheses. To make the decision, measurement data from some sensors and a model of the process is often provided. If there are two hypotheses to decide between, the problem is known as a *binary hypothesis test.* For the airplane example mentioned above, this could be

$$\begin{aligned} \mathcal{H}_0 &: \quad \text{Rudders work as normal,} \\ \mathcal{H}_1 &: \quad \text{The rudders are faulty,} \end{aligned}$$

where the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ usually are referred to as *null hypothesis* and *alternative hypothesis* respectively. Using sensor readings and the model of the airplane a *test*

statistic can be derived. The test statistic will have different distributions in case a fault is present or not. If the distributions are completely known in both cases, we will have a *simple hypothesis test.* Otherwise, if some parameters are unknown, we have a *composite hypothesis test.*

2.6.1 Simple Hypothesis Testing

Consider the case where the binary variable $\theta = \{\theta_0, \theta_1\}$ is measured with additive noise e which has a known distribution,

$$y = \theta + e.$$

The hypothesis test

$$\begin{aligned} \mathcal{H}_0 : &\quad \theta = \theta_0, \\ \mathcal{H}_1 : &\quad \theta = \theta_1, \end{aligned}$$

is an example of a simple hypothesis test if the parameters θ_0 and θ_1 are known. The probability of the measurement can then be calculated depending on which hypothesis that is assumed true. One approach to selection between the hypotheses is to select the hypothesis that gives the highest probability for the measurement. Hence, if $p(y|\,\mathcal{H}_1) > p(y|\,\mathcal{H}_0)$, the alternative hypothesis is decided. Another approach would be to be more restrictive with causing an alarm, so the requirement for an alarm is that $p(y|\,\mathcal{H}_1) > \gamma p(y|\,\mathcal{H}_0)$ where $\gamma > 1$. This can also be expressed as a ratio between the probabilities, the *likelihood ratio*

$$\frac{p(y|\,\mathcal{H}_1)}{p(y|\,\mathcal{H}_0)} > \gamma.$$

In order to statistically describe these approaches, some definitions have to be made.

Definition 2.1. Some common terms in statistics are defined as

- False alarm (FA) is to decide $\mathcal{H}_1$ when $\mathcal{H}_0$ is true.
- Miss (M) is to decide $\mathcal{H}_0$ when $\mathcal{H}_1$ is true.
- Probability of false alarm (P_{FA}); $P_{FA} = P(\mathcal{H}_1\,|\,\mathcal{H}_0)$
- Probability of miss (P_M); $P_M = P(\mathcal{H}_0\,|\,\mathcal{H}_1)$
- Probability of detection (P_D); $P_D = P(\mathcal{H}_1\,|\,\mathcal{H}_1) = 1 - P_M$

When designing a test to decide between the hypotheses, a trade-off between high P_D and low P_{FA} has to be made. If it is decided to have P_{FA} on a certain level, what is the optimal test? This question is answered by the Neyman-Pearson,NP, lemma, which was first published in the articles Neyman and Pearson (1928a,b). Here, the lemma is presented in Theorem 2.1. The NP-lemma is referred to as the *Likelihood Ratio Test,* LRT, and is the most powerful test for simple hypotheses. Most powerful means that it maximizes P_D, given a P_{FA}.

Theorem 2.1 (Neyman-Pearson lemma)
To maximize P_D for a given $P_{FA} = \alpha$ decide $\mathcal{H}_1$ if

$$L(y) = \frac{p(y|\,\mathcal{H}_1)}{p(y|\,\mathcal{H}_0)} > \gamma, \tag{2.33}$$

where the threshold γ is given by

$$P_{FA} = \int_{y:L(y)>\gamma} p(y|\,\mathcal{H}_0)\,dy = \alpha. \tag{2.34}$$

Proof: See (Kay, 1998, Appendix 3A). □

2.6.2 Composite Hypothesis Testing

Consider the measurement

$$y = \theta + e, \tag{2.35}$$

where the measurement noise e has a known distribution but θ is unknown. The one-sided hypothesis

$$\mathcal{H}_0: \quad \theta = 0, \tag{2.36}$$
$$\mathcal{H}_1: \quad \theta > 0, \tag{2.37}$$

will have an unknown parameter in the distribution of y under the alternative hypothesis. This test is therefore called a composite hypothesis test. It is sometimes possible to find an optimal detector even for composite hypotheses. A detector that yields the highest P_D given P_{FA} for all values of θ is known as *Uniformly Most Powerful*, UMP. A two-sided hypothesis test could be

$$\mathcal{H}_0: \quad \theta = 0, \tag{2.38}$$
$$\mathcal{H}_1: \quad \theta \neq 0, \tag{2.39}$$

where the alternative hypothesis has two test regions for the parameter θ. It can be shown that a UMP test does not exist for a two-sided hypothesis test, see Kay (1998). Thus, in those cases suboptimal detectors have to be implemented.

To construct a detector for a composite hypothesis test, there are two major approaches:

1. Marginalization or Bayesian approach. If the unknown parameter has a known distribution it can be marginalized.

2. Generalized Likelihood Ratio Test, GLRT.

The approaches will be discussed in the following sections.

Marginalization

Consider again the measurement

$$y = \theta + e, \tag{2.40}$$

where e is noise with known distribution but θ is unknown. The hypotheses are

$$\mathcal{H}_0 : \quad \theta = \theta_0, \tag{2.41a}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1, \tag{2.41b}$$

where θ_0 and θ_1 are unknown. Now, there are unknown parameters in the distribution of y for both hypotheses. The marginal density function of y can be computed as

$$p(y|\,\mathcal{H}_0) = \int_{-\infty}^{\infty} p(y|\,\mathcal{H}_0, \theta_0) p(\theta_0)\, d\theta_0, \tag{2.42a}$$

$$p(y|\,\mathcal{H}_1) = \int_{-\infty}^{\infty} p(y|\,\mathcal{H}_1, \theta_1) p(\theta_1)\, d\theta_1. \tag{2.42b}$$

This can be thought of as computing the "expected distribution, $p(y)$". Inserting this density function in the LR test gives

$$\tilde{L}(y) = \frac{p(y|\,\mathcal{H}_1)}{p(y|\,\mathcal{H}_0)}. \tag{2.43}$$

Generalized Likelihood Ratio

Another way of eliminating the unknown parameters in (2.41) is to use the *maximum likelihood estimate* of the parameter. The parameters can then be calculated as

$$\hat{\theta}_\star = \arg\max_{\theta_\star} p(y|\theta_\star). \tag{2.44}$$

The estimated parameters, $\hat{\theta}_\star$, are then inserted into the Likelihood Ratio test provided in the NP-lemma which yields the Generalized Likelihood Ratio test,

$$\hat{L}(y) = \frac{p(y|\hat{\theta}_1)}{p(y|\hat{\theta}_0)}. \tag{2.45}$$

Even though the NP-test is optimal, the GLRT is not always optimal but often works well in practice. It is shown in Lehmann (1986) that GLRT is UMP among the tests that are invariant.

2.7 Appendix: Distributions

2.7.1 Gaussian Distribution

The most widely used distribution for random variables is the *Gaussian* or *Normal* distribution. Consider the stochastic variable X which is Gaussian distributed with mean μ and variance σ^2, denoted

$$X \sim \mathsf{N}(\mu,\, \sigma^2).$$

The probability density function, PDF, for X is then

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}. \tag{2.46}$$

This means that the probability for X to have a value in the interval $[a,\, b]$ is

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

In a more general case, X is an $n \times 1$ vector. The distribution of the Gaussian vector is then represented by the *multivariate Gaussian* PDF

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \det^{1/2}(S)} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T S^{-1}(\mathbf{x}-\mu)}, \tag{2.47}$$

where μ is the mean vector and S is the positive definite covariance matrix. An example of the PDF for a variable distributed with the parameters $\mu = \begin{pmatrix} 0 & 0 \end{pmatrix}^T$ and $S = \text{diag}(1,\, 2)$ is plotted in Figure 2.2.

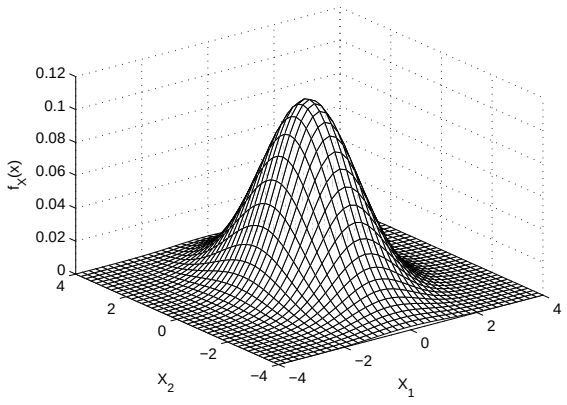


Figure 2.2: *The multivariate Gaussian distribution with parameters* $\mu = (0 \;\; 0)^T$ *and* $P = diag(1,\, 2)$.

2.7.2 Chi-square Distribution

Squared Gaussian variables are common, especially in detection theory. Often are sums of squared Gaussian residuals used as a test statistic and the result will then be chi-square distributed if the residuals are white and Gaussian. If the residuals have zero mean, the test statistic will have a *central chi-square* distribution. In case of non-zero mean, the distribution will be the *noncentral chi-square.*

Central Chi-square

Sums of squared independent and identically distributed, IID, Gaussian variables with zero mean will become chi-square distributed. That is, $X = \sum_{i=1}^{\nu} X_i^2$ with $X_i \sim \mathsf{N}(0,1)$ will be chi-square distributed with ν degrees of freedom, denoted $X \sim \chi_\nu^2$. The probability density function is

$$f_X(x) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}x} & x > 0 \\ 0 & x < 0 \end{cases}, \tag{2.48}$$

where $\Gamma(u)$ is the Gamma function, which is defined as

$$\Gamma(u) = \int_0^\infty t^{u-1} e^{-t}\, dt. \tag{2.49}$$

Noncentral Chi-square

When taking the sum of squared IID Gaussian random variables with non-zero mean, the distribution of the result will be represented by the *noncentral chi-square distribution.* Consider the case where $X_i \sim \mathsf{N}(\mu_i, 1)$, then the variable $X = \sum_{i=1}^{\nu} X_i^2$ has ν degrees of freedom and the *noncentrality parameter* $\lambda = \sum_{i=1}^{\nu} \mu_i^2$. The distribution of X is denoted $X \sim \chi_\nu'^2(\lambda)$ and the PDF can be expressed as

$$f_X(x) = \begin{cases} \frac{1}{2}\left(\frac{x}{\lambda}\right)^{\frac{\nu-2}{4}} I_{\frac{\nu}{2}-1}(\sqrt{\lambda x}) e^{-\frac{1}{2}(x+\lambda)} & x > 0 \\ 0 & x < 0 \end{cases}, \tag{2.50}$$

where $I_r(u)$ is the modified Bessel function of the first kind and order r. This function is defined as

$$I_r(u) = \frac{\left(\frac{1}{2}u\right)^r}{\sqrt{\pi}\Gamma\left(r + \frac{1}{2}\right)} \int_0^\pi \sin^{2r}(\theta) e^{u\cos\theta}\, d\theta. \tag{2.51}$$

3

Fault Observability

THE PURPOSE OF FAULT DETECTION is to discover the presence of a fault and for fault diagnosis to tell which fault is present. To get information about the system, measurements are made. If it is possible to estimate the fault from the measurements, fault detection and diagnosis becomes possible. This chapter derives tests that determine if the fault can be estimated.

3.1 Studied Systems

The systems studied are time-invariant, time-discrete linear systems in state space form as described in Section 2.1.1. Two special cases of this model structure are described. A system without inputs is referred to as the nominal system. The other system has a fault vector as input, but in order to estimate this input the fault vector is included in the state vector. Hence the name, augmented system.

3.1.1 Nominal System

The following state-space model is referred to as the nominal system

$$x_{t+1} = Fx_t + w_t, \tag{3.1a}$$
$$y_t = Hx_t + e_t, \tag{3.1b}$$

where $\dim(x_t) = n_x$ and $\dim(y_t) = n_y$, that is, F is $n_x \times n_x$ and H is $n_y \times n_x$.

3.1.2 Augmented System

A state-space model with faults as inputs would have the form

$$x_{t+1} = Fx_t + G^f f_t + w_t, \tag{3.2a}$$
$$y_t = Hx_t + H^f f_t + e_t, \tag{3.2b}$$

where $\dim(f_t) = n_f$, that is, G^f is $n_x \times n_f$ and H^f is $n_y \times n_f$.

In order to estimate the faults, the faults are included in the state vector. The augmented state vector is denoted $\bar{x}_t = \begin{pmatrix} x_t^T & f_t^T \end{pmatrix}^T$. Including the fault in the state vector implies that the dynamics of the fault has to be modeled. It is here assumed that the fault is changing according to the random walk process, which would yield the following augmented model

$$\bar{x}_{t+1} = \begin{pmatrix} F & G^f \\ 0 & I \end{pmatrix} \bar{x}_t + \bar{w}_t, \tag{3.3a}$$
$$y_t = \begin{pmatrix} H & H^f \end{pmatrix} \bar{x}_t + e_t, \tag{3.3b}$$

where $\bar{w}_t$ is a noise vector. There can be several special cases of this model. For instance, if $G^f = 0$ and $H^f = I$ only measurement faults are modeled. Limiting to faults on the input side of the system, $G^f = I$ and $H^f = 0$.

3.2 Observability for Augmented Systems

The general definition of observability is given by Definition 3.1 and the tests for observability are given by Theorem 3.1 and 3.2. This section will derive explicit conditions under which the augmented system (3.3) is observable, given that the pair $\{H, F\}$ is chosen such that the nominal system (3.1) is observable. The same problem arises in applications where biases in the measurements or dynamics are estimated. The problem is studied in for example Chmielewski Jr. and Kalata (1995), Bembeneck et al. (1998).

Definition 3.1. A system (3.1) is said to be *observable* if any initial state x_0 can be uniquely determined by the corresponding response y.

Theorem 3.1
A system (3.1) *is observable if and only if the matrix*

$$\mathcal{O} = \begin{pmatrix} H \\ HF \\ \vdots \\ HF^{n_x} \end{pmatrix} \tag{3.4}$$

has full column rank.

Proof: The proof is given in most textbooks on linear systems, see for example Kailath (1980). □

Theorem 3.2 (Popov-Belevitch-Hautus (PBH) Test)
A system (3.1) *is observable if and only if the matrix*

$$\begin{pmatrix} H \\ F - sI \end{pmatrix} \tag{3.5}$$

has full column rank for all s*.*

Proof: See (Kailath, 1980, p. 137). □

3.2.1 General Augmented System

Applying Theorem 3.2 for the augmented system (3.3), the observability is given by checking the rank of

$$\begin{pmatrix} \begin{pmatrix} H & H^f \end{pmatrix} \\ \begin{pmatrix} F & G^f \\ 0 & I \end{pmatrix} - sI \end{pmatrix} = \begin{pmatrix} H & H^f \\ F - sI & G^f \\ 0 & (1-s)I \end{pmatrix}. \tag{3.6}$$

The rank should be checked for all s, hence divide the analysis into two parts $s \neq 1$ and $s = 1$. When $s \neq 1$, the matrix $(s-1)I$ will have full rank and using row operations (rank conserving) (3.6) can be rewritten as

$$\begin{pmatrix} H & 0 \\ F - sI & 0 \\ 0 & (1-s)I \end{pmatrix}. \tag{3.7}$$

Thus, it is sufficient to check the rank of the matrix

$$\begin{pmatrix} H \\ F - sI \end{pmatrix}, \tag{3.8}$$

which we know have full rank since the nominal system is observable. The other case $s = 1$ gives the matrix

$$\begin{pmatrix} H & H^f \\ F - I & G^f \\ 0 & 0 \end{pmatrix}, \tag{3.9}$$

and thus the observability of the augmented system can be determined by checking the rank of

$$\begin{pmatrix} H & H^f \\ F - I & G^f \end{pmatrix}. \tag{3.10}$$

3.2.2 System with Measurement Faults

This section will study the special case with measurement faults and no faults in the dynamics, that is, $G^f = 0$. The faults will be assumed to be such that H^f has full column rank. With these assumptions the rank condition for the augmented system is to check whether

$$\begin{pmatrix} H & H^f \\ F - I & 0 \end{pmatrix} \tag{3.11}$$

has full column rank. An interesting observation is that if F has no eigenvalues 1, that is, no pure integrations, a sufficient condition for observability is that H^f has full (column) rank.

Full column rank of the matrix B is equivalent to

$$Bx = 0 \quad \Leftrightarrow \quad x = 0. \tag{3.12}$$

To obtain a result when F includes pure integrations, study

$$\begin{pmatrix} H & H^f \\ F-I & 0 \end{pmatrix} \begin{pmatrix} x \\ f \end{pmatrix} = \begin{pmatrix} Hx + H^f f \\ (F-I)x \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x \\ f \end{pmatrix} = 0. \tag{3.13}$$

The lower block is zero if and only if x is an eigenvector of F with eigenvalue 1. Hence, a necessary and sufficient condition for the augmented system to be observable is that

$$Hx \notin \mathcal{R}(H^f) \quad \text{for all } x \text{ such that} \quad Fx = x \neq 0, \tag{3.14}$$

which can also be expressed as

$$\mathcal{R}(HU) \cap \mathcal{R}(H^f) = \emptyset, \tag{3.15}$$

where the columns of U consist of the eigenvectors of F with eigenvalue 1. (Note that $Hx \neq 0$ since otherwise $\mathcal{O}x = 0$ in contradiction to the assumption that the nominal system is observable.) Put in another way, the fault must enter the measurements orthogonally to the integrating part of state space.

Observations

Some observations based on the discussion above:

- As long as the studied system does not contain any pure integration, that is, $Fx = x$ has no solution apart from $x = 0$, any fault is allowed if H^f has full rank.
- With an integration in the system the problem is more difficult:
  - If H^f has full row rank the augmented system is not observable, see the requirement in (3.15).
  - The fault terms may not look the same in the measurements as in the integrated state, or the system is not observable. It is impossible to tell the difference between the initial state and fault, as the effect on the system is the same.

3.2.3 Dynamics with Additive Faults

Now consider the case of additive faults only in the dynamics, that is, $H^f = 0$. The augmented system is observable if and only if

$$\begin{pmatrix} H & 0 \\ F-I & G^f \end{pmatrix}. \tag{3.16}$$

For the continued discussion, the following result related to the test for full column rank is useful.

Lemma 3.1
The matrix

$$\begin{pmatrix} A \\ B \end{pmatrix} \tag{3.17}$$

has full column rank if and only if $\mathcal{N}(A) \cap \mathcal{N}(B) = \emptyset$.

Proof: The matrix in (3.17) has full column rank if and only if $\nexists s \neq 0$ such that

$$\begin{pmatrix} A \\ B \end{pmatrix} u = \begin{pmatrix} Au \\ Bu \end{pmatrix} = 0. \tag{3.18}$$

Obtaining rank deficiency implies that $Au = 0$ and $Bu = 0$ for the same u and hence must have a common null space. The requirement for full column rank is therefore non-intersecting null spaces of A and B as written in the lemma. □

In order to determine observability of the augmented system, it is interesting to know under which circumstances (3.16) has full column rank. From Lemma 3.1, we conclude that it is enough to consider the null space of (3.16). Rank deficiency can only be obtained when the following equation has a nonzero solution

$$\begin{pmatrix} H & 0 \\ F - I & G^f \end{pmatrix} \begin{pmatrix} x \\ f \end{pmatrix} = \begin{pmatrix} Hx \\ (F - I)x + G^f f \end{pmatrix} = 0 \tag{3.19}$$

The top row can only be zero for x in the null space of H, hence it is enough to study such x. Therefore, form a matrix U which has a basis for the null space of H as columns. Then, the second row of (3.19) can be written as

$$(F - I)Us + G^f f = 0, \tag{3.20}$$

where s is an arbitrary vector. For this equation to have a solution, the matrices $(F - I)U$ and G^f have to share image spaces. The requirement for observability can then be written as

$$\mathcal{R}\left((F - I)U\right) \cap \mathcal{R}(G^f) = \emptyset. \tag{3.21}$$

The interpretation of this is explained below. Consider a state x_t that is in the null space of H and is thus not directly observable in the measurements. Consider now how this state changes in a time update. The change can be written

$$x_{t+1} - x_t = (F - I)x_t + G^f f, \tag{3.22}$$

the requirement in (3.21) states that the change introduced by the dynamics of the system must be orthogonal to the influence by the disturbance. That is, it must be possible to distinguish between contributions to the state that comes from parts of the state space not directly measured and contributions from a disturbance.

3.3 Batched Systems with Linearly Growing Faults

An alternative, but completely analogous, view of the discussion above can be obtained considering systems of batched signals described in Section 2.1.3. Consider batched signals over a window of length L, then the system can be described as

$$\mathbb{Y} = \mathcal{O}x_{t-L+1} + \bar{H}^f\mathbb{F} + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{3.23}$$

The matrix $\mathcal{O}$ will equal the observability matrix given in Theorem 3.1 and since the nominal system is assumed to be observable, it will have full column rank. If the unknown disturbance can be parameterized as

$$\mathbb{F} = \Phi^T\theta_t \tag{3.24}$$

and $\bar{H}^\theta \triangleq \bar{H}^f\Phi^T$, the system can be rewritten as

$$\mathbb{Y} = \begin{pmatrix}\mathcal{O} & \bar{H}^\theta\end{pmatrix}\begin{pmatrix}x_{t-n}\\ \theta_t\end{pmatrix} + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{3.25}$$

To determine observability of this augmented system, the length of the window has to be equal to the number of states, that is $L = n_x + n_\theta$.

If it is assumed that the fault is changing linearly, that is $f_t = ct + m$. Then, the fault can be parameterized as in (3.24) with

$$\Phi^T = \begin{pmatrix}1 & t\\ \vdots & \vdots\\ 1 & t+L-1\end{pmatrix}. \tag{3.26}$$

Now computing $\bar{H}^\theta$ with $L = n_x + n_\theta = n_x + 2$ gives

$$\begin{aligned}\bar{H}^\theta = &\begin{pmatrix}H^f & 0 & \cdots & 0\\ HG^f & H^f & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ HF^{n_x}G^f & HF^{n_x-1}G^f & \cdots & H^f\end{pmatrix}\begin{pmatrix}1 & t\\ \vdots & \vdots\\ 1 & t+n_x+1\end{pmatrix} =\\ &\begin{pmatrix}H^f & tH^f\\ HG^f + H^f & tHG^f + (t+1)H^f\\ H(F+I)G^f + H^f & tHFG^f + (t+1)HG^f + (t+2)H^f\\ \vdots & \vdots\\ H\sum_{i=0}^{n_x}F^iG^f + H^f & H\sum_{i=0}^{n_x}(t+i)F^{n_x-i}G^f + (t+n_x+1)H^f\end{pmatrix}.\end{aligned} \tag{3.27}$$

The observability of the augmented system is determined by the rank of

$$\begin{pmatrix}\mathcal{O} & \bar{H}^\theta\end{pmatrix}. \tag{3.28}$$

Simple row and column operations preserves rank. To simplify the matrix, first substract the first row from the second, then the second from the third and so on. Then, substract

t times the second column from the third column. Simultaneously done below is again substracting the first row from the second, then the second from the third and so on.

$$\begin{pmatrix}\mathcal{O} & \bar{H}^\theta\end{pmatrix} \sim \begin{pmatrix} H & H^f & tH^f \\ H(F-I) & HG^f & tHG^f + H^f \\ HF(F-I) & HFG^f & tHFG^f + HG^f + H^f \\ \vdots & \vdots & \vdots \\ HF^{n_x}(F-I) & HF^{n_x}G^f & tHF^{n_x}G^f + H\sum_i^{n_x-1} F^i + H^f \end{pmatrix} \sim$$
$$\begin{pmatrix} H & H^f & 0 \\ H(F-I) & HG^f & H^f \\ H(F-I)^2 & H(F-I)G^f & HG^f \\ \vdots & \vdots & \vdots \\ HF^{n_x-1}(F-I)^2 & HF^{n_x-1}(F-I)G^f & HF^{n_x-1}G^f \end{pmatrix} =$$
$$\begin{pmatrix} \begin{pmatrix} H & H^f & 0 \\ H(F-I) & HG^f & H^f \end{pmatrix} \\ \mathcal{O}\begin{pmatrix}(F-I)^2 & (F-I)G^f & G^f\end{pmatrix} \end{pmatrix} \quad (3.29)$$

Now, to determine if (3.29) has full column rank it suffices to see if the rows have non-intersecting null spaces according to Lemma 3.1. Since $\mathcal{O}$ has full column rank, it is sufficient to look at the matrix

$$\begin{pmatrix} H & H^f & 0 \\ H(F-I) & HG^f & H^f \\ (F-I)^2 & (F-I)G^f & G^f \end{pmatrix}. \quad (3.30)$$

3.4 Summary

This section has derived explicit tests for observability of faults with certain dynamics. The test for constant fault dynamics (moving as a random walk) is provided by checking the rank of

$$\begin{pmatrix} H & H^f \\ F-I & G^f \end{pmatrix}. \quad (3.31)$$

For the special cases that there are only measurement faults or only additive faults in the dynamics, there are interpretations given for the test requirements (see Section 3.2.2 and 3.2.3).

