AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

UNIVERSITÉ DE LORRAINE

École doctorale IAEM Lorraine

In-Orbit Data Driven Parameter Estimation for Attitude Control of Satellites

THÈSE

présentée et soutenue publiquement le 25 Mai 2020

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention Automatique)

par

Carlo Nainer

Composition du jury

Rapporteurs :	Xavier Bombois	Directeur de Recherche CNRS - Ecole Centrale de Lyon
	Marco Lovera	Professeur - Politecnico di Milano
Examinateurs :	Jean-Philippe Noël	Chercheur - Eindhoven University of Technology
	Hugues Garnier	Professeur - Université de Lorraine
	Marion Gilson	Professeur - Université de Lorraine
	Hélène Evain	Ingénieur de Recherche - CNES
Invité :	Christelle Pittet	Ingénieur de Recherche - CNES

Centre de Recherche en Automatique de Nancy CRAN — UMR 7039

Mis en page avec la classe thesul.

Acknowledgments

I've never been good at writing, and these acknowledgements are the perfect demonstration.

I would like to thank all the people that directly or indirectly contributed to the success of my PhD and to the drafting of this thesis.

Firstly, I would like to thank my supervisors: Marion, Hugues, Christelle and Hélène. They trusted me and guided me through this journey while giving precious advice. This work is also the result of their knowledge and experiences. I also thank them for their patience in following and correcting my work.

I would also like to thank the jury members for spending time reviewing my thesis and for their valuable remarks.

I thank the PhD students who shared this path with me. I will not list their names so as not to unfairly forget someone. I especially thank my office-mates, past and present, for the constructive discussions and the many coffee breaks.

I would like to thank particularly Paul, for all the good moments during my PhD and for making me discover climbing.

I'm also grateful to Andrea and Leonardo, for all the video-game nights that made me feel at home. Without them this thesis would have been ready much earlier.

Finally, my deepest thanks go to my parents, that supported me during all these years.

To my parents.

Contents

Nomenclature

Chapter 1
Background and Scope

1.1 Background . . . 1
1.2 Attitude Control and Satellite Subsystems . . . 4
1.3 System Identification Overview . . . 6
1.4 Scope of the Thesis . . . 7
1.5 Outline of the Thesis . . . 10

Chapter 2
Satellite System **13**

2.1 Attitude Representation and Satellite Rotational Kinematics . . . 13
2.1.1 Reference Systems . . . 14
2.1.2 Attitude Representation . . . 15
2.2 Satellite System and Dynamics . . . 17
2.2.1 Closed-loop System . . . 17
2.2.2 Satellite Attitude Dynamics . . . 17
2.2.3 Disturbance Torque . . . 19
2.3 Sensors . . . 19
2.3.1 Star Tracker . . . 20
2.3.2 Gyroscope . . . 21
2.3.3 Reaction Wheel Sensor . . . 21
2.3.4 Sensor Delay . . . 22
2.4 Attitude Control Actuators . . . 22
2.4.1 Reaction Wheels . . . 23
2.4.2 Momentum Wheels . . . 24
2.4.3 Control Moment Gyroscopes . . . 24
2.4.4 Magnetorquers . . . 24

Chapter 3
Overview of the Satellite Parameter Estimation Problem **27**

3.1 Satellite Parameter Estimation Overview 27
3.1.1 Satellite Modeling for Parameter Estimation 28
3.1.2 Parameter Estimation: Use of Inverse Model and other Considerations . . 29
3.2 Overview of the State of the Art 31
3.3 Parameter Estimation Methods 33
3.3.1 The Least Squares Estimation 34
3.3.2 The Instrumental Variable Method 35
3.3.3 The Total Least Squares Method 37
3.3.4 The Kalman Filter as Parameter Estimation Method 39
3.3.5 Optimal Instrumental Variable Estimation 40
3.4 Next Steps towards the Satellite Parameter Estimation 43

Chapter 4
Inertia Matrix Estimation for Gyro-equipped Satellite **45**

4.1 Satellite Model and Noise Analysis 46
4.2 Filtering for Derivative Estimation 47
4.3 Proposed Instrumental Variable for Inertia Estimation 48
4.4 Simulation Results 51
4.5 Solution for the Inertia Estimation with Biased Gyroscope 54
4.6 Simulation Results with Biased Gyroscope 56
4.7 Conclusions 57

Chapter 5
Inertia Matrix Estimation for Gyroless Satellite **59**

5.1 Satellite Closed-Loop System 60
5.2 Satellite Model and Noise Analysis 61
5.2.1 Satellite Model 61
5.2.2 Noise Analysis 61
5.3 Derivative Estimates 64
5.4 Prefilter Design for Instrumental Variable 65
5.5 Augmented Model and Algorithm Implementation 67
5.6 Numerical Results from High-Fidelity Simulator 69
5.7 Inertia Estimation Using Real Data 73
5.8 Conclusions 74