For the case of linearly growing faults, the batched system description is used to derive observability tests. The faults are observable, given that the following matrix has full rank

$$\begin{pmatrix} H & H^f & 0 \\ H(F-I) & HG^f & H^f \\ (F-I)^2 & (F-I)G^f & G^f \end{pmatrix}. \quad (3.32)$$

4

Initial State Estimation for GLR Tests

A STANDARD APPROACH TO model-based change detection is to first compute a residual from a batch of data based on the model, and then check if this residual is significantly different from zero, see Basseville and Nikiforov (1993), Gustafsson (2001). For state-space models as considered here, the initial state affects the residual significantly, and its influence has to be removed. The unknown initial state can be approached in the following ways:

- Using a projection that removes the influence of initial conditions as done in the classical parity space approach (Basseville and Nikiforov, 1993, Chow and Willsky, 1984, Ding et al., 1999, Gertler, 1997, 1998).
- Using prior knowledge of the initial state obtained with a causal Kalman filter on past data preceding the data batch under consideration (Willsky and Jones, 1976). This gives a correct state estimate if the fault was not already present at that time.
- Using a minimum variance estimator (that is, anti-causal Kalman filter) on the data in the batch, which gives the correct state distribution if there is no fault in the batch.
- Using a smoothed estimate by combining an estimate from the data batch with the prior knowledge of the initial state.

The last two approaches are presumingly new, published by the author in Törnqvist and Gustafsson (2006). However, it will be shown that the third approach is identical to the first one. The anti-causal estimate is rather to be seen as a tool to derive the smoothed estimate, which will be shown to give the best performance in simulations.

A fault entering a system does often not change arbitrarily like white noise. It rather follows a smooth profile which could be parameterized in an efficient way. The parameterizations of the fault using a low-rank orthogonal basis will be shown to increase the performance significantly.

4.1 Studied Model

The class of models studied in this chapter can be either time varying or time invariant linear models. The systems are described with stacked models as described in Section 2.1.3

$$\mathbb{Y} = \mathcal{O}x + \bar{H}^u\mathbb{U} + \bar{H}^f\mathbb{F} + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{4.1}$$

The system is assumed to be observable, that is, $\mathcal{O}$ has full column rank. For non-observable or partially observable system, see Section 4.2.4. The influence of the known input signal can always be eliminated and therefore define

$$\mathbb{Z} \triangleq \mathbb{Y} - \bar{H}^u\mathbb{U} = \mathcal{O}x + \bar{H}^f\mathbb{F} + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{4.2}$$

The fault is modeled with the low dimensional parametrization as described in Section 2.2

$$\mathbb{F} = \Phi^T\theta_t \Rightarrow \bar{H}^\theta \triangleq \bar{H}^f\Phi^T. \tag{4.3}$$

4.2 State Estimation

There are several ways of estimating the initial state of the time window. Given the data in the window, the simplest way is to make an orthogonal projection of the measurements onto the signal space as described in Section 4.2.1. Another approach would be to consider the noise distribution and to make an oblique projection in order to obtain a minimum variance estimate, which is outlined in Section 4.2.2. A third way could be to combine one of the estimates above with a prior estimate from for instance a Kalman filter. This is described in Section 4.2.3.

4.2.1 Orthogonal Projection

A system (4.2) without faults is described by

$$\mathbb{Z} = \mathcal{O}x + \bar{H}^v\mathbb{V} + \mathbb{E}. \tag{4.4}$$

The least squares estimate of the initial state, without using the knowledge about noise covariances, is to use

$$\hat{x} = \mathcal{O}^\dagger\mathbb{Z}. \tag{4.5}$$

Here, $\star^\dagger$ denotes the Moore-Penrose pseudo inverse (see Golub and van Loan (1996)). This estimate is unbiased since

$$\mathsf{E}\,\hat{x} = \mathsf{E}\,\mathcal{O}^\dagger\mathbb{Z} = \mathsf{E}\left(x + \mathcal{O}^\dagger(\bar{H}^v\mathbb{V} + \mathbb{E})\right) = x, \tag{4.6}$$

observe that $\mathcal{O}^\dagger\mathcal{O} = I$ since the system is assumed to be observable and hence $\mathcal{O}$ has full column rank. In presence of faults, the expected value would be

$$\mathsf{E}\,\hat{x} = \mathsf{E}\,\mathcal{O}^\dagger\mathbb{Z} = x + \mathcal{O}^\dagger\bar{H}^\theta\theta_t. \tag{4.7}$$

Using the estimator in (4.5), the influence of the initial state on the output is given by

$$\mathcal{O}\hat{x} = \underbrace{\mathcal{O}\mathcal{O}^\dagger}_{\triangleq\mathcal{P}_\mathcal{O}}\mathbb{Z}, \tag{4.8}$$

where $\mathcal{P}_\mathcal{O}$ denotes the orthogonal projection onto the column space of $\mathcal{O}$, which follows by the following projection properties:

- A matrix $\mathcal{P}$ is a projector if and only if

$$\mathcal{P}^2 = \mathcal{P}. \tag{4.9}$$

- An orthogonal projector, $\mathcal{P}_A$, onto $\mathcal{R}(A)$ can be derived as

$$\mathcal{P}_A = AA^\dagger. \tag{4.10}$$

For proofs of the properties above see (Meyer, 2000, p. 387, 435).

4.2.2 Minimum Variance Estimation

The estimator derived in Section 4.2.1 gives an unbiased estimate of the initial state x. An estimator that also gives minimum variance for the estimation error was derived in Section 2.3. The estimate of the initial state is given by

$$\hat{x} = (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\mathbb{Z}. \tag{4.11}$$

This estimate will be distributed according to

$$\hat{x} \sim \mathsf{N}\left(x + (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\bar{H}^\theta\theta_t, (\mathcal{O}^T S^{-1}\mathcal{O})^{-1}\right). \tag{4.12}$$

The basis for x in the measurement space is $\mathcal{O}$, so estimating the term $\mathcal{O}x$ can be performed as

$$\mathcal{O}\hat{x} = \underbrace{\mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}}_{\triangleq\mathcal{P}}\mathbb{Z}, \tag{4.13}$$

where $\mathcal{P}$ is an oblique projection onto $\mathcal{R}(\mathcal{O})$ along $\mathcal{R}\left(S^{1/2}(S^{-1/2}\mathcal{O})^\perp\right)$ according to Lemma 4.1.

Lemma 4.1
The matrix

$$\mathcal{P} = \mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2} \tag{4.14}$$

is an oblique projection onto the space $\mathcal{R}(\mathcal{O})$ along the space $\mathcal{R}\left(S^{1/2}(S^{-1/2}\mathcal{O})^\perp\right)$.

Proof: The matrix $\mathcal{P}$ is a projection since it is idempotent, i.e., $\mathcal{P}\mathcal{P} = \mathcal{P}$, see (4.9). The dimension of $\mathcal{P}$ is $n \times n$ and a basis in its image space has rank p, then a basis for the null space has to have dimension $n - p$. The proof will show that $\mathcal{O}$ is a basis in the image space and therefore must have rank p. It is also shown that $S^{1/2}(S^{-1/2}\mathcal{O})^\perp$ is in the nullspace of $\mathcal{P}$ and have rank $n - p$.

The projection has the property $\mathcal{P}\mathcal{O} = \mathcal{O}$ and since $\mathcal{O}$ has full column rank $\mathrm{rank}(\mathcal{P}) \geq \mathrm{rank}(\mathcal{O})$. An upper bound on the rank of $\mathcal{P}$ is given by

$$\mathrm{rank}(\mathcal{P}) \leq \min(\mathrm{rank}(\mathcal{O}),\ \mathrm{rank}(S^{-1/2}\mathcal{O}),\ \mathrm{rank}(S^{-1/2})) = \mathrm{rank}(\mathcal{O}), \tag{4.15}$$

where the fact that the pseudo-inverse does not change the rank is used and that

$$\operatorname{rank}(S^{-1/2}\mathcal{O}) = \operatorname{rank}(\mathcal{O}) - \dim \mathcal{N}(S^{-1/2}) \cap \mathcal{R}(\mathcal{O}) = \operatorname{rank}(\mathcal{O}). \tag{4.16}$$

Thus, $\operatorname{rank}(\mathcal{P}) = \operatorname{rank}(\mathcal{O})$ and $\mathcal{O}$ must span the image space of $\mathcal{P}$. Then, $\operatorname{rank}(\mathcal{O}) = \operatorname{rank}(S^{-1/2}\mathcal{O}) = p$.

Another property of $\mathcal{P}$ is that $\mathcal{P}S^{1/2}(S^{-1/2}\mathcal{O})^{\perp} = 0$. To show this, study

$$\mathcal{P}S^{1/2}(S^{-1/2}\mathcal{O})^{\perp} = \mathcal{O}(S^{-1/2}\mathcal{O})^{\dagger}(S^{-1/2}\mathcal{O})^{\perp}. \tag{4.17}$$

Now, take the SVD

$$S^{-1/2}\mathcal{O} = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}. \tag{4.18}$$

Then, $(S^{-1/2}\mathcal{O})^{\dagger} = V_1\Sigma^{-1}U_1^T$ and $(S^{-1/2}\mathcal{O})^{\perp} = U_2U_2^T$. Since U_1 and U_2 are orthogonal, it is clear that $\mathcal{P}S^{1/2}(S^{-1/2}\mathcal{O})^{\perp} = 0$. Thus, $S^{1/2}(S^{-1/2}\mathcal{O})^{\perp}$ must be in the space that $\mathcal{P}$ projects along and is a basis in that space if the rank is $n - p$. With the same argument as in (4.16), it is clear that $\operatorname{rank}(S^{1/2}(S^{-1/2}\mathcal{O})^{\perp}) = \operatorname{rank}((S^{-1/2}\mathcal{O})^{\perp})$ and the orthogonal complement of a matrix of rank p must have rank $n - p$. □

4.2.3 Smoothed Estimate

It is a common situation in signal processing to have several estimates of the same variable. To get a common estimate using the information in all estimates are often referred to as the *sensor fusion problem*, see for instance Gustafsson (2001).

In this case, it is assumed that an estimate, $\hat{x}^{(1)}$, from a prior or Kalman filter is available. The estimate is assumed to be Gaussian distributed as

$$\hat{x}^{(1)} \sim \mathsf{N}(x, P^{(1)}). \tag{4.19}$$

Then, the smoothed initial state is estimated by forming the joint state estimate of (4.11) and (4.19) by the standard sensor fusion formula. Denote the estimate from (4.11) with $\hat{x}^{(2)}$ and its covariance with $P^{(2)}$, then

$$\hat{x} = P\Big({P^{(1)}}^{-1}\hat{x}^{(1)} + {P^{(2)}}^{-1}\hat{x}^{(2)}\Big), \tag{4.20}$$

where

$$P \triangleq \Big({P^{(1)}}^{-1} + {P^{(2)}}^{-1}\Big)^{-1}. \tag{4.21}$$

The expected value of this estimate is derived using (4.12) and (4.19)

$$\begin{aligned} \mathsf{E}\,\hat{x} &= \underbrace{P({P^{(1)}}^{-1} + {P^{(2)}}^{-1})}_{=I} x + P{P^{(2)}}^{-1}(S^{-1/2}\mathcal{O})^{\dagger}S^{-1/2}\bar{H}^{\theta}\theta_t \\ &= x + P\mathcal{O}^T S^{-1}\bar{H}^{\theta}\theta_t \end{aligned} \tag{4.22}$$

Note that $\mathsf{Cov}(\hat{x}) = P < P^{(2)}$ which means that the covariance of the estimate is decreased when prior information is used.

4.2.4 Partially Observable Systems

In order to estimate the initial state of the system (4.1) it has to be observable, that is, all states will have a unique influence in the output. When sensors get faulty and therefore cannot be used, the observability may be lost. However, it is often possible to observe a subspace of the state space and the system can therefore be partially observable. In order to handle such a system in the framework described above, the observable subspace has to be determined. It is given by the SVD

$$\mathcal{O} = U_1 \Sigma V_1^T, \tag{4.23}$$

where U_1 spans the observable subspace. The state $\bar{x} \triangleq \Sigma V_1^T x$ ($\dim(\bar{x}) < \dim(x)$) is now an observable state. The system (4.1) can then be rewritten as

$$\mathbb{Y} = U_1 \bar{x} + \bar{H}^u \mathbb{U} + \bar{H}^f \mathbb{F} + \bar{H}^v \mathbb{V} + \mathbb{E}. \tag{4.24}$$

Now, all methods described above can be used with $\mathcal{O}$ exchanged for U_1. When a prior estimate of the original estimate is available, as in Section 4.2.3, the state estimate and its covariance have to be transformed as

$$\hat{\bar{x}}^{(1)} = \Sigma V_1^T \hat{x}^{(1)} \tag{4.25}$$

$$P^{(1)}_{\hat{\bar{x}}^{(1)}} = \Sigma V_1^T P^{(1)} V_1 \Sigma. \tag{4.26}$$

4.3 GLR Tests

Fault detection is here considered as detecting whether the fault is zero or not. This approach corresponds to the hypothesis test

$$\mathcal{H}_0 : \ \theta_t = 0 \tag{4.27a}$$

$$\mathcal{H}_1 : \ \theta_t \neq 0. \tag{4.27b}$$

This section will derive different test statistics based on the Generalized Likelihood Ratio, GLR, for the hypothesis test above, see also Section 2.6.2. To perform this test, the initial state in the window has to be estimated using for instance the methods in Section 4.2. Using the estimate from the orthogonal projection is identified as the parity space method and is described in Section 4.3.1. If the minimum variance estimate is used the test will also reduce to the parity space method, see Section 4.3.2. Finally, the smoothed estimate is used in Section 4.3.3.

4.3.1 Parity Space Approach

The parity space method is a widely used method to find a residual for fault detection. It was first described in Chow and Willsky (1984), but is described and used in many publications for example Basseville and Nikiforov (1993), Ding et al. (1999), Gertler (1997, 1998).

Here, the parity space equations are derived by estimating the initial state by an orthogonal projection as in Section 4.2.1. The prediction error is here computed as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = (I - \mathcal{P}_{\mathcal{O}})\mathbb{Z} = \mathcal{P}_{\mathcal{O}^\perp}\mathbb{Z}. \tag{4.28}$$

The projector can be computed as

$$\mathcal{P}_{\mathcal{O}^\perp} = \mathcal{B}_{\mathcal{O}^\perp}\,\mathcal{B}_{\mathcal{O}^\perp}^T, \tag{4.29}$$

where $\mathcal{B}_{\mathcal{O}^\perp}$ is an orthonormal basis for $\mathcal{R}(\mathcal{O}^\perp)$.

Residual

Since the prediction error is computed by a rank deficient projection, the information in the prediction error can be represented by a vector of lower dimension. This low-dimensional vector is here denoted the residual. Note that $\mathcal{B}_{\mathcal{O}^\perp}^T\,\mathcal{B}_{\mathcal{O}^\perp} = I$ due to orthogonality and that the dimension is the rank of the projection in (4.29). The residual can be written as

$$r = \mathcal{B}_{\mathcal{O}^\perp}^T\,\varepsilon = \mathcal{B}_{\mathcal{O}^\perp}^T\,\mathcal{B}_{\mathcal{O}^\perp}\,\mathcal{B}_{\mathcal{O}^\perp}^T\,\mathbb{Z} = \mathcal{B}_{\mathcal{O}^\perp}^T\,\mathbb{Z}, \tag{4.30}$$

which has the covariance

$$\mathsf{Cov}(r) = \mathcal{B}_{\mathcal{O}^\perp}^T\,S\,\mathcal{B}_{\mathcal{O}^\perp}\,. \tag{4.31}$$

The normalized residual is then

$$\bar{r} = \underbrace{(\mathcal{B}_{\mathcal{O}^\perp}^T\,S\,\mathcal{B}_{\mathcal{O}^\perp})^{-1/2}\,\mathcal{B}_{\mathcal{O}^\perp}^T}_{\triangleq \bar{W}_1^T}\,\mathbb{Z}. \tag{4.32}$$

Test Statistic

In the hypothesis test, (4.27), the question is whether the fault is zero or not. This can be rearranged to a hypothesis test based on the residual as

$$\mathcal{H}_0:\ \bar{r} \sim N(0, I) \tag{4.33a}$$
$$\mathcal{H}_1:\ \bar{r} \sim N(\bar{W}_1^T\bar{H}^\theta\theta_t, I). \tag{4.33b}$$

To decide between the hypotheses, a GLR test will be performed with test statistic

$$\begin{aligned} L = 2\log\left(\frac{\sup_{\theta_t} p(\bar{r} - \bar{W}_1^T\bar{H}^\theta\theta_t)}{p(\bar{r})}\right) &= \sup_{\theta_t} 2\log\frac{\frac{1}{(2\pi)^{n/2}}e^{-\frac{1}{2}\|\bar{r}-\bar{W}_1^T\bar{H}^\theta\theta_t\|_2^2}}{\frac{1}{(2\pi)^{n/2}}e^{-\frac{1}{2}\|\bar{r}\|_2^2}} \\ &= \sup_{\theta_t} -\|\bar{r} - \bar{W}_1^T\bar{H}^\theta\theta_t\|_2^2 + \|\bar{r}\|_2^2, \end{aligned} \tag{4.34}$$

where n denotes the dimension of $\bar{r}$. This function is maximized for $\theta_t = (\bar{W}_1^T\bar{H}^\theta)^\dagger\bar{r}$, which gives

$$L = -\|\bar{r} - \underbrace{\bar{W}_1^T\bar{H}^\theta(\bar{W}_1^T\bar{H}^\theta)^\dagger}_{\mathcal{P}_{\bar{W}_1^T\bar{H}^\theta}}\bar{r}\|_2^2 + \|\bar{r}\|_2^2 = \bar{r}^T\mathcal{P}_{\bar{W}_1^T\bar{H}^\theta}\bar{r}. \tag{4.35}$$

4.3.2 Minimum Variance Estimation

The prediction error can be computed in minimum variance sense with the state estimate from (4.13) as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = (I - \mathcal{P})\mathbb{Z}. \tag{4.36}$$

The matrix $I - \mathcal{P}$ is the complementary projector to $\mathcal{P}$, see Meyer (2000), which has the SVD

$$I - \mathcal{P} = \begin{pmatrix} \mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp} & \mathcal{B}_{S^{-1}\mathcal{O}} \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{B}_{\mathcal{O}^\perp}^T \\ \mathcal{B}_{\mathcal{O}}{}^T \end{pmatrix} = \mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp} \Sigma \, \mathcal{B}_{\mathcal{O}^\perp}^T . \tag{4.37}$$

where $\mathcal{B}_\star$ is an orthogonal basis in $\mathcal{R}(\star)$. The construction of the SVD can be done using the result of Lemma 4.1. To construct the residual, the dimension of ε should be reduced to the rank of $I - \mathcal{P}$. By inspection of (4.37), the natural premultiplier would be $\Sigma^{-1}\mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp}{}^T$. Since

$$\Sigma^{-1}\mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp}{}^T (I - \mathcal{P}) = \mathcal{B}_{\mathcal{O}^\perp}^T, \tag{4.38}$$

the residual will be

$$r = \mathcal{B}_{\mathcal{O}^\perp}^T \mathbb{Z}. \tag{4.39}$$

This coincides with the residual in the parity-space approach and the hypothesis test therefore coincides with the one in Section 4.3.1.

4.3.3 Estimation by Smoothing

The prediction error is formed using the smoothed estimate from (4.20) as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = \mathcal{O}(x - \hat{x}) + \bar{H}^\theta \theta_t + \bar{H}^v \mathbb{V} + \mathbb{E}, \tag{4.40}$$

then the expected value of the prediction error is obtained using (4.22) as

$$\mathsf{E}\,\varepsilon = -\mathcal{O}P\mathcal{O}^T S^{-1} \bar{H}^\theta \theta_t + \bar{H}^\theta \theta_t = \underbrace{\left(I - \mathcal{O}P\mathcal{O}^T S^{-1}\right)}_{\triangleq W_2^T} \bar{H}^\theta \theta_t. \tag{4.41}$$

To construct the test statistic, it is desirable to have ε as a function of the measurements, $\mathbb{Z}$, and the state estimate from the Kalman filter, $\hat{x}^{(1)}$. Thus,

$$\begin{aligned} \varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} &= \mathbb{Z} - \mathcal{O}P\Big(P^{(1)^{-1}}\hat{x}^{(1)} + P^{(2)^{-1}}\hat{x}^{(2)}\Big) \\ &= \Big(I - \mathcal{O}PP^{(2)^{-1}}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\Big)\mathbb{Z} - \mathcal{O}PP^{(1)^{-1}}\hat{x}^{(1)} \\ &= W_2^T\mathbb{Z} - \mathcal{O}PP^{(1)^{-1}}\hat{x}^{(1)} \end{aligned} \tag{4.42}$$

and the covariance is

$$\mathsf{Cov}(\varepsilon) = W_2^T S W_2 + \mathcal{O}\underbrace{PP^{(1)^{-1}}P}_{\triangleq Q}\mathcal{O}^T = \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} W_2 \\ \mathcal{O}^T \end{pmatrix}. \tag{4.43}$$

It is important to note that the matrix W_2^T is time varying due to the dependence of the covariance of the Kalman filter estimate. However, the Kalman filter will become stationary and then the covariance converges to a matrix which can be computed by solving the Riccati equation.

Lemma 4.2
Assume that S and Q are both positive definite matrices, then the matrix in (4.43) *is positive definite.*

Proof: If (4.43) is a positive definite matrix, the following must hold

$$x^T \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} W_2 \\ \mathcal{O}^T \end{pmatrix} x > 0 \quad \forall x \neq 0. \tag{4.44}$$

Since it is known that S and Q are positive definite, it suffices to show that

$$x^T \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} \neq 0 \quad \forall x \neq 0. \tag{4.45}$$

The matrix W_2^T can be written as $I - \mathcal{O}M$ where M is an arbitrary matrix. If $x \in \mathcal{N}(\mathcal{O}^T)$ and $x \neq 0$ then

$$x^T \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} = \begin{pmatrix} x^T & 0 \end{pmatrix} \neq 0. \tag{4.46}$$

If $x \notin \mathcal{N}(\mathcal{O}^T)$ and $x \neq 0$ then

$$x^T \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} = \begin{pmatrix} x^T - x^T\mathcal{O}M & x^T\mathcal{O} \end{pmatrix} \neq 0 \tag{4.47}$$

since at least the second element is nonzero. Thus, we have shown that there is no $x \neq 0$ such that (4.45) is invalidated. □

Since $Q = PP^{(1)^{-1}}P$ where $P^{(1)}$ and P are positive definite matrices, Q is also positive definite. Then, according to Lemma 4.2, the covariance of ε is positive definite and thus invertible. The normalized residual can therefore be expressed as

$$\bar{r} = \mathsf{Cov}(\varepsilon)^{-1/2}\varepsilon = \underbrace{\mathsf{Cov}(\varepsilon)^{-1/2}W_2^T}_{\triangleq \bar{W}_2^T}\mathbb{Z} - \mathsf{Cov}(\varepsilon)^{-1/2}\mathcal{O}PP^{(1)^{-1}}\hat{x}^{(1)}. \tag{4.48}$$

Test Statistic

Using the normalized residual (4.48), the hypothesis test (4.27) can be written as

$$\mathcal{H}_0 : \ \bar{r} \sim \mathsf{N}(0, I) \tag{4.49a}$$

$$\mathcal{H}_1 : \ \bar{r} \sim \mathsf{N}(\bar{W}_2^T\bar{H}^\theta\theta_t, I). \tag{4.49b}$$

This yields the log-likelihood ratio

$$L = \sup_{\theta_t} \log \frac{e^{-\frac{1}{2}\|\bar{r} - \bar{W}_2^T\bar{H}^\theta\theta_t\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}\|_2^2}} = \sup_{\theta_t} -\Big(\|\bar{r} - \bar{W}_2^T\bar{H}^\theta\theta_t\|_2^2 - \|\bar{r}\|_2^2\Big), \tag{4.50}$$

which is maximized for $\theta_t = (\bar{W}_2^T\bar{H}^\theta)^\dagger\bar{r}$. Then

$$L = -\Big(\|(I - \underbrace{\bar{W}_2^T\bar{H}^\theta(\bar{W}_2^T\bar{H}^\theta)^\dagger}_{\mathcal{P}_{\bar{W}_2^T\bar{H}^\theta}})\bar{r}\|_2^2 - \|\bar{r}\|_2^2\Big) = \bar{r}^T\mathcal{P}_{\bar{W}_2^T\bar{H}^\theta}\bar{r} \tag{4.51}$$

4.3.4 Statistics

To choose suitable thresholds for the test statistics above, it is necessary to compute their distributions. While having Gaussian noise, the test statistics will be chi-square distributed variables, see Section 2.7.2. For a thorough review of statistics in signal processing, see Kay (1998). The normalized residual is distributed as

$$\bar{r} \sim \mathsf{N}(\bar{W}_\star^T \bar{H}^\theta \theta_t, I), \tag{4.52}$$

where $\star$ is 1 or 2 (depending on method) and $\theta_t = 0$ under the null hypothesis (4.27a). The test statistic is then distributed as the noncentral chi-square distribution

$$L = \bar{r}^T \mathcal{P}_{W_\star^T \bar{H}^\theta} \bar{r} \sim \chi'^2_\nu(\lambda) \tag{4.53}$$

where $\nu = \operatorname{rank}(W_\star^T \bar{H}^\theta)$ and

$$\lambda = (W_\star^T \bar{H}^\theta \theta_t)^T \mathcal{P}_{W_\star^T \bar{H}^\theta} W_\star^T \bar{H}^\theta \theta_t = (W_\star^T \bar{H}^\theta \theta_t)^T W_\star^T \bar{H}^\theta \theta_t. \tag{4.54}$$

Observe that $\lambda = 0$ in the fault-free case and then the test statistic is distributed according to the central chi-square distribution $L \sim \chi^2_\nu$. The threshold is then chosen from the chi-square distribution so that the fault-free hypothesis is rejected erroneously only with a small probability.

4.4 Example

To show the performance of the fault detection algorithms, a DC-motor has been simulated. The DC-motor can have a fault which is interpreted as a torque disturbance or a fault in the voltage supply.