Chapter 6
Inertia and Actuator Alignment for Gyroless Satellite **77**

6.1 Satellite Model for Parameter Estimation . . . 77
6.2 Noise Term in the Linear Regression Model . . . 80
6.3 Filter Design . . . 82
6.4 Augmented Model and Delay Estimation . . . 85
6.4.1 Augmented Regressor and Instrument . . . 85
6.4.2 Modified Regressor for Sensor Delay Compensation . . . 86
6.4.3 Proposed Instrumental Variable Based Algorithm . . . 87
6.5 Numerical Simulations and Results . . . 87
6.6 Conclusions . . . 91

Chapter 7
Experiment Design for Satellite Inertia Matrix Estimation **95**

7.1 Problem Statement . . . 96
7.2 Parametrization Choice for The Maneuver Profile . . . 97
7.3 Experiment Design . . . 98
7.3.1 Fisher Information Matrix Based Experiment Design . . . 98
7.3.2 Noise Model Correction Filter . . . 100
7.3.3 Including Physical Constraints in the Cost Function . . . 101
7.4 Numerical and Experimental Results . . . 102
7.4.1 Numerical Simulations . . . 102
7.4.2 Zero-Gravity Experiment . . . 106
7.5 Conclusions . . . 109

Chapter 8
Conclusions and Future Works **111**

8.1 Contribution Overview . . . 111
8.1.1 Analysis, Development and Design of a Tailor-made IV Method for Satellite Parameter Estimation . . . 111
8.1.2 Satellite Inertia Estimation for Gyroless Satellites . . . 112
8.1.3 Extension to the Actuator Alignment Estimation . . . 112
8.1.4 Design of Maneuvers for Satellite Inertia Estimation . . . 113
8.2 User Parameter Guide . . . 113
8.3 Future Works . . . 114
8.3.1 Recursive Parameter Estimation . . . 114
8.3.2 Estimation with Attitude and Angular Velocity Sensors . . . 114

8.3.3 Null Space Experiment Design 114

Appendices **115**

Appendix A
Cepstrum Based Filter Synthesis **117**

Appendix B
Application of Delayed State Variables Filter **121**

Résumé en français **125**

Bibliography **131**

List of Figures

1.1 Number of satellite launches per year (source: The Economist). 2
1.2 MicroCarb Satellite from the Myriade program (source: CNES). 3
1.3 Block diagram for an attitude-control system. 5
1.4 Closed-loop satellite system with feedforward control. 8
1.5 Objectives of the work. 9

2.1 Geocentric local orbital coordinate system. 14
2.2 Example of closed-loop satellite system. 17
2.3 Star tracker with triple optical head (source: Spacecraft System Engineering [Fortescue et al., 2011]). 20
2.4 Double prism optical gyroscope (source: Spacecraft System Engineering [Fortescue et al., 2011]). 22
2.5 Common pyramidal arrangement configuration for reaction wheels. 24

3.1 Satellite system, in "open-loop". 30
3.2 Satellite model in inverse form (EIV model). 30
3.3 Satellite model in inverse form with noise term brought to the "output". 31
3.4 Box-Jenkins model. 42

4.1 IV implementation scheme with the auxiliary model. 50
4.2 Angle reference profile, represented in Euler's angles (sequence z-y-x). 51
4.3 Input-Output data: reaction wheel total angular momentum h_{rw} at the top, and satellite measured angular velocity $\overline{\omega}$ at the bottom. 52
4.4 Box plots for the errors of the inertia parameter estimates. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF. 53
4.5 Box plots of the error in the $\hat{J}_{11}$ estimate for different values of γ. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF. 53
4.6 Log-plot of the average MSE for the LS-SVF (in red) and IV-SVF (in blue) methods for different values of γ. 54
4.7 Box plots for the errors of the inertia parameter estimates with $\sigma = 34 \cdot 10^{-5}\, rad/s$. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF. 54
4.8 Box plots for the errors of the inertia parameter estimates. In red (on the left) the results for the IV-SVF (IV_1), while in blue (on the right) the results for the augmented IV-SVF (IV_2). 57

5.1 Closed-loop satellite system. The equivalent torque generated by the reaction wheels is represented by $-\dot{h}_{rw}$, while r_1 and r_2 are the reference and feedforward signals, respectively. 60
5.2 Attitude and angular rate reference profile. 70
5.3 Input (reaction wheel total angular momentum) and output (in quaternions) of the satellite. 71
5.4 Magnitude Bode plot of the filter used in the IV method. 71
5.5 Sum of the mean square errors (MSE) of the 6 inertia parameter estimates for different values of the user-parameter τ_σ. The value of τ_σ is represented as its ratio with an empirical optimal value τ_{opt}. 72
5.6 Input-output data of the Picard experiment. 74
5.7 Torque cross-validation: measured (solid lines), estimated (dash-dot lines). 75

6.1 Geometrical representation of α and β. 79
6.2 Disturbance torque spectrum approximation in solid blue line, measurement noise term in solid red line, and overall noise spectrum in dashed black line. 83
6.3 Reciprocal of the overall noise spectrum in dashed black line, and its third-order filter approximation in solid red line. 83
6.4 Estimated prefilters from the residuals. 85
6.5 Attitude and angular rate reference profile. 88
6.6 Angular momentum for each of the 4 reaction wheels (at the top). Satellite attitude expressed as Euler angles (x-y-z) at the bottom. 88
6.7 Box plot for the inertia estimate errors. 89
6.8 Box plot for the alignment estimate error vectors for each of the 4 reaction wheels. 90
6.9 Inertia error comparison. 92
6.10 Alignment error comparison. 92

7.1 Convex hull property of a cubic spline. 98
7.2 Initial reference profile used for the optimization. 103
7.3 D-opt maneuver. 104
7.4 Box plot for the inertia estimate errors. In red the results from the basic profile, in yellow the ones from the scaled basic profile, and in blue the ones of the D-opt profile. 105
7.5 CMG speeds obtained from the D-opt maneuver. 105
7.6 A picture of the SCRAT-0g platform tested on a parabolic flight (courtesy of ISAE SUPAERO and ONERA). 106
7.7 Input-Ouput data from the zero-G experiment. 107
7.8 Torque cross-validation for the zero-G experiment: measured (solid lines), one-step ahead prediction (dash-dot line). 108
7.9 Torque cross-validation for the numerical simulation: measured (solid lines), one-step ahead prediction (dash-dot line). 108

A.1 Bode plot of the true (black line) and estimated model with the proposed method (red lines) from a Monte Carlo of 100 runs. 120

B.1 The order of the nonlinear function $F(\cdot)$ and of the delay filter does not change the output. 121
B.2 Bode plot for a third-order Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$. 122

B.3 Group delay for the third-order Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$. 123
B.4 Group delay for the equalized Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$. 124

1 Schéma de contrôle d'attitude. 125
2 Satellite MicroCarb du simulateur haute fidélité (source: CNES). 128

Nomenclature

Symbols and Operators

$E(x)$	Expectation of the random vector x
$\bar{E}(\cdot)$	Total expectation. $\bar{E}(\cdot) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} E(\cdot)$
$\hat{\theta}$	Estimated model parameter vector
θ_J	Inertia parameters as vector
θ_e	Linearized misalignment parameters as vector
$\hat{\theta}$	Estimated model parameter vector
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^{n\times m}$	Set of $n \times m$ matrices with real entries
$\otimes$	Quaternion multiplication operator
a_i	i-th reaction wheel axis orientation as unit vector
θ	Model parameter vector
$F(s)$	Prefilter transfer function
h_{rw}	Reaction wheel total angular momentum (in the body frame)
h_C	Control moment gyro total angular momentum (in the body frame)
$H(s)$	Continuous-time noise transfer function
$H(z^{-1})$	Discrete-time noise transfer function
I_n	Identity matrix of size n
J	Inertia matrix
J_{rw}	Scalar reaction wheel inertia around its spinning axis
M_d	Sum of Disturbance Torques
$N(0, \sigma^2)$	Normal distribution with zero mean and variance σ^2
q	Unit quaternion vector: $q = [q_0\ ,\ q_1\ ,\ q_2\ ,\ q_3]^T$
ν	Overall noise term
ω	Angular velocity: $\omega = [\omega_x\ ,\ \omega_y\ ,\ \omega_z]^T$
Ω	reaction wheel rotation speed
z^{-1}	Delay operator. $z^{-1}x(t) = x(t-1)$
Z	Instrument

Acronyms

ADCS	Attitude determination and control system
AR	Autoregressive
ARMA	Autoregressive moving average
ARMAX	Autoregressive moving average exogenous
ARX	Autoregressive exogenous
BJ	Box-Jenkins
CAD	Computer aided design
CARMA	Continuous Autoregressive moving average
CMG	Control moment gyro
DSVF	Delay state-variable filter
EIV	Errors-in-variables
EKF	Extended Kalman Filter
FIM	Fisher information matrix
FIR	Finite impulse response
FOG	Fiber optic gyroscope
GEO	Geostationary Earth orbit
HRG	Hemispherical Resonator Gyroscope
IRWSM	Integrated random walk smoother
IV	Instrumental variable
LEO	Low Earth orbit
LIF	Linear integral filter
LS	Least squares
MEMS	Micro electro-mechanical system
MEO	Medium Earth orbit
MSE	Mean squared error
PD	Proportional-derivative
RIV	Refined instrumental variable
RW	Reaction wheel
SRIV	Simplified refined instrumental variable
SVF	State-variables filter
TF	Transfer function
TLS	Total least squares
UKF	Unscented Kalman filter

1

Background and Scope

Contents

1.1 Background **1**
1.2 Attitude Control and Satellite Subsystems **4**
1.3 System Identification Overview **6**
1.4 Scope of the Thesis **7**
1.5 Outline of the Thesis **10**

1.1 Background

A satellite is an object in space that orbits around a bigger object. There are two kinds of satellites: natural (such as the moon orbiting the Earth) or artificial (such as the International Space Station orbiting the Earth). In the following, by the term "satellite" we will implicitly mean an artificial satellite. An artificial satellite is an object that has been built and launched into orbit using rockets. There are currently thousands of active satellites orbiting the Earth.