4.4.1 Modeling

The continuous-time transfer function from input voltage to axle angle is

$$G(s) = \frac{1}{s(s+1)}. \tag{4.55}$$

Using the state variables x_1 (angle) and x_2 (angular velocity), a sampled version of the DC-motor can be written as

$$x_{t+1} = Fx_t + G^u u_t + G^f f_t + G^w w_t \tag{4.56a}$$
$$y_t = Hx_t + e_t, \tag{4.56b}$$

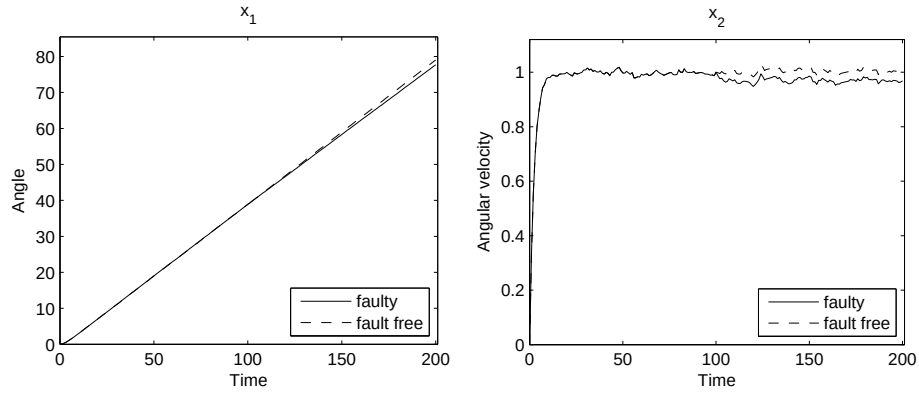
with the system matrices

$$F = \begin{pmatrix} 1 & 1 - e^{-T} \\ 0 & e^{-T} \end{pmatrix}, \qquad G^u = \begin{pmatrix} T - (1 - e^{-T}) \\ 1 - e^{-T} \end{pmatrix},$$
$$G^w = G^f = G^u, \qquad H = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

where $w_t \sim \mathsf{N}\left(0, \left(\frac{\pi}{180}\right)^2\right)$, $e_t \sim \mathsf{N}\left(0, \left(\frac{\pi}{180}\right)^2\right)$ and the sample interval used is $T = 0.4\,\mathrm{s}$.

4.4.2 Simulations

In the simulation, the system has been simulated for 200 samples in time with the unit step as input. After 100 fault-free samples, a constant fault is introduced with the magnitude of $\frac{2\pi}{180}$, which is present until the last sample. The fault can be seen as a torque disturbance which causes a drop in the angular velocity, x_2, and a slope change in the angle, x_1. With this setup, 2000 Monte Carlo (MC) simulations have been carried out. One instance of the state trajectory realizations is shown in Figure 4.1 where also the fault-free trajectory is shown using the same noise realization. The drop in angular velocity when the fault is present can be seen in Figure 4.1b. The angle, which is measured, has a small change in the slope which is hardly noticeable in Figure 4.1a. The fault detection method using



(a) *(b)*

Figure 4.1: *Plots of the state trajectories with and without fault influence for the* DC*-motor in* (4.56)*.*

smoothing, includes a time-varying term dependent on the estimation covariance from the Kalman filter. After a transient period the Kalman filter will become stationary which means that the estimation covariance converges. Figure 4.2 shows $\sqrt{\mathrm{tr}\, P^{(1)}}$, a measure of the state covariance, $P^{(1)}$. Note that $P^{(1)}$ converges already after approximately 15 samples and the smoothing method can then be considered time invariant.

4.4.3 Fault detection

Fault detection on the simulated data has been made both with and without making use of the known fault structure. Since the fault is a step disturbance, a basis of dimension one can be used to describe its behavior. For fault detection a sliding window of 8 samples is used and for the smoothed method the Kalman filter estimate from previous data is used as well. For the discussions below, some definitions have to be made. The alarm rate is calculated as

$$a_r(t) = \frac{\text{\# L(t)} > \gamma}{\text{\# MC-simulations}}, \tag{4.57}$$

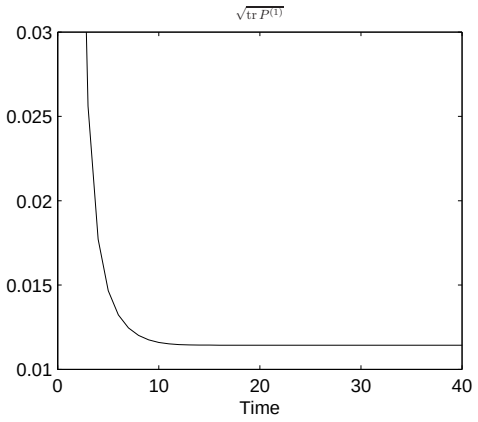


Figure 4.2: *$\sqrt{\mathrm{tr}\, P^{(1)}}$ is used as a measure of the state estimation covariance.*

where $L(t)$ is the test statistic and γ the threshold for detection. The probability of detection, P_D, is defined as the alarm rate during presence of fault. Probability of false alarm, P_{FA}, is defined as the alarm rate during fault-free periods.

A. No Knowledge about Fault Structure

First, the fault detection methods are used without assuming knowledge about the structure of the fault. The results can be seen in Figures 4.3a–4.3c. The average over the MC-simulations of the test statistics ((4.35) and (4.51)) is shown in Figure 4.3a. Note that the difference in average during the fault-free part between the methods is due to different degrees of freedom of the test statistic. The parity space method has 6 degrees of freedom whereas the smoothed method has 7. The alarm rate versus time is shown in Figure 4.3b. The false alarm rate has been set to 1% and it and the actual alarm rate during the fault-free parts of the simulation corresponds to this value. The method with smoothing outperforms the parity space approach, but the comparison is unfair since the estimation is based on a different amount of data. In Figure 4.3c, the Receiver Operation Characteristics (ROC) is plotted. The ROC-curve shows how the probability of detection for a fault varies versus the probability of false alarm. Guessing corresponds to a the straight line between $P_{FA} = P_D$. In this case the values for P_D is estimated one window after the entrance of the fault. This is done since the Kalman filter estimate, used in the smoothed method, will be influenced by the fault afterward. More about ROC-curves can be found in Kay (1998).

B. Using the Structure of the Faults

Secondly, the structure of the fault is used to enhance the fault detectability. Since it is here assumed that the fault is a step, a one dimensional basis function is used to describe the fault. It is discussed in Section 4.3.4 that the average of the test statistics should equal the degrees of freedom for the test statistic. Due to the parameterization, the degree of

freedom is 1 and the effect of this can be seen in Figure 4.3a where the averages of the test statistics ((4.35) and (4.51)) are 1 in the fault-free part. The lower average during the fault-free part makes the relative increase to the faulty part larger and the probability of detection is therefore higher than without structure. Figure 4.3b shows that the probability of detection is considerably increased for both methods due to the use of the fault structure. The ROC-curves in Figure 4.3c have therefore also shifted towards higher detection rates.

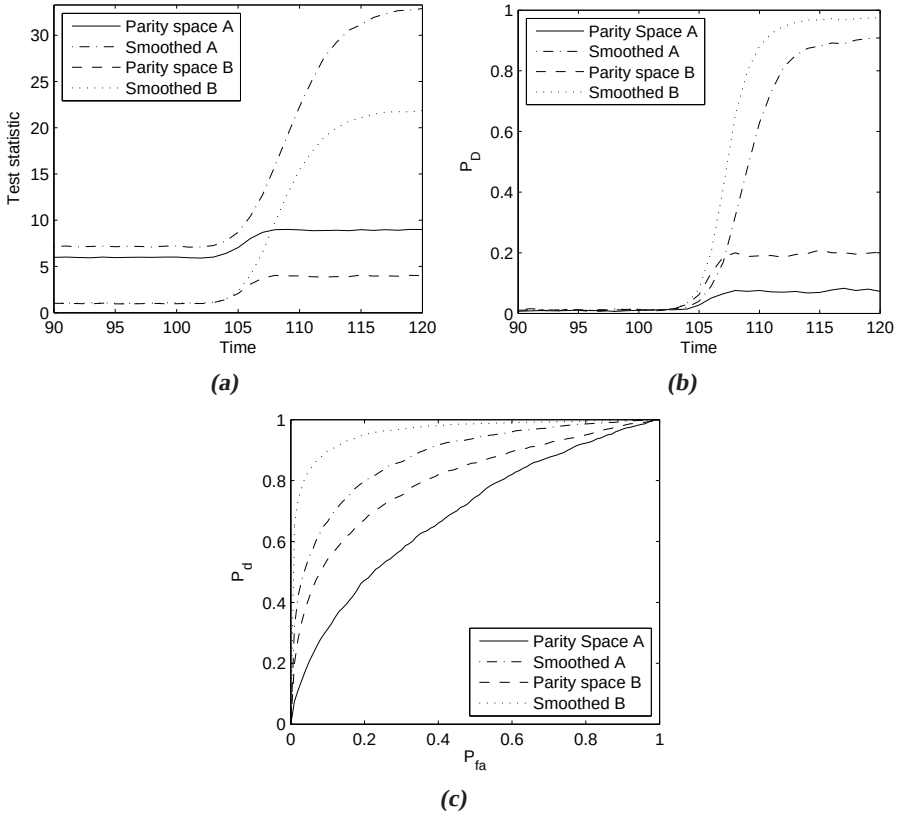


Figure 4.3: *The averages of the test statistics* (4.35) *and* (4.51) *are given in (a), the alarm rate in (b) and a plot of the* ROC *curve in (c). A and B denote with and without using the structure of the fault respectively.*

4.5 Conclusions

This chapter studied the initial state dependency for the GLR test for fault detection based on a batch of data. Previous approaches are either based on eliminating the initial state by projection (the parity space approach) or by using a prior computed from passed data (the sliding window approach). The minimum variance estimate of the initial state from the

data batch is derived. This, in itself, is shown in Section 4.3.2 to be algebraically identical to the parity space approach. However, in combination with the sliding window approach, it opens up the possibility to form a smoothed estimate. It is demonstrated in simulations that this improved performance considerably. Another contribution is to point out how a smooth parameterization of the fault profile can improve detection performance.

5

Modeling and Estimation of IMU:s

IN MANY SITUATIONS, the attitude and position of a system is of interest. For example, navigation and positioning of airplanes and cars are a mature research area today. Most systems rely partly on satellite navigation equipment, such as the Global Positioning System, GPS, and map-aided navigation systems that recognize the terrain, see for instance Bergman (1999), Karlsson (2005). To obtain complete position and attitude information, this information has to be combined with an inertial navigation system. This system contains three dimensional accelerometers and gyroscopes, and is sometimes referred to as an Inertial Measurement Unit, IMU.

Another application where an IMU is used is positioning of film cameras. The position of the camera has to be known in order to put virtual objects in the scene, which is known as augmented reality. The measurement from the IMU is combined with information from the image and in this way position and orientation can be obtained (Schön, 2006, Hol, 2005).

Human motion analysis is another area where IMU:s are used. There are scientists who wants to get a better understanding of the walking process and therefore studies how different body parts are rotated and accelerated. Athletes analyzes their movements in order to improve their performance while avoiding injury, see Luinge (2002), Roetenberg et al. (2005).

5.1 Sensors

The measurements in Chapter 6.1 are made with an IMU from Xsens Technologies, see (Xsens Technologies B.V.). This IMU, see Figure 5.1, consists of a three dimensional accelerometer, gyroscope and magnetometer (compass). All these sensors are based on the technology microelectromechanical systems, MEMS. That means they are small silicon based sensors of the strapdown type, that is, rigidly mounted onto the moving body. This

is in contrast to the old gimbaled gyroscopes and accelerometers, which always had the attitude of the fixed coordinate system. For a thorough review of strapdown inertial navigation and micro machined sensors, see Titterton and Weston (1997), Gardner (1994).

Figure 5.1: *The IMU MTi from Xsens Technologies which is used for the experiments in this thesis.*

Traditional accelerometers used a mass-spring-damper-system to measure the accelerations. The extension of the spring in steady state is proportional to the acceleration. The sensor used here is manufactured by Analog Devices (ADXL203) and has a measurement range of $\pm 1.7\,\mathrm{g}$. Although it is silicon based, the principle is the same as for the mass-spring-damper-system. A silicon mass is suspended with silicon springs between two fixed silicon plates. To read out the acceleration, the ratio of the capacitance between the fixed plate and the mass on both sides is used. (Gardner, 1994, Analog Devices)

The gyroscope is also manufactured by Analog Devices, model number ADXRS300. The gyroscope is a so called resonator gyroscope which is based on a vibrating silicon structure. There is a capacitive pickoff that senses Coriolis forces, which are dependent on the angular velocity.

The magnetometer is a Honeywell HMC1023, which is a magneto-resistive sensor. A magneto-resistive sensor is composed of thin plates in which the current lines are rotated depending on the magnetic field. The rotation of the current lines gives an increase in friction and the magnetic field can thereby be measured. (Gardner, 1994, Honeywell Solid State Electronics Center)

5.2 Modeling

The model describing the IMU consists of two parts: one describing how its body is moving (dynamic model) and one describing the measurements (measurement model).

To describe the behavior of the IMU two coordinate systems are used, one fixed to the earth and one fixed in the IMU.

5.2.1 Coordinate Systems

To describe a moving body system, two coordinate systems are used. One system, that is fixed to the world (Earth), referred to as the W-system, and one that is following the moving body, referred to as the B-system. Theoretically, the measurements from the IMU is made in a fixed inertial coordinate system. The rotation of the Earth introduces Coriolis forces, but since these effects are small, they are neglected here. Figure 5.2 illustrates the coordinate systems. The relation between the coordinate systems is described in two steps. First, the translation from the origin of the W-system to the origin of the B-system is described with the vector p. Second, the rotation (or attitude) from the W-system to the B-system is described by the unit quaternion q_{bw} and the corresponding rotation matrix $R(q_{bw})$. Unit quaternions can be used to describe attitude numerically more robust than with for example Euler angles. For more information about quaternions, see Appendix B.

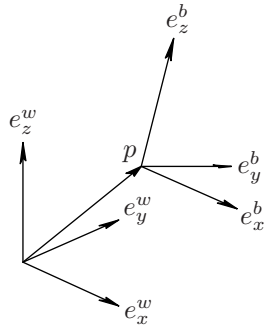


Figure 5.2: *The two coordinate systems describing the moving body system. The orthogonal unit vectors* (e_x^b, e_y^b, e_z^b) *are moving with the body and the orthogonal unit vectors* (e_x^w, e_y^w, e_z^w) *are fixed.*

5.2.2 Dynamic Model

The translation between the W-system and the B-system p is also called the position of the body. The dynamic model need to have the position p and the rotation q_{bw} as states. Since the accelerations of the B-system are measured, these have to be modeled as states as well as the velocity v^w. The state vector can then be chosen as $x = \begin{pmatrix} p^{wT} & v^{wT} & a^{wT} & {q_{bw}}^T \end{pmatrix}^T$, where p^w is translation from the W-system to the B-system, v^w is the velocity vector of the B-system relative to the W-system (described in the W-system), a^w is acceleration of the B-system relative to the W-system (described in the W-system), q_{bw} is the unit quaternion describing the orientation of the B-system.

The dynamics of the translation is easy, namely

$$\dot{p}^w = v^w$$
$$\dot{v}^w = a^w.$$

Since nothing is known about the derivative of the acceleration, this is modeled as noise,

$$\dot{a}^w = v_1,$$

where v_1 is process noise. The rotation of the B-system relative to the W-system is described using the unit quaternion, q_{bw}. The dynamics of the unit quaternion is derived in appendix B and the derivative is given in (B.26). The result is

$$\dot{q}_{bw} = \frac{1}{2} S(\omega_{bw}^b) q_{bw}, \tag{5.1}$$

where ω_{bw}^b denotes the angular velocity of the B-system relative to the W-system described in the B-system and S is a matrix which is dependent linearly on the elements in ω_{bw}^b. The angular velocity ω_{bw}^b is measured by the gyroscopes and should therefore be modeled as a state which is measured. However, to reduce the dimension of the state space and avoid the nonlinearity which would have been introduced then, the measurement of the angular velocity is used as an input in the dynamics. This approach would only be theoretical correct if ω_{bw}^b could be measured without noise. However, since the measurement noise from the gyroscopes, v_2, is small, the approximation done here is to use $\omega_{bw}^b + v_2$ as angular velocity.

$$\dot{q}_{bw} \approx \frac{1}{2} S(\omega_{bw}^b + v_2) q_{bw} \tag{5.2}$$

Using (B.26)

$$S(\omega_{bw}^b + v_2) q_{bw} = S(\omega_{bw}^b) q_{bw} + S'(q_{bw}) v_2,$$

where $S'(\star)$ is defined in (B.26). Then, (5.2) can be written as

$$\dot{q}_{bw} \approx \frac{1}{2} S(\omega_{bw}^b) q_{bw} + \frac{1}{2} S'(q_{bw}) v_2. \tag{5.3}$$

Thus, the total dynamic model is given by

$$\begin{pmatrix} \dot{p}^w \\ \dot{v}^w \\ \dot{a}^w \\ \dot{q}_{bw} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} S(\omega_{bw}^b) \end{pmatrix}}_{A} \begin{pmatrix} p^w \\ v^w \\ a^w \\ q_{bw} \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ I & 0 \\ 0 & \frac{1}{2} S'(q_{bw}) \end{pmatrix}}_{B} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{5.4}$$

5.2.3 Measurements

The sensors that are available in the system are gyroscopes, accelerometers and magnetometers (compass). The measurements from the gyroscopes are, as discussed above,

directly incorporated in the dynamics and are therefore not described in the measurement equations.

The accelerometers measures both the free acceleration of the body and the earth gravitational field described in the B-system. The measurement equation can therefore be written as

$$y_{a,t} = a_t^b - g^b + e_{a,t} = \underbrace{R(q_{bw})(a_t^w - g^w)}_{\triangleq h_a(q_{bw}, a_t^w)} + e_{a,t}, \tag{5.5}$$

where $R(\star)$ is the rotation matrix given in (B.21), a is the free acceleration, g is the gravity vector and e_a is the measurement noise.

The reading from the magnetometer is a normalized vector, $\bar{n}_{np}^b$, given in the B-system and pointing along the earth magnetic field. When using an ordinary compass, the reading will be two dimensional giving the heading towards the magnetic North Pole. Since this measurement is three dimensional, information is also given about the magnetic inclination. The inclination, also called dip angle, is the angle between the earth's tangent plane and the magnetic field vector. The magnetic inclination depends on the position on earth and is $0°$ at the magnetic equator and $90°$ at the magnetic poles. The inclination in Linköping, Sweden, is about $70°$. The heading information will of course be less accurate when the inclination approaches $90°$. For more information about the earth magnetic field, see (National Geomagnetism Program).

Since the earth magnetic field is very weak, it can easily be disturbed by an electric motor or ferromagnetic objects such as beams or other iron constructions. Such a disturbance can be represented by an unknown vector, here denoted $d_{m,t}$. The measurement equation can thus be written as

$$y_{m,t} = \bar{n}_{np}^b + d_{m,t} + e_{m,t} = \underbrace{R(q_{bw})\bar{n}_{np}^w}_{\triangleq h_m(q_{bw})} + d_{m,t} + e_m, \tag{5.6}$$

where $R(\star)$ is the rotation matrix given in (B.21), $\bar{n}_{np}$ is the normalized earth magnetic field vector and e_m is the measurement noise.

The total measurement equation becomes

$$y = \begin{pmatrix} y_a \\ y_m \end{pmatrix} = h(q_{bw}, a_t^w) + \underbrace{\begin{pmatrix} 0 \\ I \end{pmatrix}}_{\triangleq H^d} d_{m,t} + e_t. \tag{5.7}$$

5.2.4 Discrete-Time Dynamic Model

The dynamic model (5.4) is divided into two parts. One containing the translational dynamics of the B-system and the other one the attitude dynamics. The first part contains the states p^w, v^w and a^w and the second part q_{bw}. To simplify the notation and stress the time dependency q_{bw} is denoted q_t from now on.

Translational Dynamics

For the translational part of the system, the noise inputs to the system are assumed to be constant between the sampling intervals. Then the discrete time system can be written as

$$\begin{pmatrix} p^w_{t+1} \\ v^w_{t+1} \\ a^w_{t+1} \end{pmatrix} = F \begin{pmatrix} p^w_t \\ v^w_t \\ a^w_t \end{pmatrix} + G^v v_t, \tag{5.8}$$

where

$$F = e^{AT} \quad \text{and} \quad G^v = \int_0^T e^{At} B \, dt.$$

The matrices in (5.8) can be evaluated to

$$F = \begin{pmatrix} I & TI & \frac{T^2}{2} I \\ 0 & I & TI \\ 0 & 0 & I \end{pmatrix} \quad \text{and} \quad G^v = \begin{pmatrix} \frac{T^3}{6} I \\ \frac{T^2}{2} I \\ TI \end{pmatrix}.$$

Attitude Dynamics

The part of the system (5.4) describing the attitude is the last state, q_{bw}. The following derivations have been made in discussion with Schön (2005). Assume that the angular velocity ω^b_{bw} is constant during the sampling intervals, that is

$$\omega^b_{bw}(t) = \omega^b_{bw,k}, \quad \text{when } kT \leq t < (k+1)T.$$

Making this assumption, the attitude dynamics becomes time invariant during the sampling interval. Hence, considering the noise free dynamic model (5.1), the discrete time model is given by

$$q_{t+1} = e^{\frac{1}{2} S(\omega^b_{bw,t}) T} q_t. \tag{5.9}$$

To simplify calculations, the term $e^{\frac{1}{2} S(\omega^b_{bw,t}) T}$ has to be rewritten. Starting with the series expansion, we set

$$e^{\frac{1}{2} S(\omega^b_{bw,t}) T} = \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2} S(\omega^b_{bw,t}) T\right)^n}{n!} = \sum_{n=0}^{\infty} \left(\frac{T}{2}\right)^n \frac{1}{n!} S(\omega^b_{bw,t})^n. \tag{5.10}$$

The matrix $S(\omega^b_{bw,t})$ has an interesting property, namely

$$S(\omega^b_{bw,t})^2 = -\|\omega^b_{bw,t}\|^2 I. \tag{5.11}$$

Using this result and the series expansion for the cosine, the even part of the series expansion (5.10) can be written as

$$\sum_{n=0}^{\infty} \left(\frac{T}{2}\right)^{2n} \frac{1}{(2n)!} (-1)^n S(\omega^b_{bw,t})^{2n} = \cos\left(\frac{\|\omega^b_{bw,t}\| T}{2}\right) I \tag{5.12}$$

Using knowledge about the series expansion of sine, the odd part can be written as

$$\begin{aligned}\sum_{n=0}^{\infty}\left(\frac{T}{2}\right)^{2n+1}\frac{1}{(2n+1)!}S(\omega_{bw,t}^{b})^{2n+1}\\ &= \frac{T}{2}S(\omega_{bw,t}^{b})\sum_{n=0}^{\infty}\left(\frac{T}{2}\right)^{2n}\frac{1}{(2n+1)!}(-1)^{n}\frac{\|\omega_{bw,t}^{b}\|^{2n+1}}{\|\omega_{bw,t}^{b}\|}\\ &= \frac{T}{2}S(\omega_{bw,t}^{b})\frac{\sin\left(\|\omega_{bw,t}^{b}\|T/2\right)}{\|\omega_{bw,t}^{b}\|}. \quad (5.13)\end{aligned}$$

Combining the odd and even parts of the series expansion, the unit quaternion dynamics can be written as

$$q_{t+1} = \left(\cos\left(\frac{\|\omega_{bw,t}^{b}\|T}{2}\right)I_4 + \frac{\sin\left(\frac{\|\omega_{bw,t}^{b}\|T}{2}\right)}{\|\omega_{bw,t}^{b}\|}S(\omega_{bw,t}^{b})\right)q_t. \quad (5.14)$$

If the sample rate is high enough so that the product $\|\omega_{bw,t}^{b}\|T$ is small, the small angle approximations can be used to simplify the dynamic equation to

$$q_{t+1} = \left(I_4 + \frac{T}{2}S(\omega_{bw,t}^{b})\right)q_t. \quad (5.15)$$

Now, reintroducing the measurement noise on $\omega_{bw,t}^{b}$, (5.15) can be written as

$$q_{t+1} = \underbrace{\left(I_4 + \frac{T}{2}S(\omega_{bw,t}^{b})\right)}_{\triangleq F_t}q_t + \underbrace{\frac{T}{2}S'(q_t)}_{\triangleq G_t^{v}}v_t. \quad (5.16)$$

5.2.5 Linearization

To estimate the state trajectory using the EKF and to use linear theory for fault detection, the system must be linearized. The EKF is described in Section 2.5. The dynamic equations (5.8) and (5.16) are already linear time-varying equations, but the measurement equation is nonlinear. It is therefore linearized with a first order Taylor series. Starting with the accelerometer part of the measurement equation, it can be approximated as

$$y_{a,t} = h_a(q_t) + e_{a,t} \approx h_a(\hat{q}_{t|t-1}) + H_{a,t}(q_t - \hat{q}_{t|t-1}) + e_{a,t}, \quad (5.17)$$

where

$$H_{a,t} = \left.\frac{\partial h_a(q)}{\partial q}\right|_{q=\hat{q}_{t|t-1}}. \quad (5.18)$$

The magnetometer part of the measurement equation is approximated as

$$\begin{aligned}y_{m,t} &= h_m(q_t, a_t^{w}) + e_{m,t}\\ &\approx h_m(\hat{q}_{t|t-1}) + H_t^{m}(q_t - \hat{q}_{t|t-1}) + H_{d,t}d_{m,t} + e_{m,t}, \quad (5.19)\end{aligned}$$

where

$$H_t^m = \left.\frac{\partial h_m(q)}{\partial q}\right|_{q=\hat{q}_{t|t-1}}.$$

To form the linearized system, a new measurement variable is computed with known information as

$$\tilde{y}_t \triangleq y_t - \begin{pmatrix} h_a(\hat{q}_{t|t-1}) \\ h_m(\hat{q}_{t|t-1}) \end{pmatrix} + \begin{pmatrix} H_t^a \\ H_t^m \end{pmatrix} \hat{q}_{t|t-1} = \begin{pmatrix} H_t^a \\ H_t^m \end{pmatrix} q_t + \begin{pmatrix} 0 \\ I \end{pmatrix} d_{m,t} + e_t. \tag{5.20}$$

Then the linearized system can now be written as

$$q_{t+1} = F_t q_t + G_t^v v_t \tag{5.21a}$$

$$\tilde{y}_t = \underbrace{\begin{pmatrix} H_t^a \\ H_t^m \end{pmatrix}}_{\triangleq H_t} q_t + \underbrace{\begin{pmatrix} 0 \\ I \end{pmatrix}}_{\triangleq H_t^d} d_{m,t} + e_t. \tag{5.21b}$$

The Jacobians in the measurement equation (5.21b) are

$$H_t^a = 2g \begin{pmatrix} -q_2 & q_3 & -q_0 & q_1 \\ q_1 & q_0 & q_3 & q_2 \\ q_0 & -q_1 & -q_2 & q_3 \end{pmatrix} \tag{5.22}$$

and

$$H_t^{mT} = \begin{pmatrix} n_x q_0 + n_y q_3 - n_z q_2 & -n_x q_3 + n_y q_0 + n_z q_1 & n_x q_2 - n_y q_1 + n_z q_0 \\ n_x q_1 + n_y q_2 + n_z q_3 & n_x q_2 - n_y q_1 + n_z q_0 & n_x q_3 - n_y q_0 - n_z q_1 \\ -n_x q_2 + n_y q_1 - n_z q_0 & n_x q_1 + n_y q_2 + n_z q_3 & n_x q_0 + n_y q_3 - n_z q_2 \\ -n_x q_3 + n_y q_0 + n_z q_1 & -n_x q_0 - n_y q_3 + n_z q_2 & n_x q_1 + n_y q_2 + n_z q_3 \end{pmatrix}, \tag{5.23}$$

where the magnetic field vector in the W-system has the following coordinates $\bar{n}_{np}^f = \begin{pmatrix} n_x & n_y & n_z \end{pmatrix}^T$.

5.3 Attitude Estimation

In order to use the accelerometers for attitude information the direction of the gravity field is used. However, it is impossible to distinguish between the acceleration of the body and the gravitational acceleration only using accelerometers. Therefore, we assume that the body is not accelerated or that the acceleration is so small that it can be approximated by noise. The noise generated by movements of the sensor can also be modeled, see Section 6.2. The measurement equation due to the accelerometers then becomes

$$y_t^a = -g^b + e_a = -R(q_t) g^w + e_t^a, \tag{5.24}$$

where g^b and g^w are the gravitational acceleration in the B-system and the W-system respectively and e_t^a is measurement noise. The rotation matrix $R(q_{bw})$ is given in (B.21) and since $g^w = \begin{pmatrix} 0 & 0 & -g \end{pmatrix}^T$ the measurement equation can be written as

$$y_t^a = g \begin{pmatrix} 2(q_1 q_3 - q_0 q_2) \\ 2(q_2 q_3 + q_0 q_1) \\ q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} + e_t^a, \tag{5.25}$$

where $g \approx 9.82$ is the constant of gravity.

Attitude estimation is done with an EKF. An additional normalization step for the unit quaternion is needed after both time update and measurement update. That is, add the operation

$$\hat{q} := \hat{q}/\|\hat{q}\|$$

after each update.

6

Measurement Data and Motion Modeling

THE WORK WITH REAL-WORLD applications requires extensive amounts of measuring. This section provides illustrative data sets which shows the characteristics of the sensors in the IMU.

6.1 Measurement Data

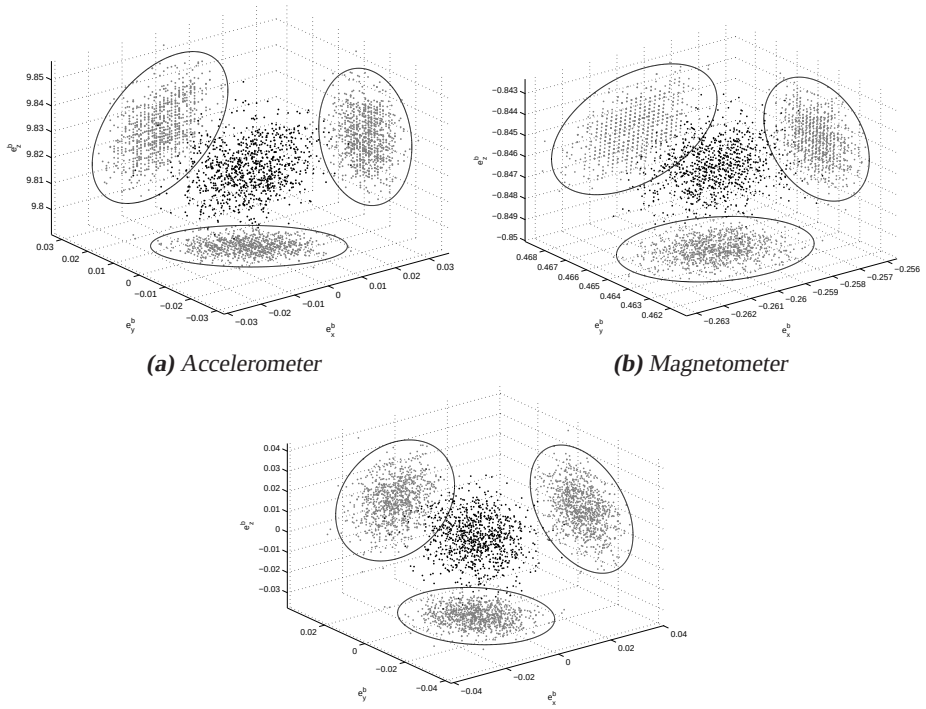
To show the performance of the disturbance detection algorithms described in previous chapter and to show the characteristics of the sensors, some measurement data is collected. The data is collected under different circumstances, for example when the sensor is lying still, is slowly rotated, with and without disturbances. The datasets are presented in Table 6.1.

Table 6.1: *Data sets collected using the IMU.*

Data Set	Disturbance	Movement
UD1	Undisturbed	Lying still
D1	Magnetometer disturbance	Lying still
D2	Light magnetometer disturbance	Lying still
UD2	Undisturbed	Holding still in hand
UD3	Undisturbed	Rotated around e_x^b
D3	Magnetometer disturbance	Rotated around e_x^b
D4	Acceleration disturbance	Moved without rotation

The purpose of the first data set, UD1, is to explore the noise characteristics of the sensors. The IMU is here lying still on a table and data is collected during 9 seconds with a sampling rate of 100 Hz. The orientation of the sensor is set so that the e_z^b-axis is pointing outwards from the earth along the gravity field (perpendicular to the table). The

acceleration measurements from the undisturbed data set UD1 are plotted in Figure 6.1a. Note that gravity is observed as a positive acceleration since only real forces acting on the sensor can be observed. The force observed is thus the normal force given by the table rather than gravity. Furthermore, the confidence ellipsoids of the measurements are quite circular indicating that the noise has low spatial correlation between the components. The confidence ellipsoids are plotted on the distance 3σ which, if the distribution is Gaussian, means that 1.1% of the measurements should be outside of this ellipsoid. The magnetometer readings are presented in Figure 6.1b. Also these confidence ellipsoids are circular and thus indicate low spatial correlation between the components in the magnetometer readings.



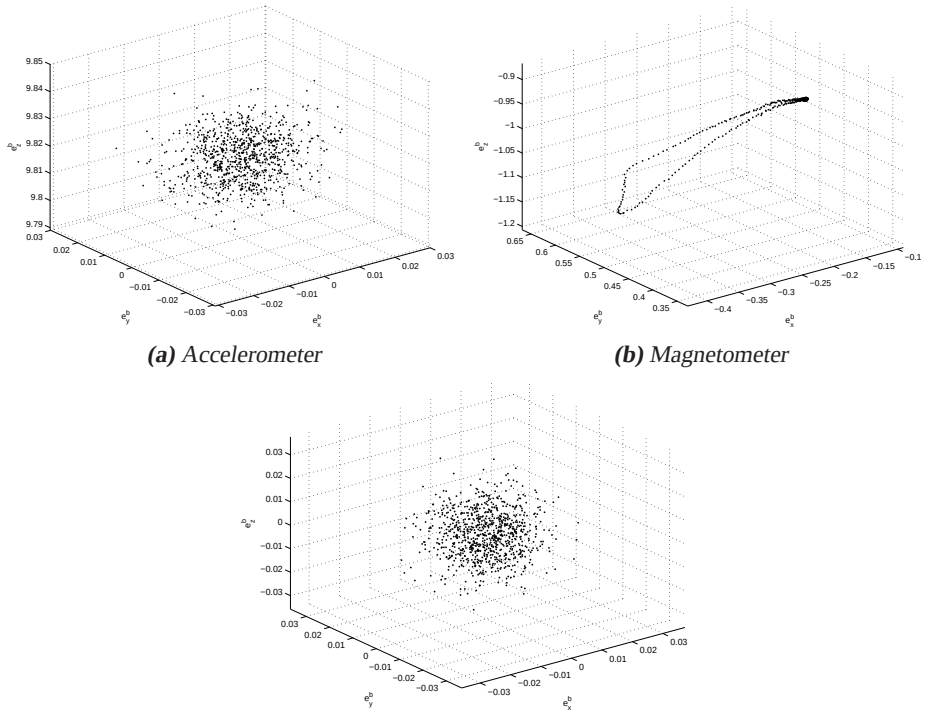
(a) *Accelerometer* ***(b)*** *Magnetometer*

(c) *Gyroscope*

Figure 6.1: *Accelerometer (a), magnetometer (b) and gyroscope (c) measurements, all from data set UD1. The ellipses represent confidence ellipsoids on the distance* 3σ*.*

The disturbed data set, D1, also represents a case where the IMU is lying still on a table, like in data set UD1. However, in this set, the magnetometer is disturbed by a metal object approaching the sensor. The metal object is here represented by a pair of scissors and the closest distance is approximately 5 cm. As can be seen in Figure 6.2a, the acccelometer readings are similar to the ones given in data set UD1. The magnetometer readings should nominally be a Gaussian cloud around one point, but the disturbance makes them follow a trajectory given in Figure 6.2b. Each component of the magnetome-

ter measurement is plotted separately in Figure 6.3. Note that the profile of the disturbance is a rather smooth function which will be used in the disturbance detection later.



(a) *Accelerometer*

(b) *Magnetometer*

(c) *Gyroscope*

Figure 6.2: *Accelerometer (a), magnetometer (b) and gyroscope (c) measurements, all from data set D1.*

The experiment setup for data set D2 is similar to D1. The difference is that the disturbance is much weaker. The disturbance was also here caused by a pair of scissors, but the closest distance was approximately 30 cm. The components of the magnetometer reading are plotted in Figure 6.4.

In applications, it is not often the case that the IMU is lying still on a table. For instance, in one application it is carried around by a person (Roetenberg et al., 2005). To illustrate the influence of human motion on the sensor readings, data set UD2 is collected when the sensor is handheld and the person holding it is trying to hold it still. Figure 6.5 shows the measurements, the characteristics of these are discuessed in Section 6.2.

Having seen data from the relatively still measurements in data set UD2, next step is to make a rotation. Data set UD3 is collected when the IMU is rotated around the e_x^b-axis. The accelerometer and magnetometer readings form a circle as shown in Figure 6.6.

In data set D3, the setup is similar to UD3, but the magnetometer is disturbed by a scissor. The accelerometer readings are unaffected as seen in Figure 6.7a, but the magnetometer is severely disturbed which can be seen in Figure 6.7b.

The data set D4 has been collected when the IMU is accelerated. The orientation of

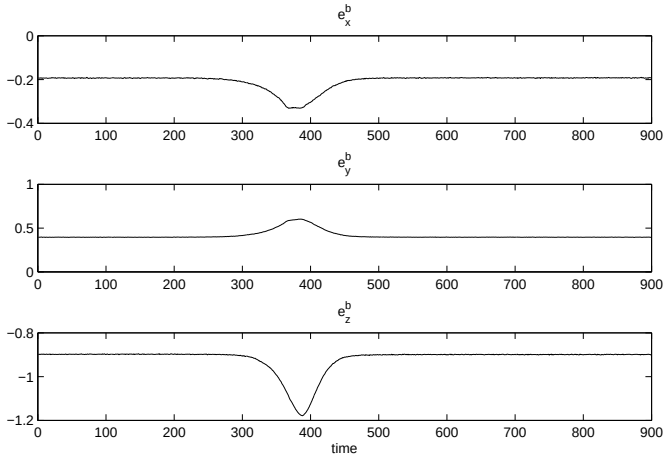


Figure 6.3: *The components of the disturbed magnetometer readings in data set D1.*

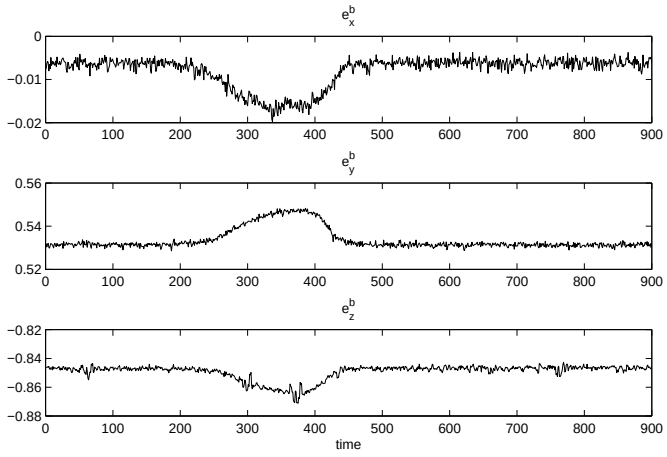
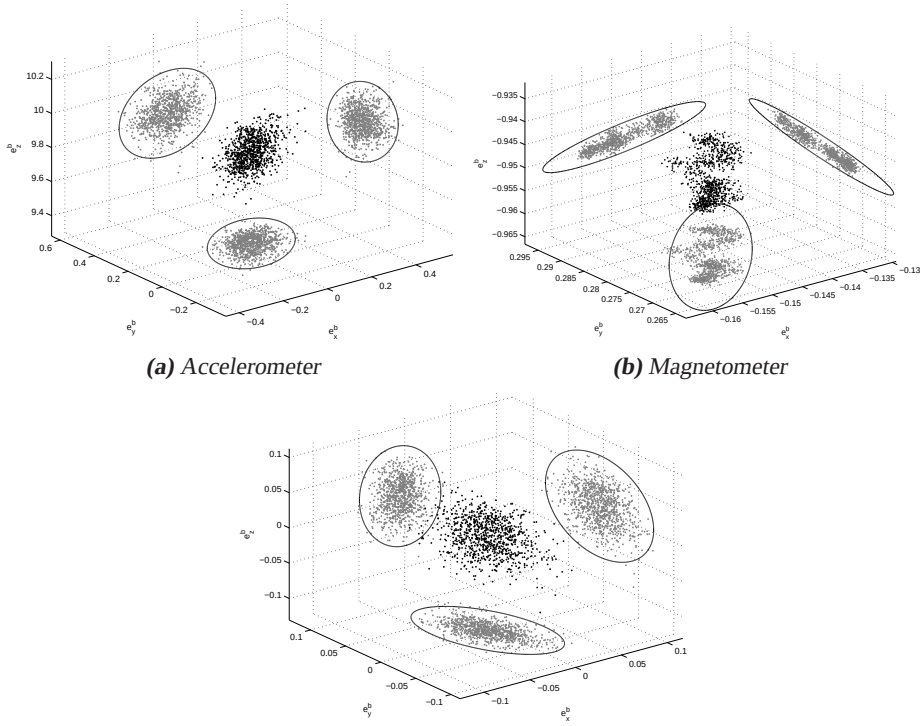


Figure 6.4: *The components of the disturbed magnetometer readings in data set D2.*

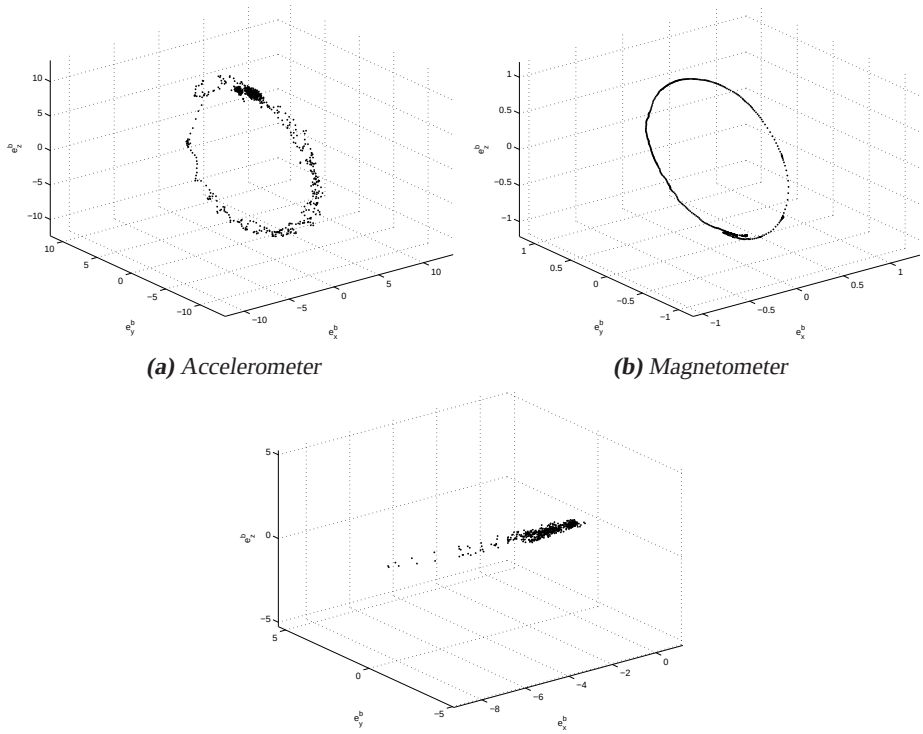


(a) *Accelerometer*

(b) *Magnetometer*

(c) *Gyroscope*

Figure 6.5: *Accelerometer (a), magnetometer (b) and gyroscope (c) measurements from data set UD2. The ellipses represent confidence ellipsoids on the distance* 3σ.

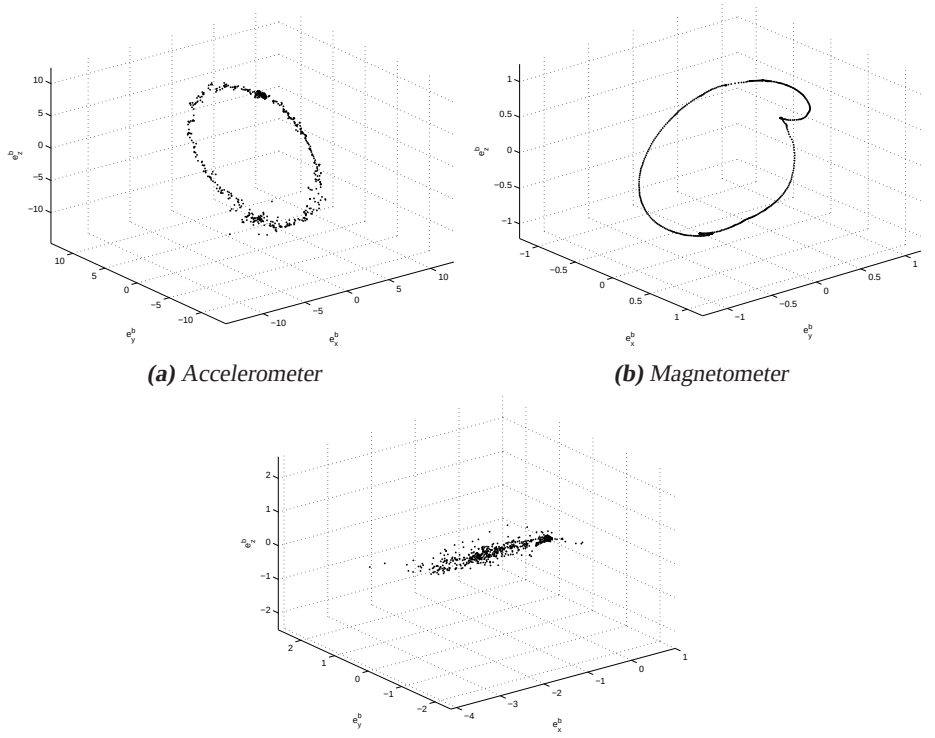


(a) *Accelerometer*

(b) *Magnetometer*

(c) *Gyroscope*

Figure 6.6: *Accelerometer measurements (a), magnetometer measurements (b) and gyroscope measurements (c) from data set UD3.*



(a) *Accelerometer*

(b) *Magnetometer*

(c) *Gyroscope*

Figure 6.7: *Accelerometer measurements (a), magnetometer measurements (b) and gyroscope measurements (c) from data set D3. Observe that the point of view for the magnetometer is different from the others in order to show the disturbance.*

the sensor is kept constant while it is accelerated along axis e_x^b. The time plot of the accelerometer readings are shown in Figure 6.8.

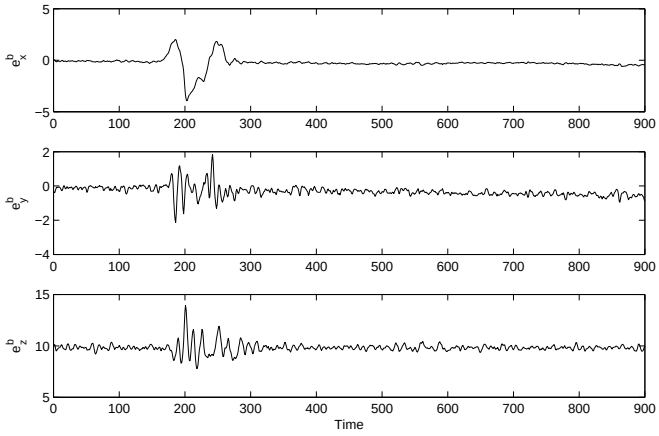


Figure 6.8: *The disturbed accelerometer measurements from data set D4.*

6.2 Noise Model

In this section, the noise is studied and noise models are built when necessary. To examine the noise characteristics of the sensors, the measurements from data set UD1 are examined. The measurement equations are given by (5.5) and (5.6). Also the measurements from the gyroscope are treated in this section even though this measurement is an input to the model. Since there are no rotation, physical acceleration or disturbance in data set UD1, the measurements will only contain noise and a constant contribution from gravity and the earth magnetic field. The measurement data was previously presented in Figure 6.1. Here, the auto covariance functions of the measurements are shown in Figure 6.9. The measurement noise is close to white since the covariance functions show small correlation in time. The spatial distribution of the measurements is close to Gaussian as can be seen in Figure 6.10. The noise characteristics of the accelerometer can thus be concluded as

$$\begin{aligned} e_{a,t} &\approx \mathsf{N}(0, R_a) \\ \mathsf{E}(e_{a,t} e_{a,t-\tau}^T) &\approx R_a \delta(\tau), \end{aligned}$$

where R_a is estimated from the data to be

$$\hat{R}_a = 10^{-4} \begin{pmatrix} 0.4524 & -0.1873 & 0.1594 \\ -0.1873 & 0.3617 & -0.0900 \\ 0.1594 & -0.0900 & 0.7855 \end{pmatrix} \tag{6.1}$$

and is diagonal dominant. The noise characteristics of the magnetometer can be concluded as

$$\begin{aligned} e_{m,t} &\approx \mathsf{N}(0, R_m) \\ \mathsf{E}(e_{m,t} e_{m,t-\tau}^T) &\approx R_m \delta(\tau), \end{aligned}$$

where R_m is estimated from the data to be

$$\hat{R}_m = 10^{-5} \begin{pmatrix} 0.1025 & -0.0015 & 0.0122 \\ -0.0015 & 0.0707 & -0.0049 \\ 0.0122 & -0.0049 & 0.0865 \end{pmatrix} \tag{6.2}$$

and is also diagonally dominant. Also for the gyroscope, the noise can be approximated with the following model

$$\begin{aligned} e_{\omega,t} &\approx \mathsf{N}(0, Q_\omega) \\ \mathsf{E}(e_{\omega,t} e_{\omega,t-\tau}^T) &\approx Q_\omega \delta(\tau), \end{aligned}$$

where Q_ω is estimated to be

$$\hat{Q}_\omega = 10^{-4} \begin{pmatrix} 0.5713 & 0.0004 & -0.0167 \\ 0.0004 & 0.7255 & 0.0308 \\ -0.0167 & 0.0308 & 0.9364 \end{pmatrix}. \tag{6.3}$$