The first artificial satellite, the Sputnik 1, was launched on 4th October 1957 by the Soviet Union and till now more than 8000 satellites have been launched by different countries. Around 5000 of them are still in orbit, out of which only 7 are revolving around planets other than Earth. Among them, only 2218 were operational at December 2019, according to the UCS (Union of Concerned Scientists) Satellite database (the typical lifetime of a satellite is around 15 years) [Angelo, 2003]. Figure 1.1 shows the trend of satellite launches from the first launch till now.

The world of satellites can be divided into two broad areas: scientific satellites and application satellites.

Scientific satellites help to acquire new information about our world, our solar system, our galaxy, and the universe.

Application satellites provide practical and business services. These are for example the communication satellites, the remote sensing satellites, the space navigation and positioning satellites, and the meteorological satellites. Thousands of application satellites have now been launched over the past half century. These satellites now represent a huge global industry.

Every satellite has some of the following basic parts [Fortescue et al., 2011]:

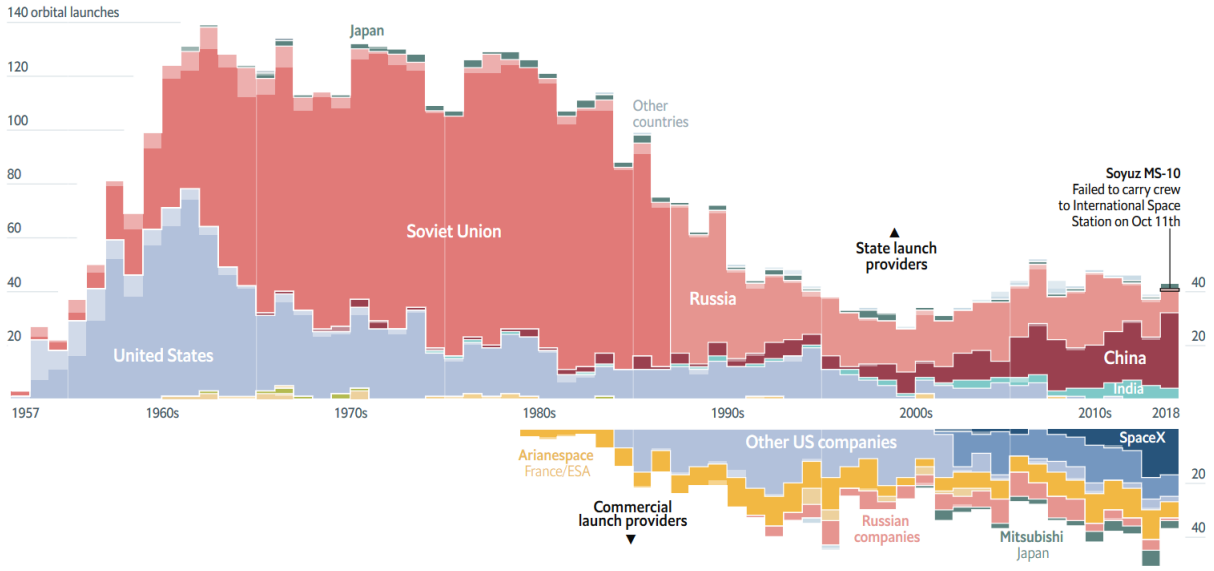


Figure 1.1: Number of satellite launches per year (source: The Economist).

- The bus: it is the frame and structure of the satellite to which all the other parts are attached.
- A power supply system: this is composed by solar panels to generate electricity and a set of batteries to store some of this energy for times when the satellite is in the shadow of the Earth.
- Heat control system: this system has the function to protect the satellite from the extremely high temperatures due to exposure to the Sun. Also the electrical components of the satellite produce heat that must be dissipated. Heat is dissipated by reradiating it towards the deep space.
- Computer system: this system has several functions. It controls the satellite operation, it monitors and controls the orbit, the attitude and the temperature, but many other functions exist, depending on the purpose of the satellite.
- Communication system: this system allows the satellite to send and receive data to ground stations or to other satellites.
- Attitude control system: this is the system that keeps a satellite pointed in the right direction. Gyroscopes and reaction wheel actuators are commonly used to change orientation. Star sensors are commonly used to determine the direction a satellite is pointing.
- A propulsion system: it is composed by a rocket engine which is used to place the satellite into the correct orbit. Once in orbit, satellites may use small rockets called thrusters to slightly correct the orbit.
- A payload: the payload depends on the purposes of satellites. It may be a camera, a telescope, or some scientific sensors for research purposes.