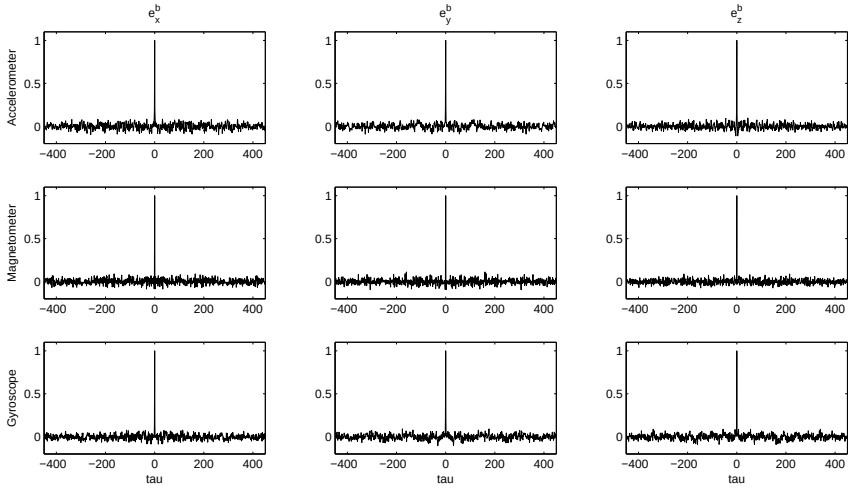


Figure 6.9: *Auto covariance for the measurements from data set UD1. Accelerometer, magnetometer and gyroscope measurements are presented.*

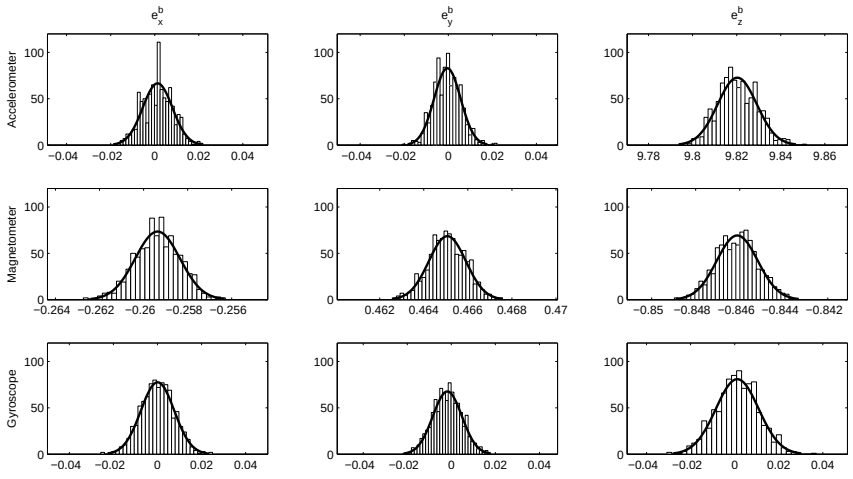


Figure 6.10: *Histograms for the measurements from data set UD1. Accelerometer, magnetometer and gyroscope measurements are presented, together with a Gaussian approximation.*

6.3 Motion Model for IMU Held in Hand (UD2)

Human motion is seen in data set UD2, where a person is trying to hold the IMU at rest by hand. Since it is hard to keep the sensor completely at rest, the motion of the hand will affect the measurements. This can be modeled as a stochastic process, that is, colored noise. First, consider the case where the motion is modeled as white Gaussian noise as in Section 6.2. Then, the covariance matrices estimated from data set UD2 becomes

$$\hat{R}_a = \begin{pmatrix} 0.0053 & 0.0025 & 0.0003 \\ 0.0025 & 0.0049 & -0.0018 \\ 0.0003 & -0.0018 & 0.0069 \end{pmatrix} \tag{6.4a}$$

$$\hat{R}_m = 10^{-4} \begin{pmatrix} 0.2064 & 0.2053 & 0.0339 \\ 0.2053 & 0.2757 & 0.0514 \\ 0.0339 & 0.0514 & 0.0170 \end{pmatrix} \tag{6.4b}$$

$$\hat{Q}_\omega = 10^{-3} \begin{pmatrix} 0.2192 & -0.1136 & -0.0485 \\ -0.1136 & 0.6585 & -0.0131 \\ -0.0485 & -0.0131 & 0.5778 \end{pmatrix}. \tag{6.4c}$$

As was seen in Figure 6.5a, where the acceleration measurements were plotted, these have greater covariance compared to data set UD1. The spatial correlation still seems to be quite low since the covariance ellipses are fairly circular. The magnetometer measurements, shown in Figure 6.5b, also have greater covariance compared to data set UD1. Here, the covariance matrices are not circular, indicating that there are spatial correlation between the components of the sensor. Trying to hold the sensor still by hand, it is easy to drift a little bit in attitude. This will give a linear trend in the accelerometer and magnetometer measurement. These are therefore removed before computing the autocorrelation function of the measurements, which are shown in Figure 6.11. It can be seen that the accelerometer and gyroscope readings are fairly white whereas the magnetometer readings are strongly correlated in time. The distributions of the measurements, plotted in Figure 6.12, show that the accelerometer measurements are close to Gaussian whereas at least one of the magnetometer components seems non-Gaussian. The conclusion drawn from this is that the movements in the accelerometer measurements can be approximated by white Gaussian noise, whereas the magnetometer measurements needs to be dynamically modeled.

The natural choice of model structure for the magnetometer measurements is an autoregressive, AR, model since only a time series of data is available. The AR-model can be written as

$$y_t = \frac{1}{1 + a_1 q^{-1} + a_2 q^{-2}} e_t. \tag{6.5}$$

Before modeling, average and linear trends are removed from data. For model order selection, the fit (see Ljung (2006)) for the output to the 5 step ahead prediction is evaluated. For the e_x^b and e_y^b axes it does not seem possible to gain much by choosing a model of higher order than 2. The model fit obtained is slightly more than 40%. The situation for the e_z^b-axis is different and it is hard to get good fit with any model order. One possible reason for this difference could be that the person holding the IMU is shaking more in the e_x^b, e_y^b-plane than in the e_z^b direction. This can also be seen in Figure 6.6b. The estimated model parameters are presented in Table 6.2 with model fit for the 5 step prediction.

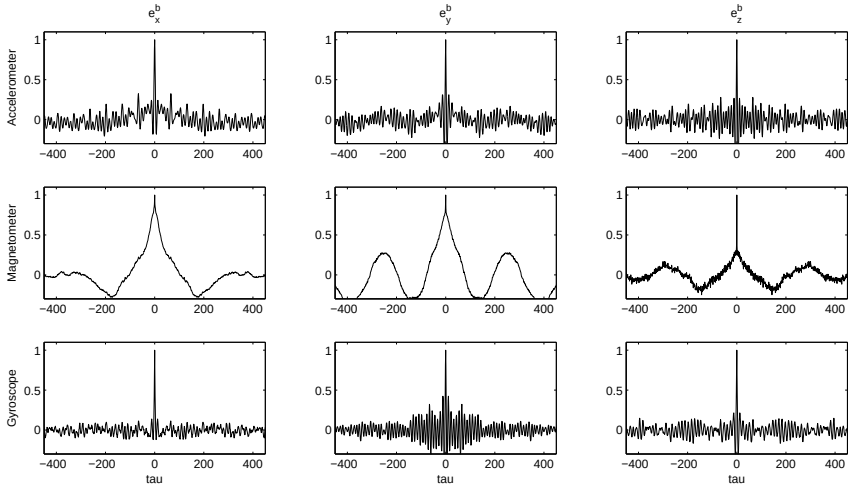


Figure 6.11: *Auto covariance for the measurements from data set UD2.*

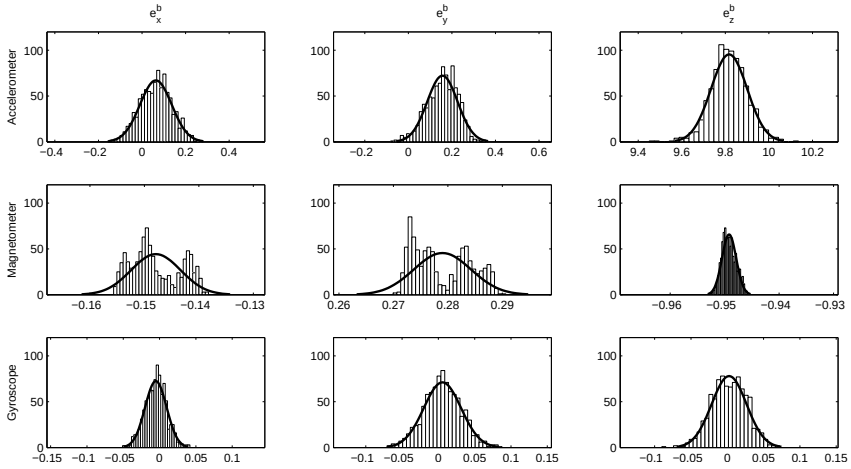


Figure 6.12: *Distributions for the measurements from data set UD2.*

Table 6.2: *Noise model parameters, estimated from magnetometer measurements in data set UD2.*

Axis	a_1	a_2	**Model fit**	**Poles**
e_x^b	-0.6494	-0.3348	43.16%	$0.9882, -0.3388$
e_y^b	-0.6268	-0.3589	40.92%	$0.9895, -0.3627$
e_z^b	-0.4794	-0.3339	-4.90%	$0.8653, -0.3859$

The AR-model can also be written as a state space model, the observability canonical form (see Section 2.1.2) is

$$x_{t+1} = \begin{pmatrix} -a_1 & 1 \\ -a_2 & 0 \end{pmatrix} x_t + \begin{pmatrix} 0 \\ e_t \end{pmatrix} \tag{6.6a}$$

$$y = \begin{pmatrix} 1 & 0 \end{pmatrix} x_t. \tag{6.6b}$$

The auto correlation functions for the AR-model residuals are shown in Figure 6.13 where it can be seen that they are fairly white. Compare with the plots in Figure 6.11 to see the difference by using a movement model. The distributions of the residuals are shown in Figure 6.14, using the motion model these can be approximated as white and Gaussian.

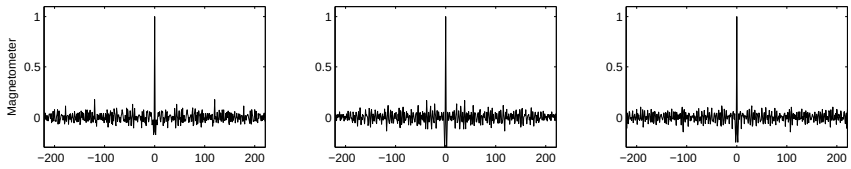


Figure 6.13: *Auto correlation functions for the magnetometer residuals using motion model.*

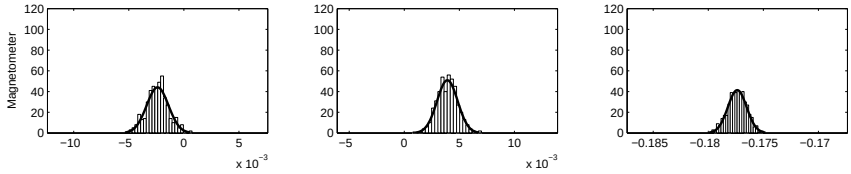


Figure 6.14: *The distributions of the residuals for the magnetometer, using motion model.*

7

Disturbance Detection for IMU

THE MAGNETOMETER IS USED to measure the earth magnetic field which is very weak. If the sensor is close to a ferromagnetic object or another magnetic field it is, as said before, easily disturbed. Such a disturbance must be detected to avoid erroneous attitude estimation. The accelerometer is affected both by gravity and body accelerations. For attitude estimation, the direction of the gravity vector is used. Body accelerations are therefore seen as disturbances that must be detected.

7.1 IMU Model and Detection Algorithm

This section will give a summary of the attitude model used for disturbance detection on the magnetometers and accelerometers. The total derivation of the model is given in Chapter 5. The linearized state space model is given by

$$q_{t+1} = \underbrace{\left(I_4 + \frac{T}{2} S(\omega^b_{bw,t})\right)}_{\triangleq F_t} q_t + \underbrace{\frac{T}{2} S'(q_t)}_{\triangleq G^v_t} v_t \tag{7.1a}$$

$$\tilde{y}_t = \begin{pmatrix} \tilde{y}_{1,t} \\ \tilde{y}_{2,t} \end{pmatrix} = \underbrace{\begin{pmatrix} H^a_t \\ H^m_t \end{pmatrix}}_{\triangleq H_t} q_t + \underbrace{I}_{\triangleq H^d_t} \begin{pmatrix} d_{a,t} d_{m,t} \end{pmatrix} + e_t, \tag{7.1b}$$

where the matrices H^a_t and H^m_t are given by (5.22) and (5.23) and

$$S(\omega^b_{bw,t}) = \begin{pmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{pmatrix}, \tag{7.2}$$

where $\omega^b_{bw,t} = \begin{pmatrix}\omega_x & \omega_y & \omega_z\end{pmatrix}^T$. The set of sensors, $\mathcal{S}$, are then $\{\tilde{y}_{1,t}, \tilde{y}_{2,t}\}$ which will be used in the detection filter algorithm. The batched form discussed in Section 2.1.3 will have the following form

$$\tilde{\mathbb{Y}} = \mathcal{O}_t x_{t-L+1} + \bar{H}^u_t \mathbb{U} + \bar{H}^d_t \mathbb{D} + \bar{H}^v_t \mathbb{V} + \mathbb{E}. \tag{7.3}$$

The linearization is done around a predicted trajectory, that is, given an estimate of the initial state in the window the trajectory is predicted in the window only using (7.1a).

If the fault is parameterized, it is assumed that the profile of the disturbance is a smooth function with respect to time. Thus, the disturbance could be parameterized as in Section 2.2. Each dimension in the disturbance vector is modeled separately, so $\bar{H}^{d^i}_t$ denotes the stacked system matrix for dimension i. This matrix is built up as (2.10) but using the i:th column of H. The stacked system can now be written as

$$\tilde{\mathbb{Y}} = \mathcal{O}_t x_{t-L+1} + \sum_{i=1}^{n_d} \bar{H}^{d^i}_t \mathbb{D}^i + \bar{H}^v_t \mathbb{V} + \mathbb{E}, \tag{7.4}$$

where $\mathbb{D}^i = \left({d^i_{t-L+1}}^T, \ldots, {d^i_t}^T\right)^T$. Each component of the disturbance is then modeled as

$$d^i_t = \phi^T_{i,t}\theta_i. \tag{7.5}$$

The influence of the disturbances can now be described as

$$\sum_{i=1}^{n_d} \bar{H}^{d^i}_t \mathbb{D}^i = \underbrace{\left(\bar{H}^{d^1}_t \cdots \bar{H}^{d^{n_d}}_t\right) \operatorname{diag}\left(\phi^T_{1,t} \cdots \phi^T_{n_d,t}\right)}_{\triangleq \bar{H}^\theta} \underbrace{\begin{pmatrix}\theta_1 \\ \vdots \\ \theta_{n_d}\end{pmatrix}}_{\triangleq \Theta}. \tag{7.6}$$

In the tests in Section 7.2, a third order basis is chosen for the fault. A plot of the basis functions are shown in Example 2.1, Figure 2.1a.

The state estimation filter, including the disturbance detection, can be described as in Algorithm 7.1. The fault parameterization, described above can either be used or not in step 3.

7.2 Test Results

The detection method presented in Section 7.1 is evaluated on the data sets presented in Section 6.1. The algorithm is tested both with and without parameterization of the faults.

7.2.1 Linearized Hypothesis Testing

The detection filter described in Algorithm 7.1 without fault parameterization is used throughout this section. The window length is 10 samples and the threshold for detection is set to a confidence level of 1 %.

Algorithm 7.1 Detection Filter

1. Time update according to the EKF in Algorithm 2.2.
2. Measurement update using non-disturbed sensors. This is done according to the EKF in Algorithm 2.2, but the measurement equation is limited to the set of non-disturbed sensors, $\mathcal{S}$. That is

$$y = \begin{pmatrix} \vdots \\ h_i(q_t) \\ \vdots \end{pmatrix}, \quad i \in \mathcal{S}. \tag{7.7}$$

3. Detection of disturbed sensors. Detection is done either according to the parity space method described in Section 4.3.1 or using the smoothing approach in Section 4.3.3. The set $\mathcal{S}$ is updated accordingly.

Data set UD1

The first test is done on the undisturbed data set UD1. In order to tune the detection filter, the noise covariances of the sensors have to be known. The noise covariances are estimated in Section 6.2 and given by (6.1), (6.2) and (6.3). Note that the noise of the gyroscope is treated as process noise since the gyroscope signal is modeled as an input. Using these noise covariances, the test statistic for the magnetometer is shown in Figure 7.1 and the accelerometer in Figure 7.2. The test statistic for the parity space method and smoothing method has 26 and 30 degrees of freedom respectively. The degrees of freedom should theoretically be the average of the Chi-square distributed test statistic, see discussion in Section 4.3.4. The averages for the test applied to the magnetometer and accelerometer correspond well to the theoretical values which can be seen in Table 7.1.

Table 7.1: *Test statistic data for data set UD1, μ is the theoretical average, $\hat{\mu}$ the obtained average and h the threshold for detection.*

Sensor	**Method**	μ	$\hat{\mu}$	h
Magnetometer	parity	26	26.75	45.6
Magnetometer	smoothing	30	30.47	50.9
Accelerometer	parity	26	25.28	45.6
Accelerometer	smoothing	30	28.68	50.9

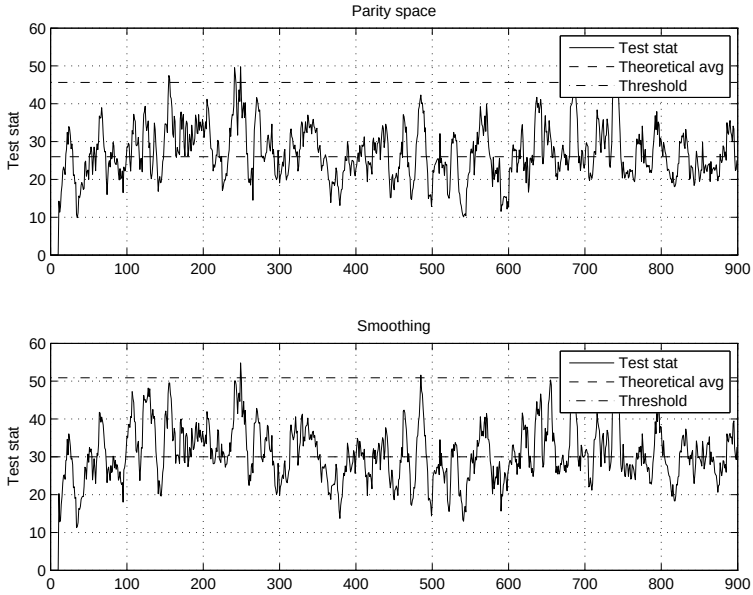


Figure 7.1: *The test statistics for the magnetometer in data set UD1. The averages are 26.75 for parity space and 30.47 for smoothing.*

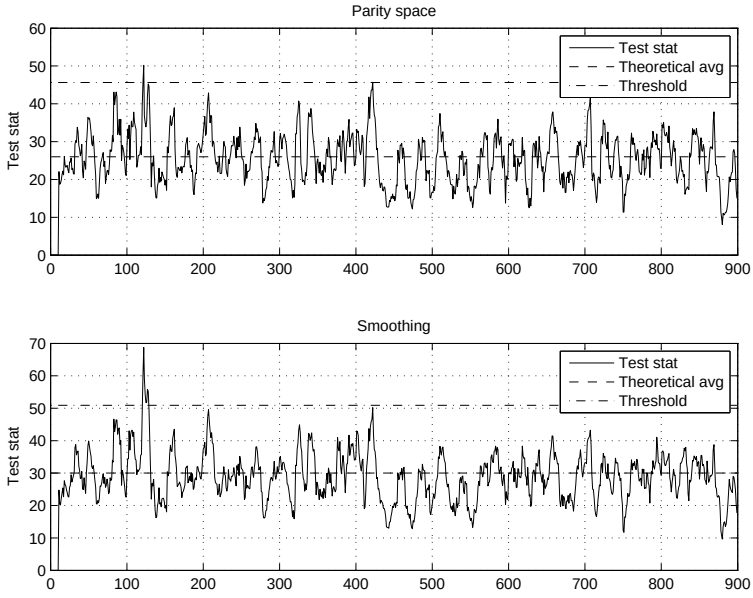


Figure 7.2: *The test statistics for the accelerometer in data set UD1. The averages are 25.28 for parity space and 28.68 for smoothing.*

Data set D1

A strong magnetometer disturbance is present in data set D1, otherwise the setup is the same as in data set UD1. The influence of the disturbance on the test statistic of the magnetometer can be seen in Figure 7.3. It can be noted that the test statistic for the smoothing method reacts more than for the parity space method. This is partly because the effective window for the smoothing method is larger than for parity space. Using a Kalman filter estimate of the state in combination with the data in the window, as done in the smoothing method, can be seen as a way of extending the window. See also the discussion in Section 4.4.3. The use of the Kalman filter estimate can also be a drawback. During the period of disturbance, the estimation of the states can only be done using the accelerometer. Since this implies only partial observability of the states, there will be a drift in the estimate of the unobservable part. The effect of this drift can be seen in Figure 7.3, where the test statistic is greater than its theoretical average even after the disturbance has disappeared. The accelerometer has no disturbance and that is also shown by the test statistics presented in Figure 7.4.

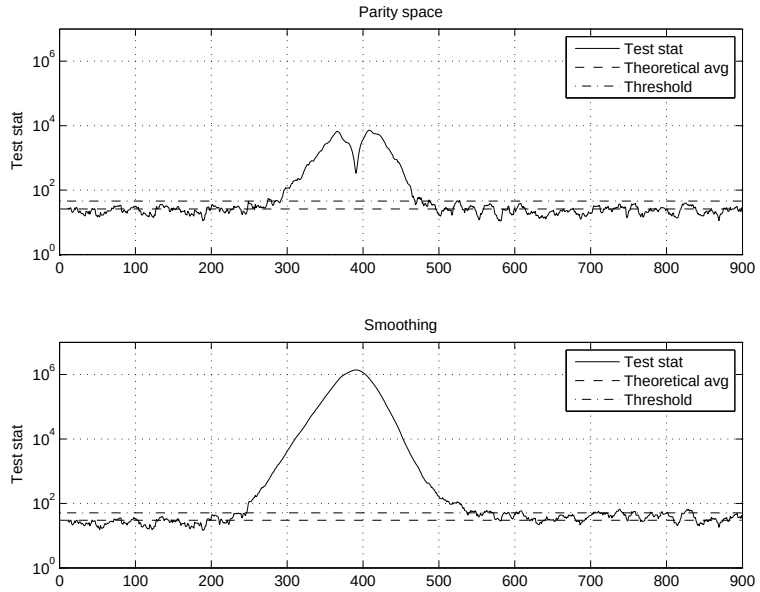


Figure 7.3: *The test statistics for the magnetometer in data set D1.*

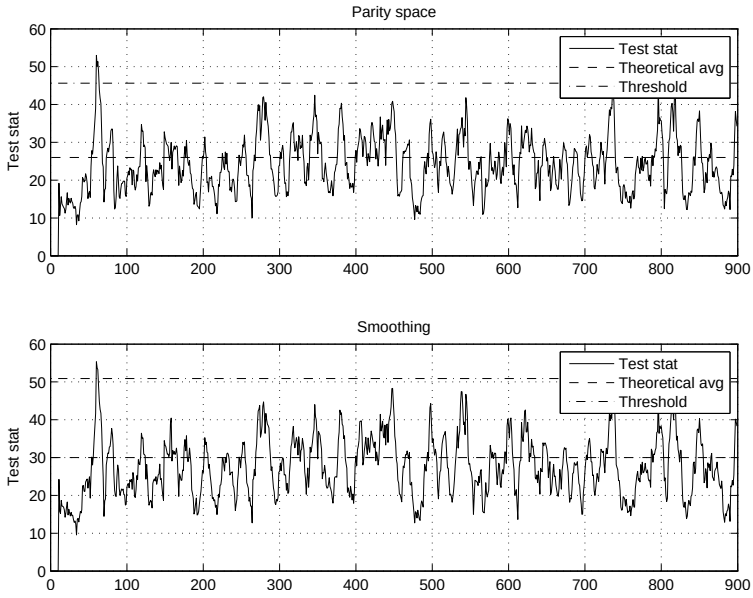


Figure 7.4: *The test statistics for the accelerometer in data set D1.*

Data set D2

The data set D2 has a weaker disturbance on the magnetometer. The test statistics for the magnetometer is shown in Figure 7.5. The parity space method has problems detecting the disturbance with the given window length. Using the smoothing method, the disturbance becomes visible. The test statistic for the accelerometer is unaffected as can be seen in Figure 7.6.

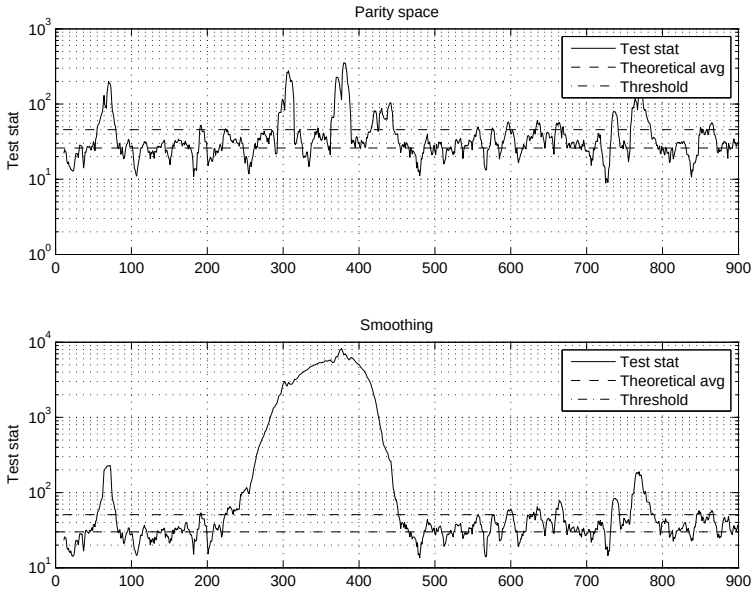


Figure 7.5: *The test statistics for the magnetometer in data set D2.*

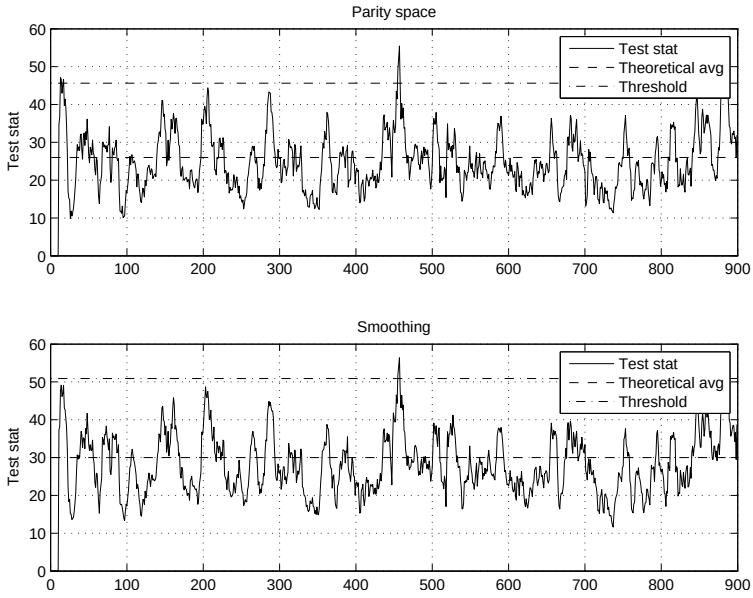


Figure 7.6: *The test statistics for the accelerometer in data set D2.*

Data set UD2

For the data set UD2, the IMU is held still by hand. The movements introduced by the person holding the IMU are modeled as white Gaussian noise and the covariance matrices estimated from the data is given in (6.4). Interesting to note is that the cross covariances are big, especially for the gyroscopes and the magnetometer. Using these covariance matrices, the test statistics for the magnetometer can be seen in Figure 7.7. The averages of the test statistics can be seen in Table 7.2 and it can be noted that they are much lower than the theoretical values. It was however seen in Section 6.2 that the noise for the magnetometer was poorly modeled with white Gaussian noise. This modeling error is likely to cause the deviations from the theoretical thresholds.

Table 7.2: *Test statistic data for data set UD2, μ is the theoretical average, $\hat{\mu}$ the obtained average and h the threshold for detection.*

Sensor	**Method**	μ	$\hat{\mu}$	h
Magnetometer	parity	26	10.24	45.6
Magnetometer	smoothing	30	14.93	50.9
Accelerometer	parity	26	18.20	45.6
Accelerometer	smoothing	30	23.72	50.9

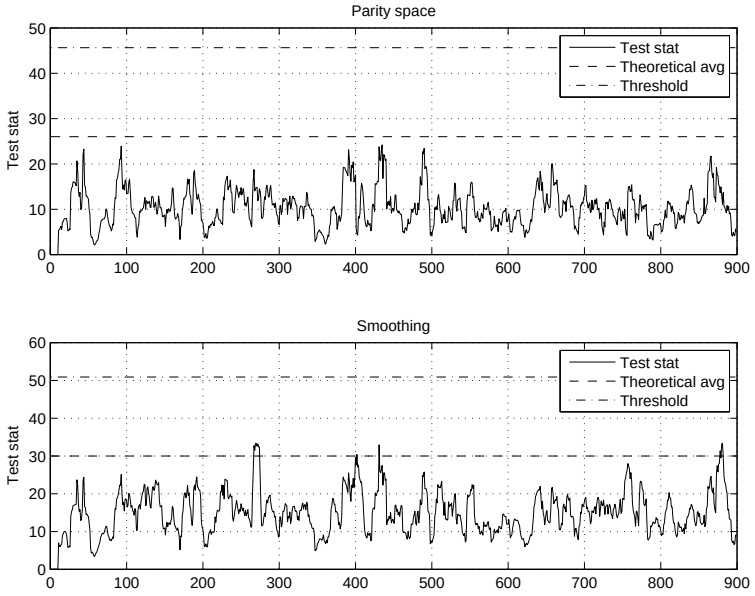


Figure 7.7: *The test statistics for the magnetometer in data set UD2. The averages are 10.8 for parity space and 21.1 for smoothing.*

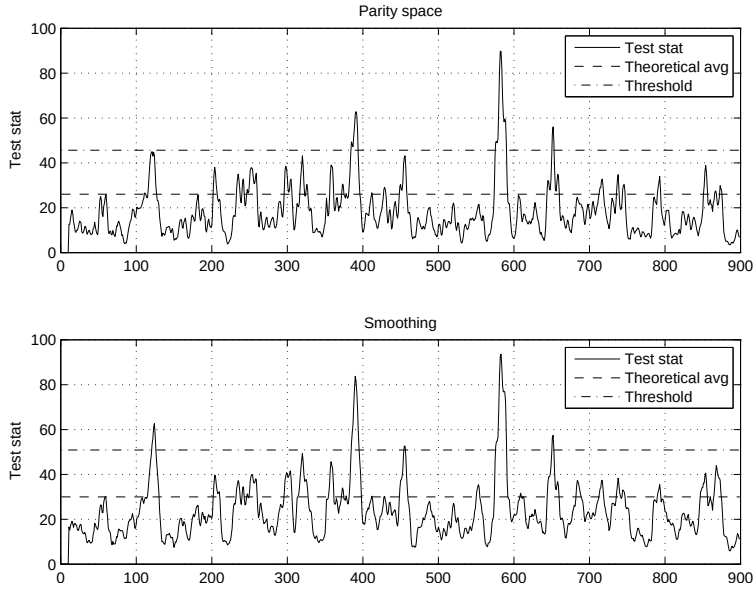


Figure 7.8: *The test statistics for the accelerometer in data set UD2. The averages are 20.5 for parity space and 28.5 for smoothing.*

Data set UD3

For data set UD3, the IMU is rotated by hand around axis e_x^b. The test statistics for the magnetometer and accelerometer can be seen in Figure 7.9 and 7.10 respectively. It can be seen that the test statistic both for parity space and smoothing for the magnetometer is above the theoretical average during the rotation period (sample 100–550). This discrepancy is partly due to linearization errors since the linearization is very sensitive to erroneous state estimates. It is also due to alignment errors of the magnetometer, see the discussion in Section 7.3. The test statistic for the accelerometer is also above the theoretical mean during the rotation. This is because it is hard not to shake the sensor during hand-held rotation. Due to these errors in the fault-free case, the thresholds are set to 100 and 500 for the magnetometer and accelerometer respectively.

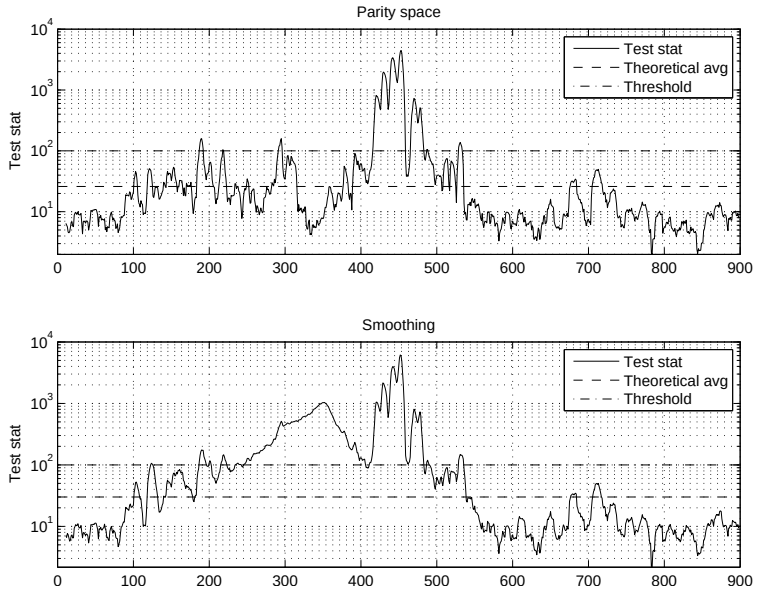


Figure 7.9: *The test statistics for the magnetometer in data set UD3.*

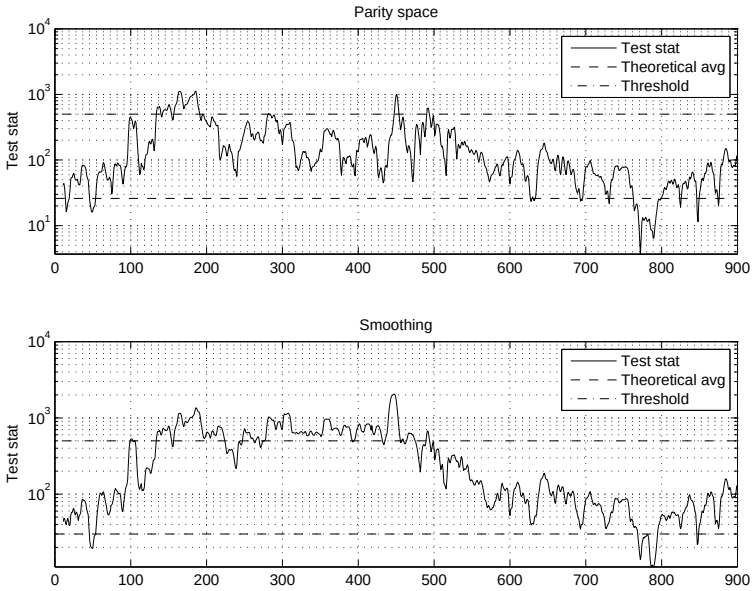


Figure 7.10: *The test statistics for the accelerometer in data set UD3.*

Data set D3

Data set D3 is collected when the IMU is rotated by hand around axis e_x^b and during the rotation, the magnetometer is disturbed. The thresholds used are the same as for data set UD3 since the conditions (except the disturbance) are similar. The test statistic for the magnetometer is shown in Figure 7.11. It can be seen that the disturbance is detected by the parity space method and that the test statistic gets a higher value than for the undisturbed case (UD3). The smoothing method has a higher value of the test statistic during the disturbance period. The test statistics for the accelerometer is shown in Figure 7.12 and it can be seen here as well that the movements of the IMU increases during the rotation period (sample 100–600).

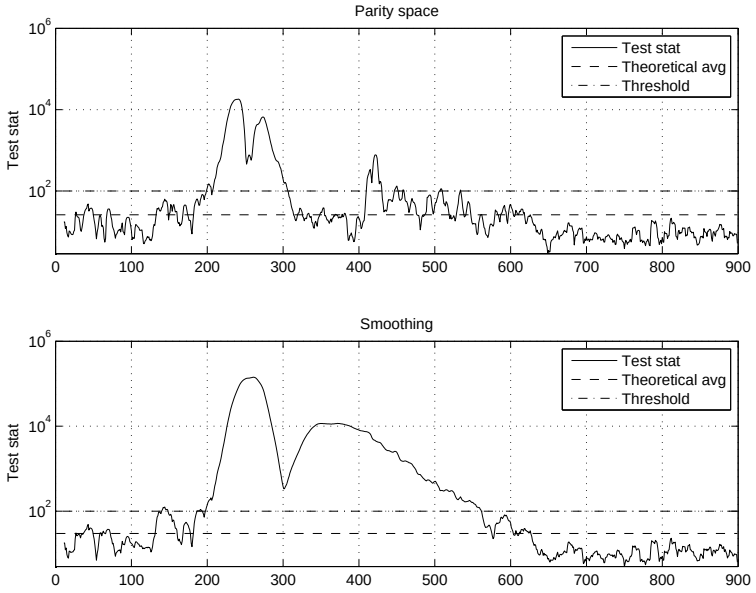


Figure 7.11: *The test statistics for the magnetometer in data set D3.*

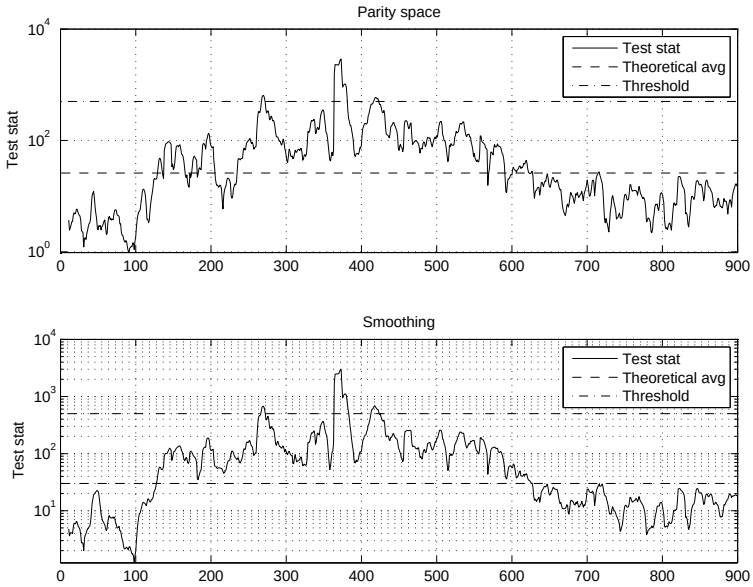


Figure 7.12: *The test statistics for the accelerometer in data set D3.*

Data set D4

The data set D4 is again hand-held measurements but contains no rotations. The IMU is here accelerated along the e_x^b-axis. The test statistic for the magnetometer can be seen in Figure 7.13 and is clearly not disturbed. The test statistics for the accelerometer is shown in Figure 7.14. It can be seen that the disturbance is detected both with parity space and smoothing.

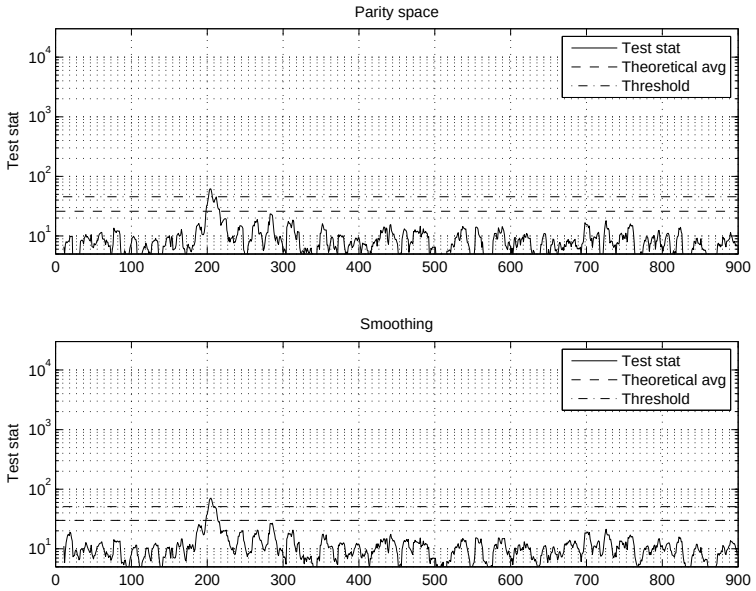


Figure 7.13: *The test statistics for the magnetometer in data set D4.*

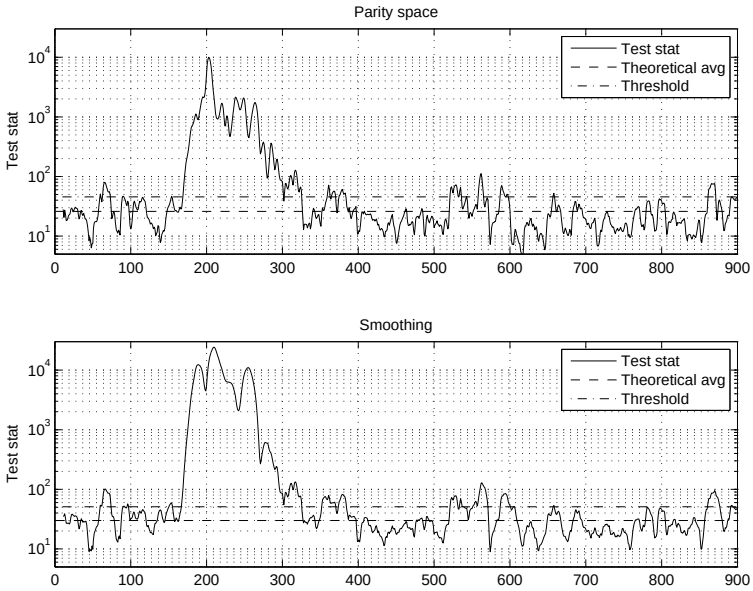


Figure 7.14: *The test statistics for the accelerometer in data set D4.*

7.2.2 Linearized Hypothesis Testing with Fault Parameterization

In this section, the fault is parameterized as described in Section 7.1. A third order model is used to describe the disturbance in each dimension. Each sensor, having three dimensions, will then have a ninth order model. The Chi-square distributed test statistic will then have 9 degrees of freedom which also should be the average, see Section 2.7.2. The window length is 10 samples and the threshold for detection is set to a confidence level of 1 %.

The first test is done on data set UD1, where the sensor is undisturbed. The test statistics for the magnetometer can be seen in Figure 7.15. The averages for the test statistics are 9.28 and 9.75 for parity space and smoothing respectively, which correspond quite good to the theoretical value. The test statistics for the accelerometer are shown in Figure 7.16. The averages are somewhat larger, 10.39 and 10.87.

Table 7.3: *Test statistic data for data set UD1, μ is the theoretical average, $\hat{\mu}$ the obtained average and h the threshold for detection.*

Sensor	**Method**	μ	$\hat{\mu}$	h
Magnetometer	parity	9	9.28	21.66
Magnetometer	smoothing	9	9.75	21.66
Accelerometer	parity	9	10.39	21.66
Accelerometer	smoothing	9	10.87	21.66

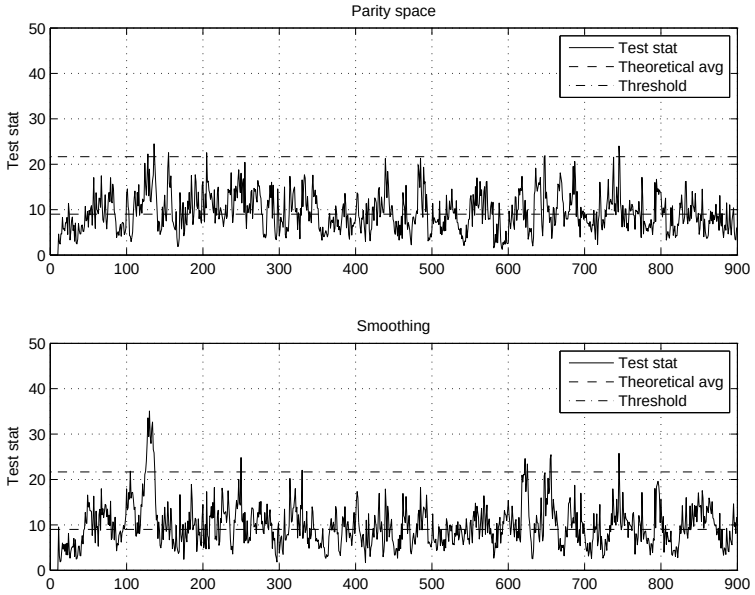


Figure 7.15: *The test statistics for the magnetometer in data set UD1 using a third order basis in each dimension. The averages are 9.28 for parity space and 9.75 for smoothing.*

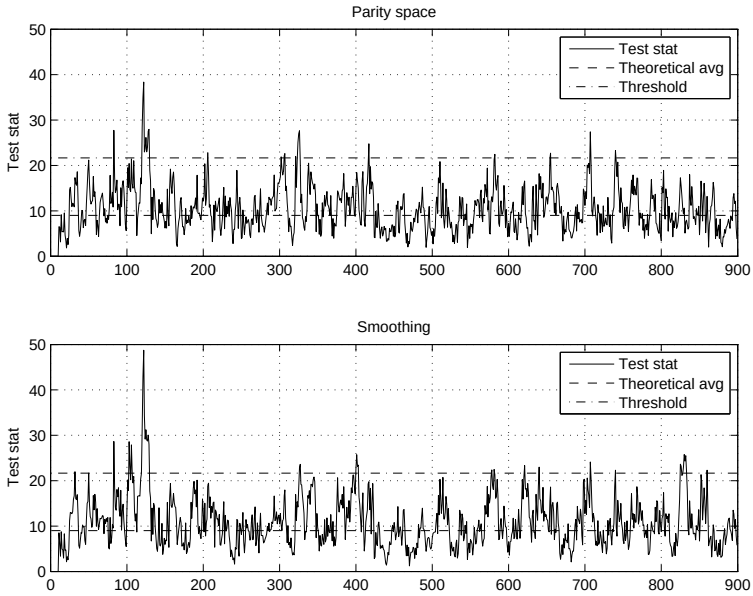


Figure 7.16: *The test statistics for the accelerometer in data set UD1 using a third order basis in each dimension. The averages are 10.39 for parity space and 10.87 for smoothing.*

Data set D1

The results for the disturbed data set D1 can be seen in Figure 7.17 and 7.18 for the magnetometer and accelerometer respectively. It is interesting to note that the value of the test statistics for the magnetometer during the disturbance are equally large as when the disturbance is not parameterized (see Figure 7.3). This means that the model of the disturbance is a good description of the real disturbance since the part of the signal not described by the model is projected away. During the undisturbed periods, the average is lowered to approximately $1/3$ compared to not using parameterization. A bigger ratio between disturbed and undisturbed periods increases the power of the test. The test statistic for the accelerometer indicates no disturbances.

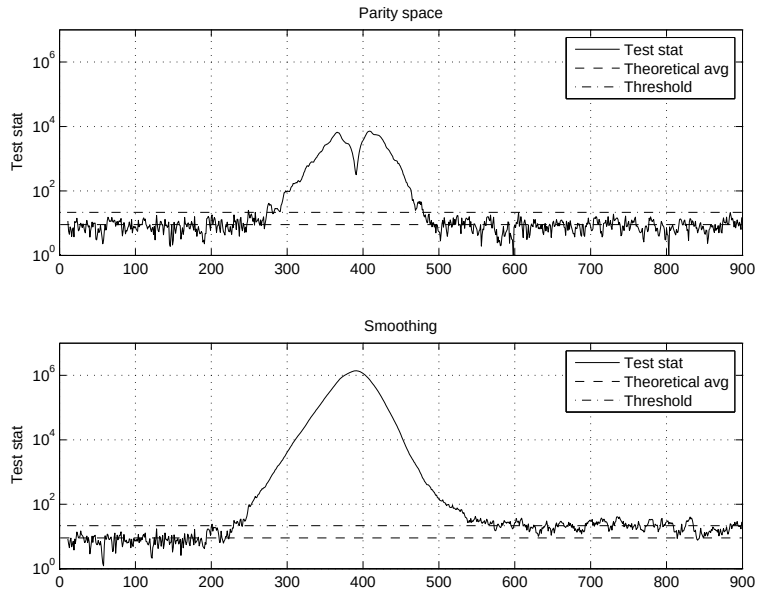


Figure 7.17: *The test statistics for the magnetometer in data set D1 using a third order basis in each dimension.*

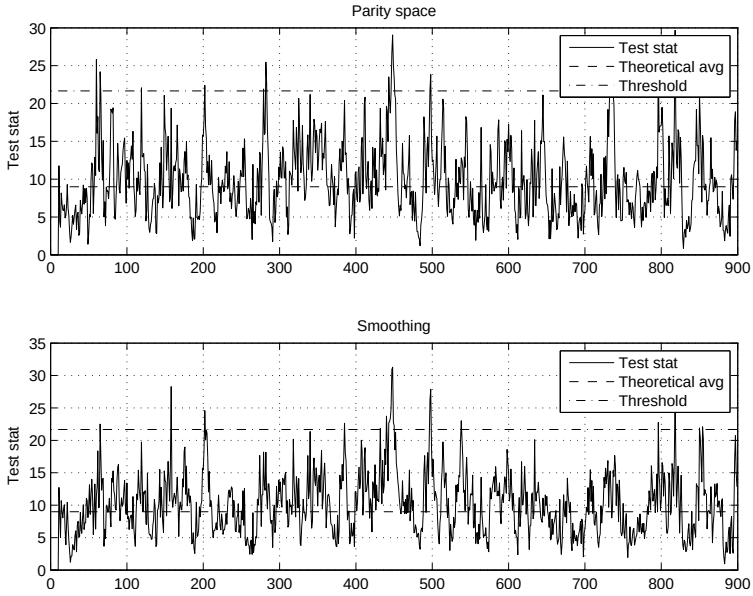


Figure 7.18: *The test statistics for the accelerometer in data set D1 using a third order basis in each dimension.*

Data set UD2

The IMU is held still by hand in data set UD2. The test statistics using a third order parameterization for disturbances are shown in Figure 7.19 and 7.20 for the magnetometer and accelerometer respectively. It can be seen that both test statistics for the magnetometer is lowered to approximately $1/2$ by using the fault parameterization, compare to Table 7.2. The average for the accelerometer is just slightly lowered which indicates that the fault model parameterizes the noise. This could probably be avoided choosing a longer time window for the detection algorithm.