Satellites may vary widely in size. Some cube satellites are as small as $10\,cm \times 10\,cm \times 10\,cm$. Some communication satellites are about $7\,m$ long and have solar panels that extend up to $50\,m$.

The largest artificial satellite is the International Space Station (ISS). Including solar panels, it is 109 m wide.

A mass based satellite classification is the following [Fortescue et al., 2011]:

- Large satellites: with a mass greater than 1000 kg
- Medium satellites: from 500 to 1000 kg
- Small satellites: less than 500 kg
  - Minisatellite: from 100 to 500 kg
  - Microsatellite: from 10 to 100 kg
  - Nanosatellite: from 1 to 10 kg
  - Picosatellite: from 0.1 to 1 kg

Among microsatellites, an important class is represented by the Myriade program of CNES[1] (Figure 1.2). This platform has been designed in order to offer to scientists the basis of a versatile tool for testing small payload instruments, up to a mass of 60 kg, for short-duration missions, with short development time and reduced costs. Most of the simulations in this work make reference to satellites of the Myriade platform. For example, in Section 5.7, the proposed algorithm is applied to real telemetry data from the Picard satellite, a spacecraft launched in 2010 with the objective to accurately measure the solar activity. Other simulations use the parameters of the MicroCarb satellite, which will be launched in 2021 with the goal to monitor the fluxes of carbon dioxide at the surface between the atmosphere and the oceans, as well as the vegetation.

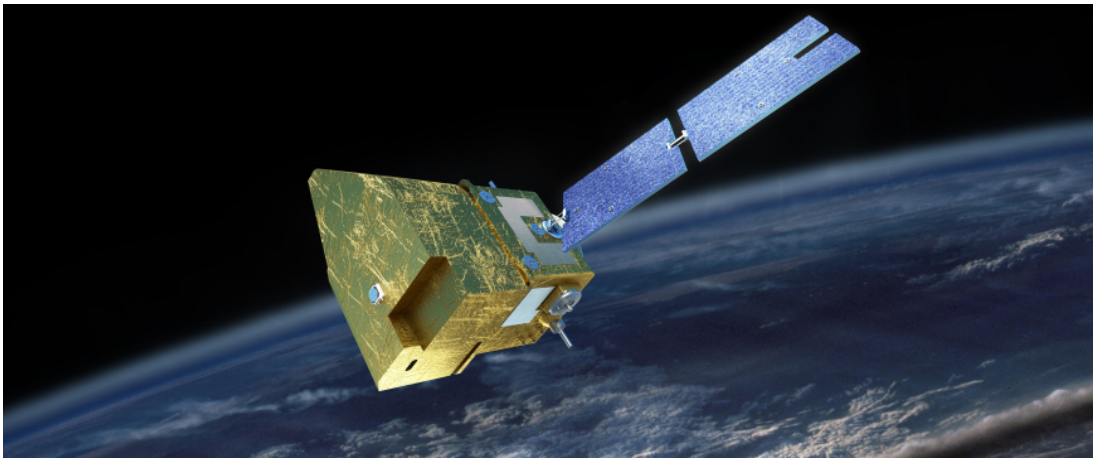


Figure 1.2: MicroCarb Satellite from the Myriade program (source: CNES).

The altitude of a satellite above the Earth's surface also varies and depends on the mission the satellite is designed for.

- Low Earth orbit (LEO): from 200 to about 1000 km. A LEO is above the outer limits of the atmosphere (about 80 km), and well below the hazardous Van Allen radiation belts, the innermost of which begins at about 2400 km. For example, the ISS orbits at 390 km. The rotation period of satellites in this orbit is in a range between 90 and 110 minutes. The

[1] French Space Agency

choice of an orbit for a LEO remote sensing spacecraft is governed by the mission objectives and payload operational requirements. Very often, to achieve a near-global coverage, the plane of the orbit must be inclined at about 97° with respect to the Earth's equator. This is for example the orbit of the *Sun synchronous satellites*, which are characterized by the fact of passing over the same point of Earth at the same local time. Their orbit has a precession with a period of one year in order to always maintain the same orientation with respect to the sun with the advantage to see the Earth surface with almost the same lighting conditions.

- Geostationary orbit (GEO): its altitude is $35786\,km$ and the time for one orbit is exactly $24\,hours$. This is to match the rotation of the Earth so that the satellite appears to stay above the same point on the Earth's equator. Communication satellites and global weather satellites are placed in geostationary orbit because of the large portion of the Earth's surface visible from that altitude and the fact that ground stations do not have to track the satellite, which appears fixed in the sky. Since Earth rotates at a constant angular rate, a geostationary satellite must move at a constant speed in its orbit, therefore a geostationary orbit must be circular [Wertz, 2012].

- Medium Earth orbit (MEO): Although over 90% of the satellites are positioned in LEO or in GEO, the space between these two most popular orbit can be the ideal environment for a smaller subset of satellites. Satellites in this middle region, denominated MEO, have a larger footprint than LEO satellites (meaning that they can see a wider part of the Earth surface at a time) and a lower transmission lag than GEO satellites. One important example of MEO satellites is the Global Positioning System (GPS).

1.2 Attitude Control and Satellite Subsystems

Once in orbit, satellites could need an adjustment of their orbit and a stabilization/correction of their attitude.

Orbit adjustment could be necessary for different reasons. It may be necessary for deorbiting a satellite at the end of his life or to compensate the effect of perturbing forces (e.g. atmospheric drag). Sometimes a maneuver called orbital phasing is necessary, moving a spacecraft to a different location within the same orbit, as in a rendezvous to a space station in a docking maneuver.

Attitude is the three-dimensional orientation of a spacecraft with respect to a specified reference frame. The ability to point and do so accurately is essential in a satellite built for applications that involve looking at or measuring something on the ground or in space. In general, the pointing accuracy required of the spacecraft structure is determined by its mission. Pointing errors affects the quality of the image in an observation satellite, for example (image blur is related with the pointing control stability). The most demanding application regarding the pointing accuracy are the astronomy missions. For example, the James Webb Telescope is expected to achieve a pointing accuracy of $10^{-6}\,deg$ or $0.0036\,arcsec$ [Wertz, 2012].

The pointing accuracy, agility, pointing stability, as well as the budget of the mission, condition the choice of attitude control system and the choice of sensors and actuators. Different types of sensors are available. The attitude can be measured by Sun and Earth sensors, with a limited accuracy, or by star tracker with very high accuracy. Several types of gyroscopes are available, with different levels of accuracy, that provide angular velocity measurements. Also the different

types of actuators affect the pointing accuracy: magnetorquers (also known as magnetic torque rods), thrusters, cold gas, electric propulsion, reaction wheels and control moment gyros.

Some satellite subsystems must be pointed towards specific directions, as solar arrays must be pointed to the Sun, thermal radiators must be pointed to deep space, antennas pointed to ground stations and, prior to firing, thrusters must be pointed to the correct direction. The required orientation will frequently be related to an Earth-based frame of reference, for example one face of the spacecraft must point down the local vertical. In general, separate pointing mechanisms are to be avoided. However, they are often needed in order to enable a part of the satellite to remain pointed in one direction whilst the main structure changes its orientation. For example, a solar array needs to remain pointed to the Sun whilst the main structure turns to align with the local vertical.

The attitude determination and control system (ADCS) plays an indispensable role in the satellite on-orbit operation which could greatly affect the satellite performance. The ADCS includes sensors, actuators, computer, software, and ground support equipment. Its general structure is represented in Figure 1.3. The links between components identify the major interactions, while arrows indicate the flow of information.