Table 7.4: *Test statistic data for data set UD2, μ is the theoretical average, $\hat{\mu}$ the obtained average and h the threshold for detection.*

Sensor	**Method**	μ	$\hat{\mu}$	h
Magnetometer	parity	9	4.81	21.66
Magnetometer	smoothing	9	7.60	21.66
Accelerometer	parity	9	15.82	21.66
Accelerometer	smoothing	9	20.01	21.66

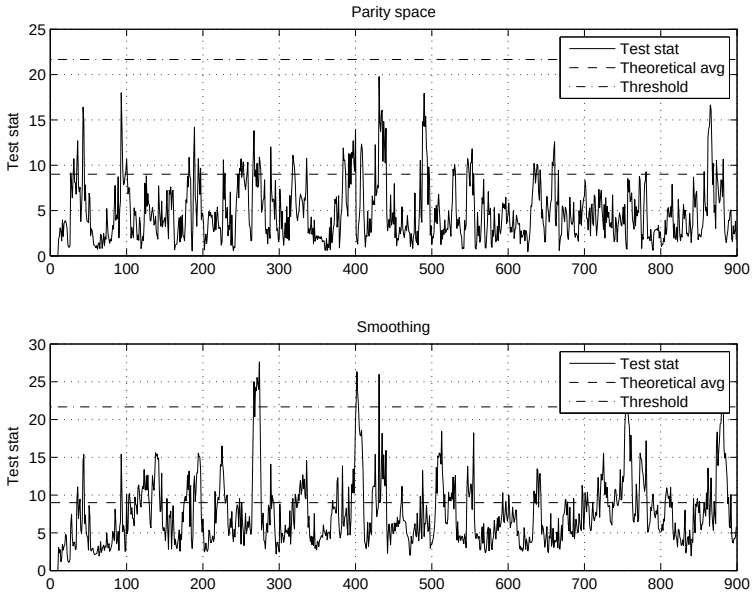


Figure 7.19: *The test statistics for the magnetometer in data set UD2 using a third order basis in each dimension.*

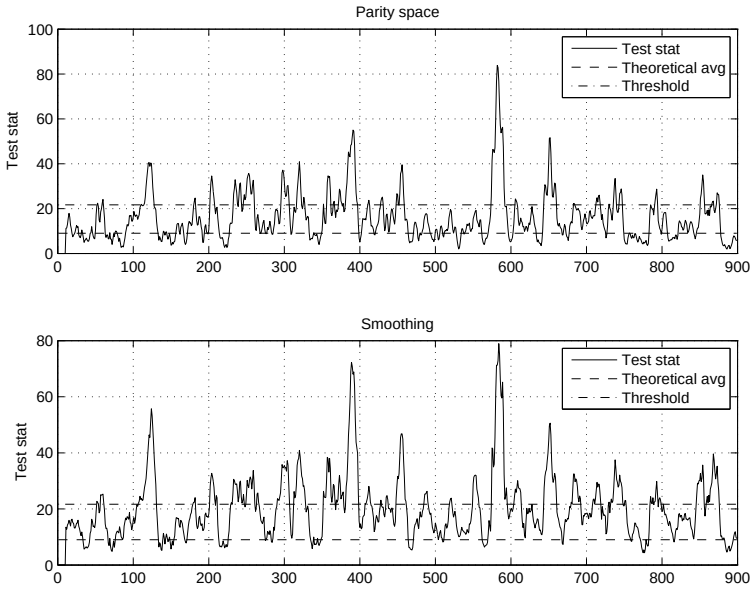


Figure 7.20: *The test statistics for the accelerometer in data set UD2 using a third order basis in each dimension.*

Data set D2

The test statistics from data set D2, which have a lighter disturbance, are presented in Figure 7.21 and 7.22. The magnetometer disturbance is hard to detect with the parity space method, but is well detected by the smoothing method. If the test statistic for the smoothing method is compared to the unparameterized test in Figure 7.5, it can be seen that the value in presence of fault is the same whereas the average during fault-free periods are lowered. The accelerometer is undisturbed which is reflected in the tests.

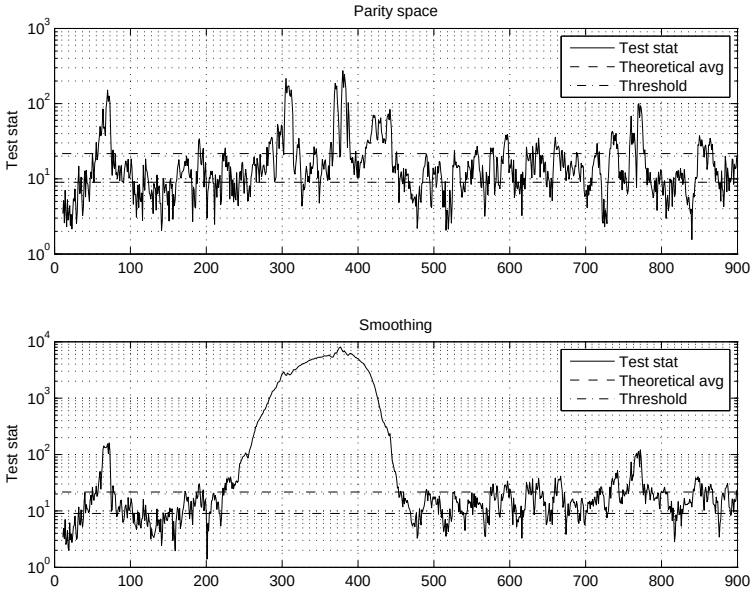


Figure 7.21: *The test statistics for the magnetometer in data set D2 using a third order basis in each dimension.*

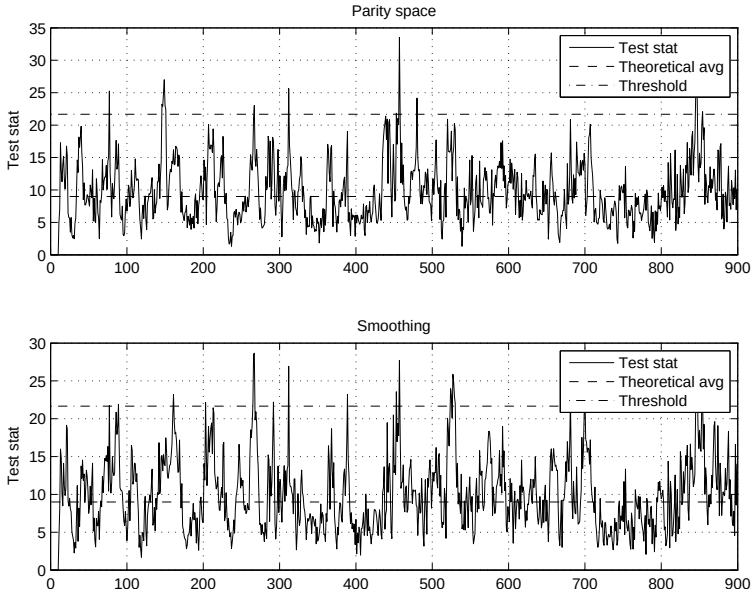


Figure 7.22: *The test statistics for the accelerometer in data set D2 using a third order basis in each dimension.*

Data set UD3

Data set UD3 is measured during a rotation of the IMU around the e_x^b-axis. The test statistics are shown in Figure 7.23 and 7.24. The comments in Section 7.2.1 for this data set are valid even here. The test reacts to changes in the magnetic field due to alignment errors and to acceleration induced by the person rotating the IMU. To avoid drift in the state estimates, the threshold are set to 100 for the magnetometer and 500 for the magnetometer. Using the theoretical thresholds would lead to detection which prevents measurement updates from the sensor. Since the system is not observable without both sensors, to long periods without measurement updates would lead to drift in the state estimates.

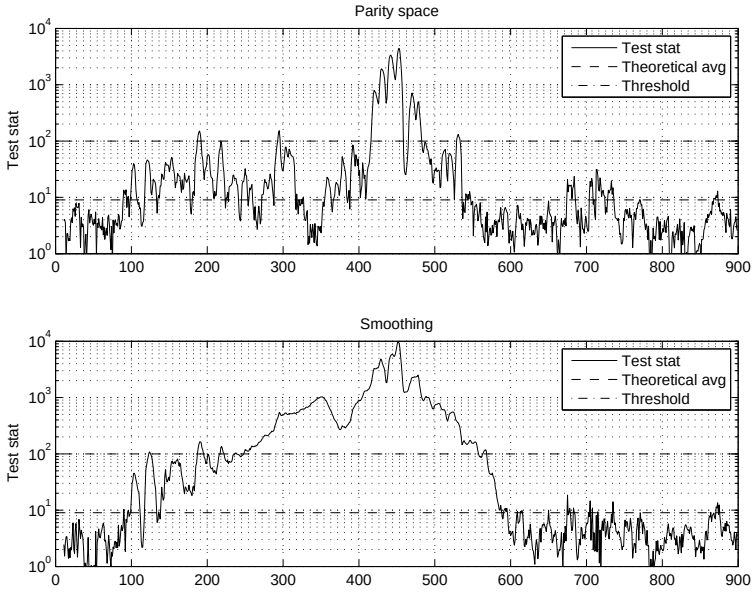


Figure 7.23: *The test statistics for the magnetometer in data set UD3 using a third order basis in each dimension.*

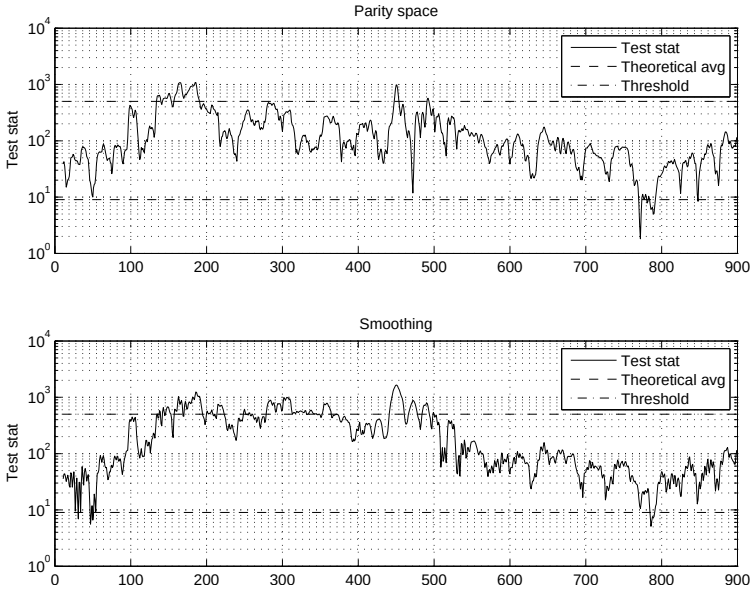


Figure 7.24: *The test statistics for the accelerometer in data set UD3 using a third order basis in each dimension.*

Data set D3

The conditions for data set D3 are similar to UD3, except that the magnetometer is disturbed in D3. The test statistics using third order basis functions are shown in Figure 7.25 and 7.26 for the magnetometer and accelerometer respectively. If the test statistics is compared to the unparameterized tests (Figure 7.11 and 7.12) it can be seen that the averages are lower during fault-free periods, but remains unchanged during the faulty period.

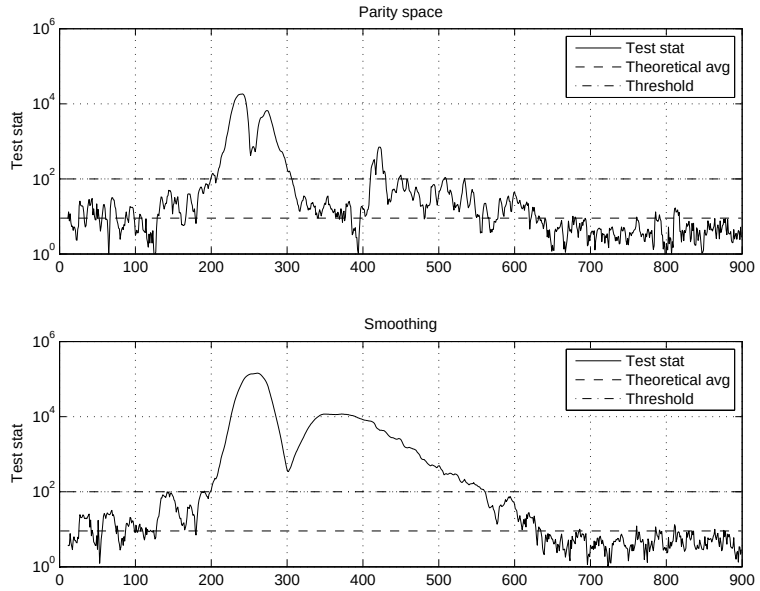


Figure 7.25: *The test statistics for the magnetometer in data set D3 using a third order basis in each dimension.*

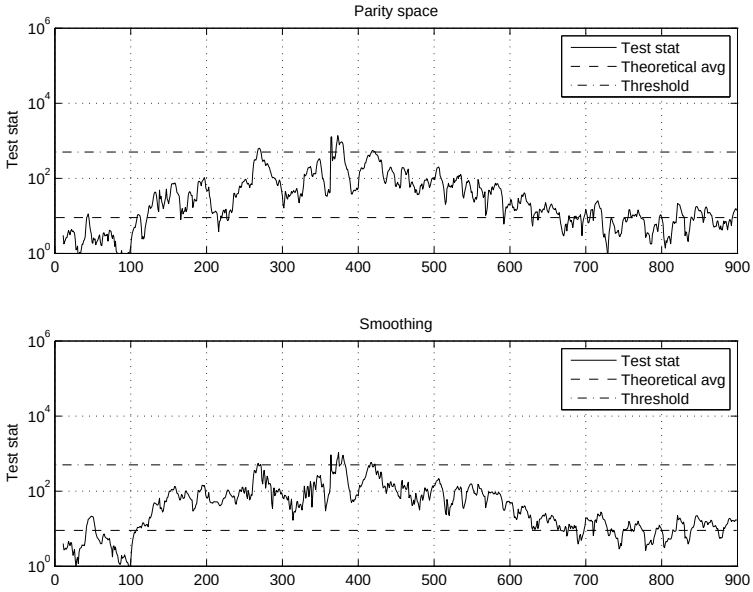


Figure 7.26: *The test statistics for the accelerometer in data set D3 using a third order basis in each dimension.*

Data set D4

In data set D4, the orientation is kept fixed but the IMU is accelerated for a short period of time. The test statistics for the magnetometer are shown in Figure 7.27. It can be seen that the magnetometer is not disturbed and that the average of the test statistics are lower than the unparameterized. The test statistics for the accelerometer are presented in Figure 7.28. The acceleration disturbances are clearly seen but the average during fault-free periods is not lowered. This indicates, as said earlier, that the noise can be parameterized by the fault model.

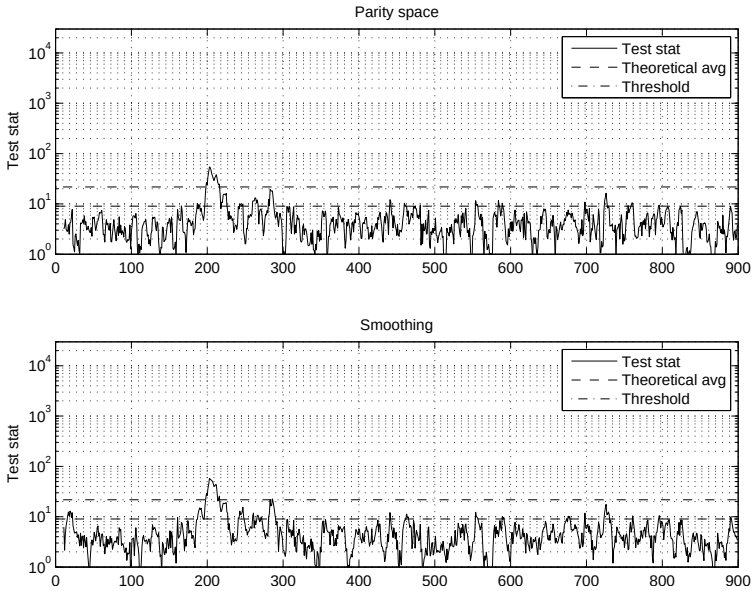


Figure 7.27: *The test statistics for the magnetometer in data set D4 using a third order basis in each dimension.*

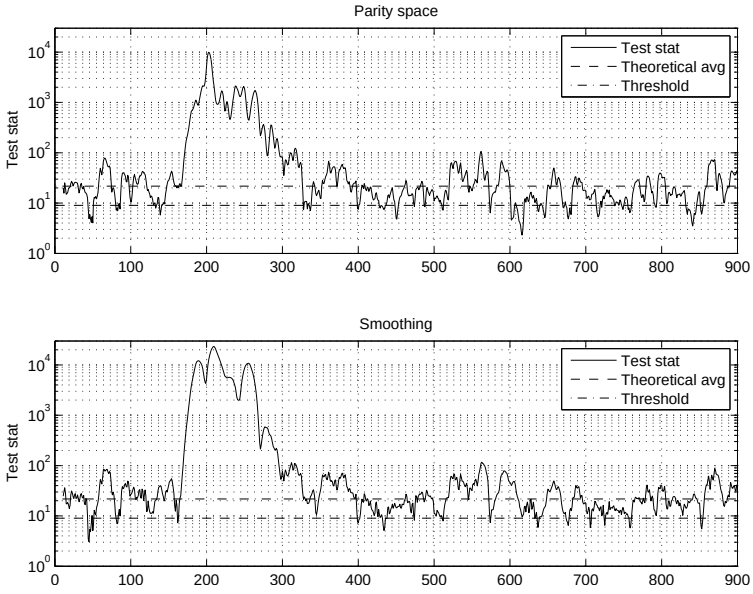


Figure 7.28: *The test statistics for the accelerometer in data set D4 using a third order basis in each dimension.*

7.3 Discussion

First consider the data sets where the IMU lies still, that is, data set UD1, D1 and D2. For those data sets, faults could be detected both with parity space and smoothing. However, the smoothing method gives a greater residual during faulty periods since the effective window (the number of samples the initial state is estimated from) is much bigger. It could also be seen that fault parameterization gives a lower average of the test statistic during fault-free periods, but the value during fault remains the same. This gives a greater power of the test.

The data sets where the IMU is held in hand are UD2, UD3, D3 and D4. New covariance matrices were estimated for the analysis of these data sets since the movements of the sensor introduced unmodeled disturbances. The covariance matrices were estimated from the raw measurements. However, the magnetometer will only react on attitude changes which also will be seen through the gyroscope signals. Since the connection between gyroscope and magnetometer is modeled the covariance will be estimated to large. Looking at the test statistics for the magnetometer in data set UD2 (Figure 7.7 and 7.19), it can be seen that the average is much lower than the theoretical threshold. This is caused by the overestimated covariance. Regarding the accelerometer, real disturbances in form of physical acceleration is introduced when the IMU is held in hand. These disturbances are more realistic to model as an increased covariance matrix. When using the fault model, the average of the test statistic becomes significantly greater than the theoretical value (Figure 7.20). This is due to that the shakings of the IMU is modeled by the parameterization, see also Figure 7.29.

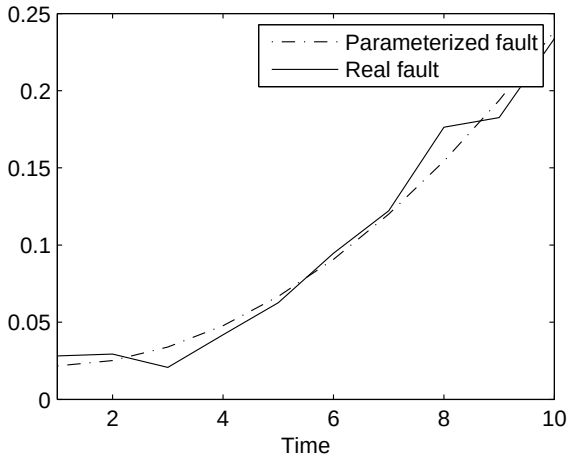


Figure 7.29: *An example of the parameterization of acceleration measurements from data set UD2.*

When rotating the sensor (data set UD3, D3) more unmodeled disturbances are introduced. Disturbances are detected on the magnetometer when the sensor is rotated. This can be due to modeling/linearization errors but probably mostly on alignment errors. The

IMU has an internal model of the sensor which essentially is

$$y_{m,t} = A^{-1}u, \tag{7.8}$$

where u is the raw sensor data and A is an alignment matrix. Temperature models are omitted here. The alignment matrix is used to compensate for the case where the components in the magnetometer are nonorthogonal. Small errors in this matrix would give different strength of the magnetic field depending on the rotation of the sensor. Data set UD3 is collected in the middle of a soccer field to avoid disturbances from magnetic objects. Even then, the norm of the magnetic field vector changes when the sensor is rotated (see Figure 7.30). A possible cause of this could be an erroneous alignment matrix. Also for the accelerometer, faults are detected when the sensor is rotated even though the sensor should not be accelerated (see Figures 7.10 and 7.24). This is explained by the fact that it is hard to rotate the sensor by hand without also moving it.

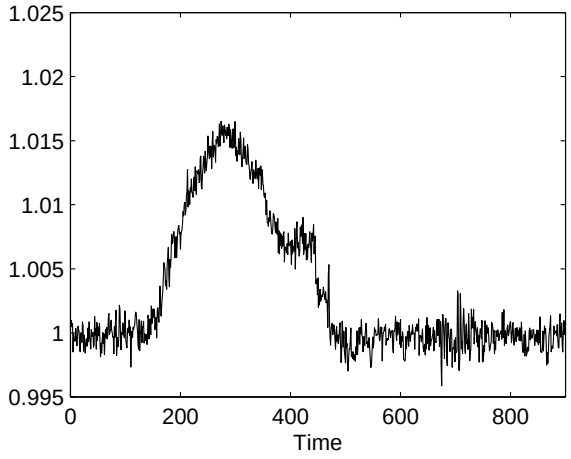


Figure 7.30: *Norm of the magnetometer measurements during roation (data set UD3), it should nominally be 1.*

Another observation made during measurements is that the magnetic field is very inhomogeneous indoors. Changing the position of the IMU also changes the magnitude of the magnetic field vector. High frequency accelerations of the sensor also tend to disturb the magnetometer. For example moving the sensor along a non-smooth surface will give high-frequency acceleration due to the friction. During these circumstances the magnetometer is clearly disturbed.

8

Concluding Remarks

ATTITUDE AND POSITIONING ESTIMATES subject to disturbances are discussed in the introduction. The thesis has discussed how these disturbances and other faults can be detected. This chapter will conclude the results discussed in the thesis and point toward future directions of research.

8.1 Conclusions

The use of statistical tests for fault detection with batched data sets have been discussed in the thesis. Estimating the initial state in the batch by minimum variance estimation from data in the batch have been shown to be identical to parity space. It was also shown that using prior knowledge about the initial state in the data batch increases the detection performance. The prior estimate are in the examples here obtained by Kalman filtering of data preceding the data batch. However, if the preceding data contain faults, the test statistic will be influenced even during fault-free periods (see Chapter 7). Hence, the reliability to the prior should be decreased after a fault is detected.

The fault parameterization introduced in Section 2.2 provides a low order description of smooth faults. This increases the power of the GLR-test (discussed in Chapter 4) which results in a higher probability of detection.

The methods discussed above are implemented for disturbance detection of an IMU. The IMU is modeled using unit quaternions to represent the orientation. Movements (acceleration disturbance) and magnetometer disturbances are detected, in order not to affect the estimated orientation. The noise distribution of the sensors are well approximated with white Gaussian noise which gives good correspondence between practical and theoretical results when the sensor is kept at rest. By rotating the IMU, alignment errors of the sensor as well as linearization errors become visible. To avoid false alarms and drift of the estimate, the thresholds then has to be higher than the theoretical values.

The observability of a fault inputs is discussed in Chapter 3. The special cases of constant and linearly increasing faults are treated and the conditions for observability are derived. Interpretations are given for the case of constant faults.

8.2 Future Work

There are many interesting questions to continue with in the area of fault detection and fault diagnosis, some of the questions raised during this work are:

- The smoothing approach to fault detection introduced in Chapter 4 may be possible to rewrite recursively. One possible way of doing this might be to use a Kalman smoother, but this introduces extra states in the filter. Would this be a possible approach and does it lower the computational complexity?
- The detection problem with applications to IMU disturbances are originally a nonlinear problem that is linearized to suit the methods presented here. One interesting extension would be to compare the results using nonlinear detection theory.
- Particle filters used for fault detection is a rather unexplored area. Can this be a good way of handling nonlinear systems, such as the IMU?
- The observability test for constant (random walk) faults and linearly growing faults are derived in Chapter 3. For constant faults, interpretations of the conditions were given. Can intuitive interpretations be derived even for linearly growing faults? It would also be interesting to apply this theory to a real-world example.

Appendices

A

Prerequisites in Vector Kinematics and Mathematics

A.1 Cross-product

Note the following properties of the cross-product

$$\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u} \tag{A.1}$$
$$a(\mathbf{u} \times \mathbf{v}) = (a\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (a\mathbf{v}) \tag{A.2}$$
$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w}) \tag{A.3}$$
$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) \tag{A.4}$$
$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = \mathbf{v}(\mathbf{w} \cdot \mathbf{u}) - \mathbf{w}(\mathbf{u} \cdot \mathbf{v}). \tag{A.5}$$

Note also that it is possible to write the cross-product as a matrix multiplication, i.e., for the 3-dimensional case

$$\mathbf{u} \times \mathbf{v} = -\begin{pmatrix} 0 & u_3 & -u_2 \\ -u_3 & 0 & u_1 \\ u_2 & -u_1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}. \tag{A.6}$$

A.2 Vector Rotation

The vector $\mathbf{u}$ is rotated around $\mathbf{n}$ as shown in Figure A.1. The new vector $\mathbf{v}$ can be described as

$$\begin{aligned} \mathbf{v} &= ON + NW + WV = \\ &= (\mathbf{u} \cdot \mathbf{n})\mathbf{n} + \frac{\mathbf{u} - (\mathbf{u} \cdot \mathbf{n})\mathbf{n}}{|\mathbf{u} - (\mathbf{u} \cdot \mathbf{n})\mathbf{n}|} \underbrace{|NV| \cos \mu}_{|NW|} + \underbrace{\frac{\mathbf{u} \times \mathbf{n}}{|\mathbf{u}| \sin \phi}}_{\frac{WV}{|WV|}} \underbrace{|NV| \sin \mu}_{|WV|}, \end{aligned} \tag{A.7}$$

where ON denotes the vector that points from O to N and so on. Note that

$$|NV| = |NU| = |\mathbf{u} - (\mathbf{u} \cdot \mathbf{n})\mathbf{n}| = |\mathbf{u}| \sin \phi.$$

Then,

$$\mathbf{v} = (1 - \cos \mu)\mathbf{n}(\mathbf{n} \cdot \mathbf{u}) + \cos \mu \mathbf{u} - \sin \mu (\mathbf{n} \times \mathbf{u}), \tag{A.8}$$

which is often referred to as the *rotation formula.*

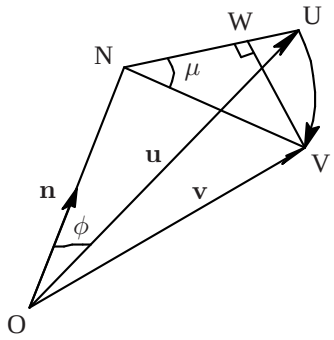


Figure A.1: *The vector* $\mathbf{u}$ *is rotated around* $\mathbf{n}$.