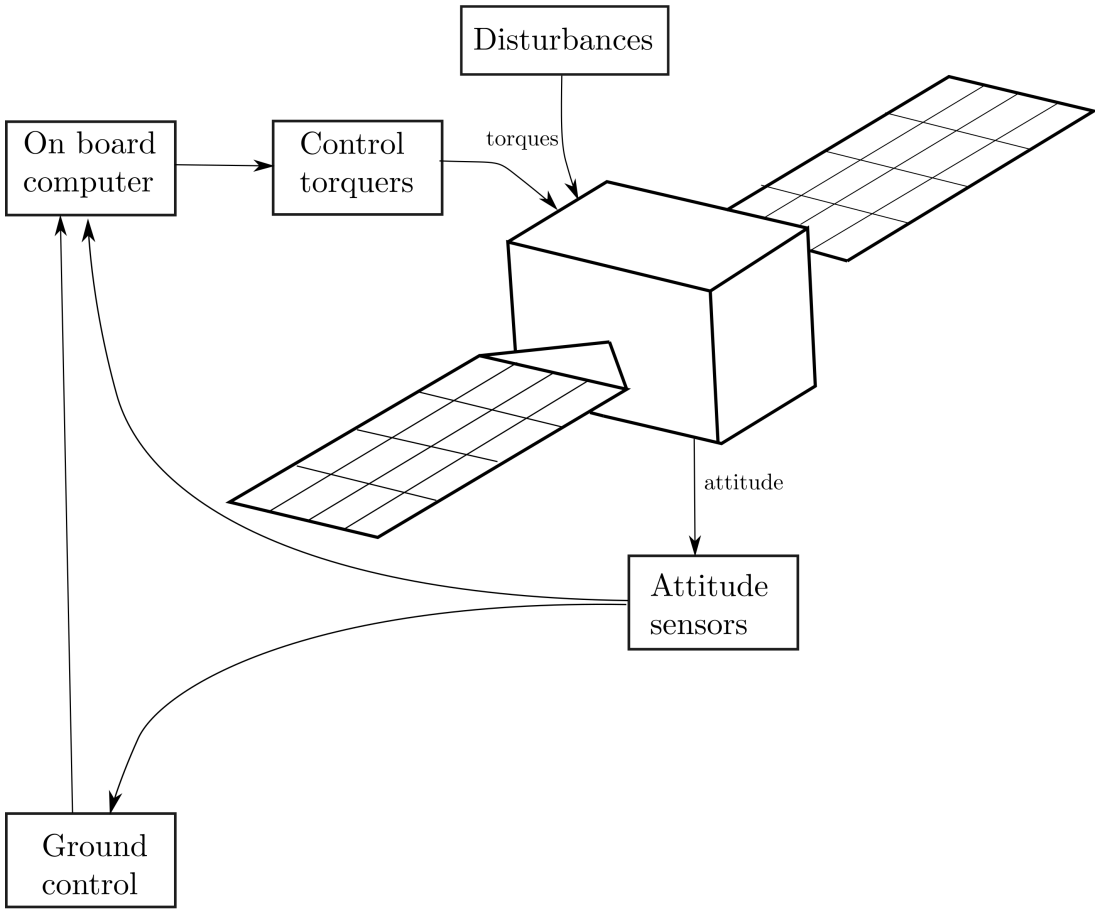


Figure 1.3: Block diagram for an attitude-control system.

The torques arising from moments of forces about the center of mass, or pure torques, can be distinguished as external or internal to the spacecraft. The former affect the satellite total angular momentum, whereas the latter affect only the distribution of momentum between its

moving parts (as in the case of reaction wheels or control moment gyros). Since the internal sources are subject to saturation (saturations that are mainly due to disturbance torques that increase the total angular momentum), it is necessary to include controllable external torquers, at least for desaturation purposes.

1.3 System Identification Overview

The accurate knowledge of the satellite model is fundamental for an optimal tuning of the ADCS. This knowledge can be obtained by means of system identification techniques. The basic objective of system identification is to develop a mathematical model of a physical system by observing its input-output relationship. A mathematical model of a system is required for many purposes, such as diagnosis, simulation, prediction, and control. Often in practice, one concerns with linear models since many physical systems can be described by linear or approximately linear equations. Moreover, linear systems are easier to handle.

The construction of a model from input-output data involves the following basic steps [Ljung, 1999]:

- The selection of the input-output data through a specific design of experiment. The objective of the experiment design is to make the data maximally informative, subject to possible physical constraints. This step is covered in Chapter 7.

- The selection of a set of candidate models. This is the one of the most important and difficult choices of the system identification procedure. Models can be classified according to their physical meaning (we consider here only parametric models) [Isermann and Münchhof, 2010]:

  white box: If the physical laws that govern the behavior of a system (or any other theoretical considerations about the system) are completely known, they can be used to construct a so called *white-box* model, also called *first principles model*.

  black box: At the other end of the modeling scale there are the so called *black-box* models. A black-box model does not directly use the actual mechanisms which are supposed to underly the studied system, but it merely intends to make good predictions of the system behavior: the internal description merely serves to relate input signals to output signals.

  grey box: An intermediate level between white and black box modeling is also possible. This approach is often denoted by the term *grey-box* modeling. In this case the model is constructed from basic physical principles and the parameters represent unknown values of the system coefficients that, at least in principle, have a direct physical interpretation.

  In this work the satellite is represented as a grey-box model. The physical laws that govern the satellite dynamics are well known and they have been used to construct the model as described in Chapter 2. Some physical parameters of this model, as the inertia matrix, are not exactly known and must be estimated. Moreover, the disturbances that affect the satellite are the overall effect of several complex physical effects, which are quite difficult to model with sufficient reliability. In this case a black-box approach has been chosen, where a reasonable and simplified model is used, based on some insights instead of an accurate knowledge of the reality.

- Selection of the estimator: a suitable cost function is chosen, which reflects how well the model fits the experimental data. Among the selection criteria, these are the most commonly used (we consider here only the time domain estimators):
  - Prediction error minimization: the model is selected based on the smallest prediction errors when applied to the observed data.
  - Least squares criterion for the linear regression: The least-squares method is a special case of the prediction-error minimization method. The important feature of this criterion is that it produces a quadratic function in the parameter and therefore it can be minimized analytically.
  - Maximum likelihood optimization: The model that maximizes the probability of the observed data is defined maximum likelihood estimator.
  - Instrumental variables: It is based on a cost function very similar to the least squares estimator. Like the least squares method, it has an analytical solution.
- The selection of the best model in the set, based on the data. This is an optimization problem, which is solved to obtain the numerical values of the model parameters.
- Model validation: this is the assessment of model quality, which indicates how the model performs when it attempts to reproduce the measured data.