A.3 Direction Cosines

To describe the direction of a vector in a unique way, the direction cosines is a useful tool. The angles α, β and γ in Figure A.2 are used. The cosine of the angles will be the projections onto the basis axes. It is possible to describe the direction of $\mathbf{n}$ using only two angles, this is the case when using spherical coordinates as in right part of Figure A.2. The problem with this description is that a vector pointing in the direction of f_z can have any angle ρ, which is clearly not a unique description.

Observe the following property of the direction cosines. The theorem of Pythagoras can be written

$$|\mathbf{n}|^2 \cos^2 \alpha + |\mathbf{n}|^2 \cos^2 \beta + |\mathbf{n}|^2 \cos^2 \gamma = |\mathbf{n}|^2.$$

This yields that

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1. \tag{A.9}$$

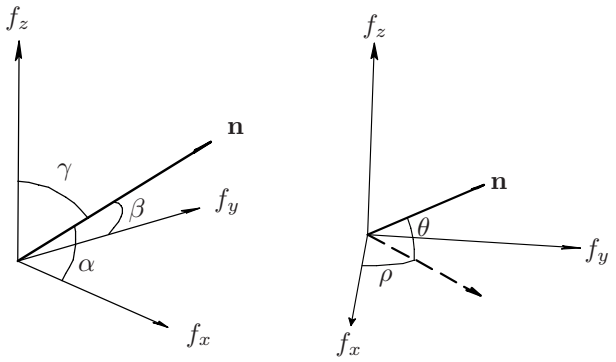


Figure A.2: *Describing the vector* $\mathbf{n}$ *can be done using the direction angles* α, β *and* γ *(left) or using only two angles (right).*

B

Quaternion preliminaries

THE QUATERNIONS WERE INVENTED by Sir William Rowan Hamilton as a way to extend the imaginary numbers, see Hamilton (1844). The unit quaternion have later shown to be a useful tool for attitude representations. A comprehensive description of unit quaternions is given by Kuipers (1999) and a nice survey of attitude representations is given by Shuster (1993). A good primer on unit quaternions for attitude representations of airplanes is given by Stevens and Lewis (2003).

B.1 Operations and Properties

The quaternion is a four-tuple with the elements $q_0, \ldots, q_3$. It can also be viewed as a vector consisting of the scalar q_0 and the vector $\mathbf{q}$.

$$q = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix}. \tag{B.1}$$

Multiplication of two quaternions denoted with $\odot$ is defined as

$$p \odot q = \begin{pmatrix} p_0 \\ \mathbf{p} \end{pmatrix} \odot \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} p_0 q_0 - \mathbf{p} \cdot \mathbf{q} \\ p_0 \mathbf{q} + q_0 \mathbf{p} + \mathbf{p} \times \mathbf{q} \end{pmatrix}. \tag{B.2}$$

The quaternions have the following properties:

$$p \odot q \neq q \odot p \tag{B.3}$$

$$\text{norm}(q) = \sum_{i=0}^{3} q_i^2 \tag{B.4}$$

$$\text{norm}(p \odot q) = \text{norm}(p) \cdot \text{norm}(q) \tag{B.5}$$

$$(p \odot q) \odot r = p \odot (q \odot r) \tag{B.6}$$

$$q^{-1} = \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix}^{-1} = \begin{pmatrix} q_0 \\ -\mathbf{q} \end{pmatrix} \tag{B.7}$$

$$(q_a \odot q_b)^{-1} = q_b^{-1} \odot q_a^{-1}. \tag{B.8}$$

B.2 Describing a Rotation with Quaternions

Consider the vector

$$q = \begin{pmatrix} \cos\delta \\ \sin\delta\mathbf{n} \end{pmatrix}, \tag{B.9}$$

which describes a rotation around the unit vector $\mathbf{n}$. This unit vector can be described as in Section A.3 with the directional cosines

$$\mathbf{n} = \begin{pmatrix} \cos\alpha \\ \cos\beta \\ \cos\gamma \end{pmatrix}. \tag{B.10}$$

Note that

$$q^T q = \cos^2\delta + \sin^2\delta(\cos^2\alpha + \cos^2\beta + \cos^2\gamma) = 1, \tag{B.11}$$

which is also a necessity for the quaternion in order to represent a rotation. The variable δ is a measure of the rotation angle as we will see later.

There are two ways of visually describing a rotation. Either a vector is rotated around another vector or the coordinate frame is rotated around a vector, see Figure B.1. When a vector is rotated the transformation gives the new vector coordinates in the frame which remains fixed. When the frame is rotated the vector coordinates are transformed to the new frame. Mathematically, the difference is only the sign of the rotation angle.

To rotate the vector $\mathbf{u}$ around $\mathbf{n}$, (A.8) and Figure A.1 can be used. Observe that the vector rotation is done counterclockwise, which could equally be seen as rotating the coordinate frame clockwise. This rotation can be done using quaternion algebra. If q represents a rotation around $\mathbf{n}$, the rotated vector can be described as

$$v = q^{-1} \odot u \odot q, \tag{B.12}$$

where

$$v = \begin{pmatrix} 0 \\ \mathbf{v} \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} 0 \\ \mathbf{u} \end{pmatrix}. \tag{B.13}$$

An alternative description is to use

$$v = q \odot u \odot q^{-1}. \tag{B.14}$$

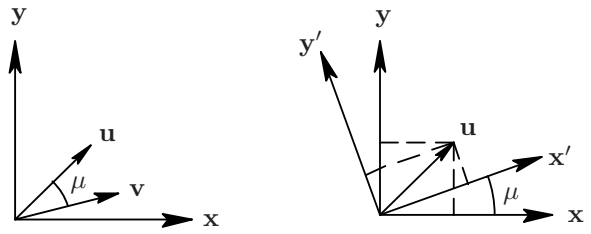


Figure B.1: *In the left figure, a vector is rotated counterclockwise with angle μ. In the right figure, the coordinate frame is rotated clockwise with angle μ.*

With this representation vector rotation becomes clockwise and coordinate frame rotation counterclockwise. However, in this work, (B.12) is viewed as the standard rotation.

The result of the quaternion multiplication in (B.12) is

$$v = \begin{pmatrix} \mathbf{q} \cdot \mathbf{u} q_0 - (q_0 \mathbf{u} - \mathbf{q} \times \mathbf{u}) \cdot \mathbf{q} \\ (\mathbf{q} \cdot \mathbf{u})\mathbf{q} + q_0(q_0 \mathbf{u} - \mathbf{q} \times \mathbf{u}) + (q_0 \mathbf{u} - \mathbf{q} \times \mathbf{u}) \times \mathbf{q} \end{pmatrix}, \tag{B.15}$$

which simplifies to (note the use of (A.5) for the last term in the second row)

$$v = \begin{pmatrix} 0 \\ 2(\mathbf{q} \cdot \mathbf{u})\mathbf{q} + (q_0^2 - \mathbf{q} \cdot \mathbf{q})\mathbf{u} - 2q_0 \mathbf{q} \times \mathbf{u} \end{pmatrix}. \tag{B.16}$$

If (B.9) is used, this can be rewritten as

$$v = \begin{pmatrix} 0 \\ 2\sin^2\delta \mathbf{n}(\mathbf{n} \cdot \mathbf{u}) + (\cos^2\delta - \sin^2\delta)\mathbf{u} - 2\cos\delta\sin\delta(\mathbf{n} \times \mathbf{u}) \end{pmatrix}. \tag{B.17}$$

This should be compared with (A.8), since the second row should give the same result. If $\delta = \mu/2$ and some trigonometric identities are applied the results are equal. Thus, inserting $\delta = \mu/2$ in (B.9), the unit quaternion

$$q = \begin{pmatrix} \cos(\mu/2) \\ \sin(\mu/2)\mathbf{n} \end{pmatrix} \tag{B.18}$$

describes a counterclockwise vector rotation or a clockwise coordinate frame rotation around $\mathbf{n}$ with angle μ.

Observe that rotating first with q_a and then with q_b can be done in one step by rotating by $q_a \odot q_b$. Using (B.8) this is easily seen since

$$q_b^{-1} \odot (q_a^{-1} \odot u \odot q_a) \odot q_b = (q_a \odot q_b)^{-1} \odot u \odot (q_a \odot q_b). \tag{B.19}$$

B.3 Rotation Matrix

It is possible to rewrite the quaternion multiplication in (B.12) as a matrix multiplication. Rewriting (B.16) on the form of

$$\begin{pmatrix} 0 \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} 0 \\ R(q)\mathbf{u} \end{pmatrix} \tag{B.20}$$

using (A.6) yields

$$R(q) = \begin{pmatrix} (q_0^2 + q_1^2 - q_2^2 - q_3^2) & 2(q_1 q_2 + q_0 q_3) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & (q_0^2 - q_1^2 + q_2^2 - q_3^2) & 2(q_2 q_3 + q_0 q_1) \\ 2(q_1 q_3 + q_0 q_2) & 2(q_2 q_3 - q_0 q_1) & (q_0^2 - q_1^2 - q_2^2 + q_3^2) \end{pmatrix}. \tag{B.21}$$

This is also described in (Kuipers, 1999, p. 158).

B.4 Dynamics

Consider a moving body system. The description of the rotation between the F-system (fixed system) and the B-system (body system) will then contain some dynamics. In order to derive the dynamic equations, the time derivative of the elements of the unit quaternion will be calculated.

Let $q_{bf}(t)$ represent the rotation of the B-system with respect to the F-system at time t. Furthermore, let the instantaneous angular velocity of the B-system be in the direction of the unit vector $\hat{\mathbf{s}}^b$, with magnitude ω, which is represented in the B-system. Then, the unit quaternion δq_{bf} describes the rotation around $\hat{\mathbf{s}}^b$ during the small time interval δt. Using small angles approximation and (B.18) this can be written as

$$\delta q_{bf}(\delta t) \approx \begin{pmatrix} 1 \\ \hat{\mathbf{s}}^b \omega \delta t/2 \end{pmatrix}. \tag{B.22}$$

Then, the rotated unit quaternion can be represented as

$$q_{bf}(t + \delta t) = q_{bf}(t) \odot \delta q_{bf}(\delta t). \tag{B.23}$$

The derivative of $q_{bf}(t)$ can then be written as

$$\begin{aligned} \frac{dq_{bf}(t)}{dt} &= \lim_{\delta t \to 0} \frac{q_{bf}(t + \delta t) - q_{bf}(t)}{\delta t} = \lim_{\delta t \to 0} \frac{q_{bf}(t) \odot (\delta q_{bf}(\delta t) - I_q)}{\delta t} = \\ &= \lim_{\delta t \to 0} q_{bf}(t) \odot \left(\frac{\delta q_{bf}(\delta t)}{\delta t} - \frac{I_q}{\delta t} \right) = \frac{1}{2} q_{bf}(t) \odot \begin{pmatrix} 0 \\ \hat{\mathbf{s}}^b \omega \end{pmatrix} = \frac{1}{2} q_{bf}(t) \odot \omega_{bf}^b, \end{aligned} \tag{B.24}$$

where I_q represents the unit quaternion and ω_{bf}^b represents the angular velocity of the B-system relative to the F-system expressed in the B-system. Decompose the unit quaternions as

$$q_{bf} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix} \quad \text{and} \quad \omega_{bf}^b = \begin{pmatrix} 0 \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\omega} \end{pmatrix}. \tag{B.25}$$

Then, the quaternion multiplication in (B.24) can be written as a matrix multiplication

with help of (B.2), (A.6) and (A.1).

$$\begin{aligned}
\dot{q}_{bf}(t) = \frac{1}{2} q_{bf}(t) \odot \omega_{bf}^{b} = \frac{1}{2} \begin{pmatrix} -\mathbf{q} \cdot \boldsymbol{\omega} \\ q_0 \boldsymbol{\omega} + \mathbf{q} \times \boldsymbol{\omega} \end{pmatrix} = \\
= \frac{1}{2} \begin{pmatrix} -(q_1\omega_x + q_2\omega_y + q_3\omega_z) \\ q_0 \boldsymbol{\omega} - \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \end{pmatrix} = \\
= \underbrace{\frac{1}{2} \begin{pmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{pmatrix}}_{S(\omega_{bf}^{b})} \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \\
= \underbrace{\frac{1}{2} \begin{pmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix}}_{\triangleq S'(q_{bf})} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}
\end{aligned} \tag{B.26}$$

Bibliography

M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1965.

Analog Devices. Online: http://www.analog.com.

B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Inc, Englewood Cliffs, NJ, 1979. ISBN 0-13-638122-7.

M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Information and system sciences series. Prentice Hall, Englewood Cliffs, NJ, 1993.

C. Bembeneck, T. A. Chmielewski Jr., and P. R. Kalta. Observability conditions for biased linear time invariant systems. In *Proceedings of the American Control Conference*, pages 1180–1184, Philadelphia, Pennsylvania, June 1998.

N. Bergman. *Recursive Bayesian Estimation: Navigation and Tracking Applications*. Dissertations no 579, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden, May 1999.

Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM Publications, 1996.

T. A. Chmielewski Jr. and P. R. Kalata. On the identification of stochastic biases in linear time invariant systems. In *Proceedings of the American Control Conference*, pages 4067–4071, Seattle, Washington, June 1995.

A. Y. Chow and A. S. Willsky. Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control*, 29(7):603–614, 1984.

X. Ding, L. Guo, and T. Jeinsch. A characterization of parity space and its application to robust fault detection. *IEEE Transactions on Automatic Control*, 44(2):337–343, 1999.

J. W. Gardner. *Microsensors: Principles and Applications.* John Wiley & Sons, Ltd, 1994.

J. Gertler. Fault detection and isolation using parity relations. *Control Engineering Practice*, 5(5):653–661, 1997.

J. J. Gertler. *Fault Detection and Diagnosis in Engineering Systems.* Marcel Dekker, Inc., 1998.

G. H. Golub and C. F. van Loan. *Matrix Computations.* John Hopkins University Press, 3 edition, 1996. ISBN 0-2018-54-14-8.

F. Gunnarsson, D. Törnqvist, E. Geijer Lundin, G. Bark, N. Wiberg, and E. Englund. Uplink transmission timing in WCDMA. In *Proceedings of IEEE Vehicular Technology Conference Fall*, Orlando, FL, USA, Oct. 2003.

F. Gustafsson. *Adaptive filtering and change detection.* John Wiley & Sons, Inc., 2 edition, 2001.

S. W. R. Hamilton. On quaternions; or on a new system of imaginaries in algebra. *Philosophical Magazine*, xxv:10–13, July 1844.

J. D. Hol. Sensor fusion for camera pose estimation. Master's thesis, University of Twente, June 2005.

Honeywell Solid State Electronics Center. Online: http://www.ssec.honeywell.com.

T. Kailath. *Linear Systems.* Prentice-Hall, Inc, 1980. ISBN 0-13-536961-4.

T. Kailath, , A. H. Sayed, and B. Hassibi. *Linear Estimation.* Prentice Hall, 2000.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Tranactions of the American Society of Mechanical Engineering — Journal Basic Engieering*, 82 (Series D):35–45, Mar. 1960.

R. Karlsson. *Particle Filtering for Positioning and Tracking Applications.* Dissertations no 924, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden, Mar. 2005.

S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*, volume 2. Prentice-Hall, Inc, 1998. ISBN 0-13-504135-X.

J. B. Kuipers. *Quaternions and Rotation Sequences.* Princeton Univerity Press, 1999.

E. L. Lehmann. *Testing Statistical Hypotheses.* Probability and Mathematical Statistics. John Wiley & Sons, Ltd, 2 edition, 1986. ISBN 0-471-84083-1.

L. Ljung. *System Identification Toolbox: User's Guide.* The Mathworks, Inc., 2006.

H. J. Luinge. *Inertial Sensing of Human Movement.* PhD thesis, University of Twente, 2002.

MATRIS. Online: http://www.ist-matris.org.

C. Meyer. *Matrix Analysis and Applied Linear Algebra.* SIAM, 2000.

National Geomagnetism Program. Online: http://geomag.usgs.gov.

J. Neyman and E. S. Pearson. On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika*, 20A(1/2):175–240, July 1928a.

J. Neyman and E. S. Pearson. On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika*, 20A(3/4):263–294, Dec. 1928b.

T. J. Rivlin. *The Chebyshev Polynomials.* John Wiley & Sons, Inc, 1974.

D. Roetenberg. *Inertial and Magnetic Sensing of Human Motion.* PhD thesis, University of Twente, 2006.

D. Roetenberg, H. J. Luinge, C. T. M. Baten, and P. H. Veltink. Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(3): 395–405, 2005.

W. J. Rugh. *Linear System Theory*. Prentice-Hall, Inc, 1996. ISBN 0-13-441205-2.

T. Schön. Personal communication, 2005.

T. B. Schön. *Estimation of Nonlinear Dynamic Systems: Theory and Applications.* Dissertations no 998, Department of Electrical Engineering, Linköpings universitet, Sweden, SE-581 83 Linköping, Sweden, Feb. 2006.

M. D. Shuster. A survey of attitude representations. *The Journal of Astronautical Sciences*, 41(4):439–517, 1993.

B. L. Stevens and F. L. Lewis. *Aircraft control and simulation.* John Wiley & Sons, Inc., 2 edition, 2003.

D. H. Titterton and J. L. Weston. *Strapdown Inertial Navigation Technology*. IEEE, 1997.

D. Törnqvist and F. Gustafsson. Eliminating the initial state for the generalized likelihood ratio test. In *Proceedings of IFAC Symposium SAFEPROCESS*, Beijing, China, Aug. 2006. To appear.

D. Törnqvist, F. Gustafsson, and I. Klein. GLR tests for fault detection over sliding data windows. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 2005.

D. Törnqvist, E. Geijer Lundin, F. Gunnarsson, and F. Gustafsson. Transmission timing - a control approach to distributed uplink scheduling in WCDMA. In *Proceedings of American Control Conference*, Boston, MA, USA, June 2004.

A. Willsky and H. Jones. A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, 21: 108–112, 1976.

Xsens Technologies B.V. Online: http://www.xsens.com.

Licentiate Theses
Division of Automatic Control
Linköpings universitet

P. Andersson: Adaptive Forgetting through Multiple Models and Adaptive Control of Car Dynamics. Thesis No. 15, 1983.
B. Wahlberg: On Model Simplification in System Identification. Thesis No. 47, 1985.
A. Isaksson: Identification of Time Varying Systems and Applications of System Identification to Signal Processing. Thesis No. 75, 1986.
G. Malmberg: A Study of Adaptive Control Missiles. Thesis No. 76, 1986.
S. Gunnarsson: On the Mean Square Error of Transfer Function Estimates with Applications to Control. Thesis No. 90, 1986.
M. Viberg: On the Adaptive Array Problem. Thesis No. 117, 1987.
K. Ståhl: On the Frequency Domain Analysis of Nonlinear Systems. Thesis No. 137, 1988.
A. Skeppstedt: Construction of Composite Models from Large Data-Sets. Thesis No. 149, 1988.
P. A. J. Nagy: MaMiS: A Programming Environment for Numeric/Symbolic Data Processing. Thesis No. 153, 1988.
K. Forsman: Applications of Constructive Algebra to Control Problems. Thesis No. 231, 1990.
I. Klein: Planning for a Class of Sequential Control Problems. Thesis No. 234, 1990.
F. Gustafsson: Optimal Segmentation of Linear Regression Parameters. Thesis No. 246, 1990.
H. Hjalmarsson: On Estimation of Model Quality in System Identification. Thesis No. 251, 1990.
S. Andersson: Sensor Array Processing; Application to Mobile Communication Systems and Dimension Reduction. Thesis No. 255, 1990.
K. Wang Chen: Observability and Invertibility of Nonlinear Systems: A Differential Algebraic Approach. Thesis No. 282, 1991.
J. Sjöberg: Regularization Issues in Neural Network Models of Dynamical Systems. Thesis No. 366, 1993.
P. Pucar: Segmentation of Laser Range Radar Images Using Hidden Markov Field Models. Thesis No. 403, 1993.
H. Fortell: Volterra and Algebraic Approaches to the Zero Dynamics. Thesis No. 438, 1994.
T. McKelvey: On State-Space Models in System Identification. Thesis No. 447, 1994.
T. Andersson: Concepts and Algorithms for Non-Linear System Identifiability. Thesis No. 448, 1994.
P. Lindskog: Algorithms and Tools for System Identification Using Prior Knowledge. Thesis No. 456, 1994.
J. Plantin: Algebraic Methods for Verification and Control of Discrete Event Dynamic Systems. Thesis No. 501, 1995.
J. Gunnarsson: On Modeling of Discrete Event Dynamic Systems, Using Symbolic Algebraic Methods. Thesis No. 502, 1995.
A. Ericsson: Fast Power Control to Counteract Rayleigh Fading in Cellular Radio Systems. Thesis No. 527, 1995.
M. Jirstrand: Algebraic Methods for Modeling and Design in Control. Thesis No. 540, 1996.
K. Edström: Simulation of Mode Switching Systems Using Switched Bond Graphs. Thesis No. 586, 1996.
J. Palmqvist: On Integrity Monitoring of Integrated Navigation Systems. Thesis No. 600, 1997.
A. Stenman: Just-in-Time Models with Applications to Dynamical Systems. Thesis No. 601, 1997.
M. Andersson: Experimental Design and Updating of Finite Element Models. Thesis No. 611, 1997.
U. Forssell: Properties and Usage of Closed-Loop Identification Methods. Thesis No. 641, 1997.

M. Larsson: On Modeling and Diagnosis of Discrete Event Dynamic systems. Thesis No. 648, 1997.

N. Bergman: Bayesian Inference in Terrain Navigation. Thesis No. 649, 1997.

V. Einarsson: On Verification of Switched Systems Using Abstractions. Thesis No. 705, 1998.

J. Blom, F. Gunnarsson: Power Control in Cellular Radio Systems. Thesis No. 706, 1998.

P. Spångéus: Hybrid Control using LP and LMI methods – Some Applications. Thesis No. 724, 1998.

M. Norrlöf: On Analysis and Implementation of Iterative Learning Control. Thesis No. 727, 1998.

A. Hagenblad: Aspects of the Identification of Wiener Models. Thesis No. 793, 1999.

F. Tjärnström: Quality Estimation of Approximate Models. Thesis No. 810, 2000.

C. Carlsson: Vehicle Size and Orientation Estimation Using Geometric Fitting. Thesis No. 840, 2000.

J. Löfberg: Linear Model Predictive Control: Stability and Robustness. Thesis No. 866, 2001.

O. Härkegård: Flight Control Design Using Backstepping. Thesis No. 875, 2001.

J. Elbornsson: Equalization of Distortion in A/D Converters. Thesis No. 883, 2001.

J. Roll: Robust Verification and Identification of Piecewise Affine Systems. Thesis No. 899, 2001.

I. Lind: Regressor Selection in System Identification using ANOVA. Thesis No. 921, 2001.

R. Karlsson: Simulation Based Methods for Target Tracking. Thesis No. 930, 2002.

P.-J. Nordlund: Sequential Monte Carlo Filters and Integrated Navigation. Thesis No. 945, 2002.

M. Östring: Identification, Diagnosis, and Control of a Flexible Robot Arm. Thesis No. 948, 2002.

C. Olsson: Active Engine Vibration Isolation using Feedback Control. Thesis No. 968, 2002.

J. Jansson: Tracking and Decision Making for Automotive Collision Avoidance. Thesis No. 965, 2002.

N. Persson: Event Based Sampling with Application to Spectral Estimation. Thesis No. 981, 2002.

D. Lindgren: Subspace Selection Techniques for Classification Problems. Thesis No. 995, 2002.

E. Geijer Lundin: Uplink Load in CDMA Cellular Systems. Thesis No. 1045, 2003.

M. Enqvist: Some Results on Linear Models of Nonlinear Systems. Thesis No. 1046, 2003.

T. Schön: On Computational Methods for Nonlinear Estimation. Thesis No. 1047, 2003.

F. Gunnarsson: On Modeling and Control of Network Queue Dynamics. Thesis No. 1048, 2003.

S. Björklund: A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems. Thesis No. 1061, 2003.

M. Gerdin: Parameter Estimation in Linear Descriptor Systems. Thesis No. 1085, 2004.

A. Eidehall: An Automotive Lane Guidance System. Thesis No. 1122, 2004.

E. Wernholt: On Multivariable and Nonlinear Identification of Industrial Robots. Thesis No. 1131, 2004.

J. Gillberg: Methods for Frequency Domain Estimation of Continuous-Time Models. Thesis No. 1133, 2004.

G. Hendeby: Fundamental Estimation and Detection Limits in Linear Non-Gaussian Systems. Thesis No. 1199, 2005.

D. Axehill: Applications of Integer Quadratic Programming in Control and Communication. Thesis No. 1218, 2005.

J. Sjöberg: Some Results On Optimal Control for Nonlinear Descriptor Systems. Thesis No. 1227, 2006.