Some additional operations are often necessary before performing an identification experiment and after collecting the data, in order to polish the data to make them suitable for system identification. This step is referred to as data pre-processing. Below, some preprocessing procedures are described:

- Decimation: When the sampling frequency is too high with respect to the actual bandwidth of interest of the system under investigation, one may resample the data by selecting one every i-th sample from the original data sequences.
- Detrending the data: This process consists in removing long term drifts or biases.
- Pre-filtering the data: Distortions in the data, as trends and noise, that reside in a frequency band that can be separated from the frequency band of interest of the system can be removed by filtering the data with a suitable filter. Applying such filtering to the input and output sequences has some implications for the identification procedure.
- Removing the outliers. Outliers are data points that differ significantly from other observations. The removal of the outliers helps to improve the accuracy of the estimated results.

1.4 Scope of the Thesis

The correct tuning of the attitude determination and control system is of primary importance for the satellite operation. An accurate and fast pointing is only possible with a precise and accurate knowledge of satellite inertia and of the actuator directions. The importance of these parameters becomes even more critical when a feedforward control is used in addition to the closed-loop control, as shown in Figure 1.4[2]. This control scheme has the advantage to minimize

[2]The second inputs in the sensor boxes, q_n and $\tilde{\omega}$, represent the sensor measurement noise, whereas M_d represents the overall disturbance torques.

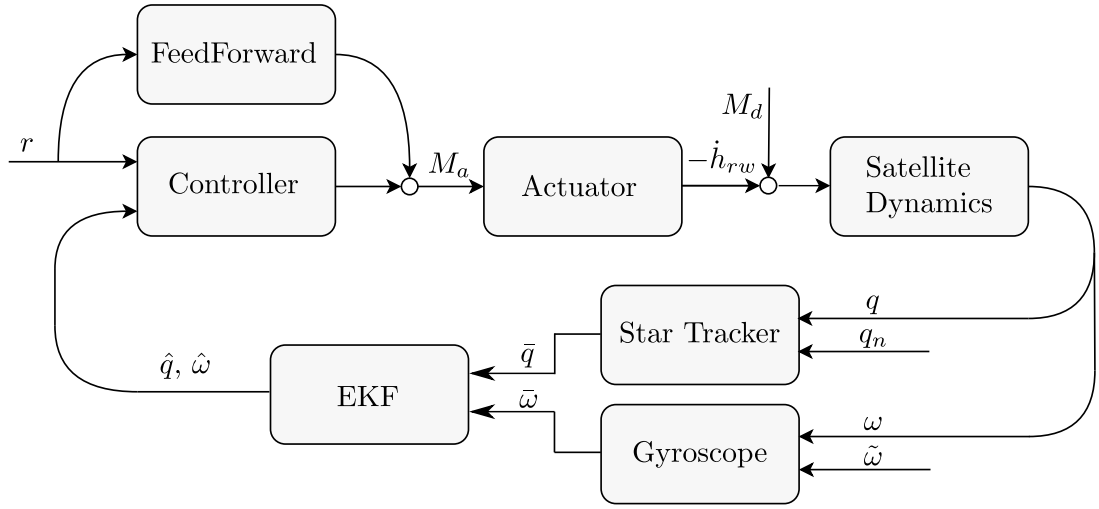


Figure 1.4: Closed-loop satellite system with feedforward control.

the negative effects of a known disturbance and to improve the reaction time and the control accuracy. In fact, even if the robustness of closed-loop controllers is usually guaranteed even for uncertainties of 15% on the inertia matrix, inertia errors below 1% would allow for a faster pointing with lower tranquilization time after the maneuver.

Unfortunately, the inertia matrix and the actuator alignments cannot be reliably estimated on ground. The gravity on earth prevents the possibility to accurately measure the inertia. Moreover, these parameters change after launch, due to fuel consumption, deployment of the solar panels and to a possible deformation of the satellite structure. Also the actuator alignment may be affected by the structural deformation. The estimation of these parameters must therefore be done in orbit, using the telemetry data obtained during a suitable maneuver.

The attempt to provide methods and algorithms for obtaining the best possible estimation of these parameters is the main topic of this thesis.

The work can be divided into two main parts, as shown in Figure 1.5: the search for a method that produces reliable and consistent estimators and the attempt to improve the quality (reduce the variance) of the estimates by means of filtering and a proper selection of the excitation maneuver (also known as experiment design).

In order to develop an estimation algorithm, several non-idealities of the satellite system and its telemetry data must be considered:

- a measurement noise is present on all the sensors (attitude sensors and actuator sensors), which cannot always be approximated by a simple white noise source;
- several external disturbances are present, which can hardly be estimated;
- some internal disturbances could be present (i.e. oscillation of the solar panels or of some internal moving part);
- some constraints limit the amplitude and the speed of the satellite maneuvers: physical constraints, due to actuators sizing with respect to satellite inertia, limit the angular momentum and angular acceleration; operational constraints (such as sensors blinding, or thermal constraints) limit the reachable attitude;

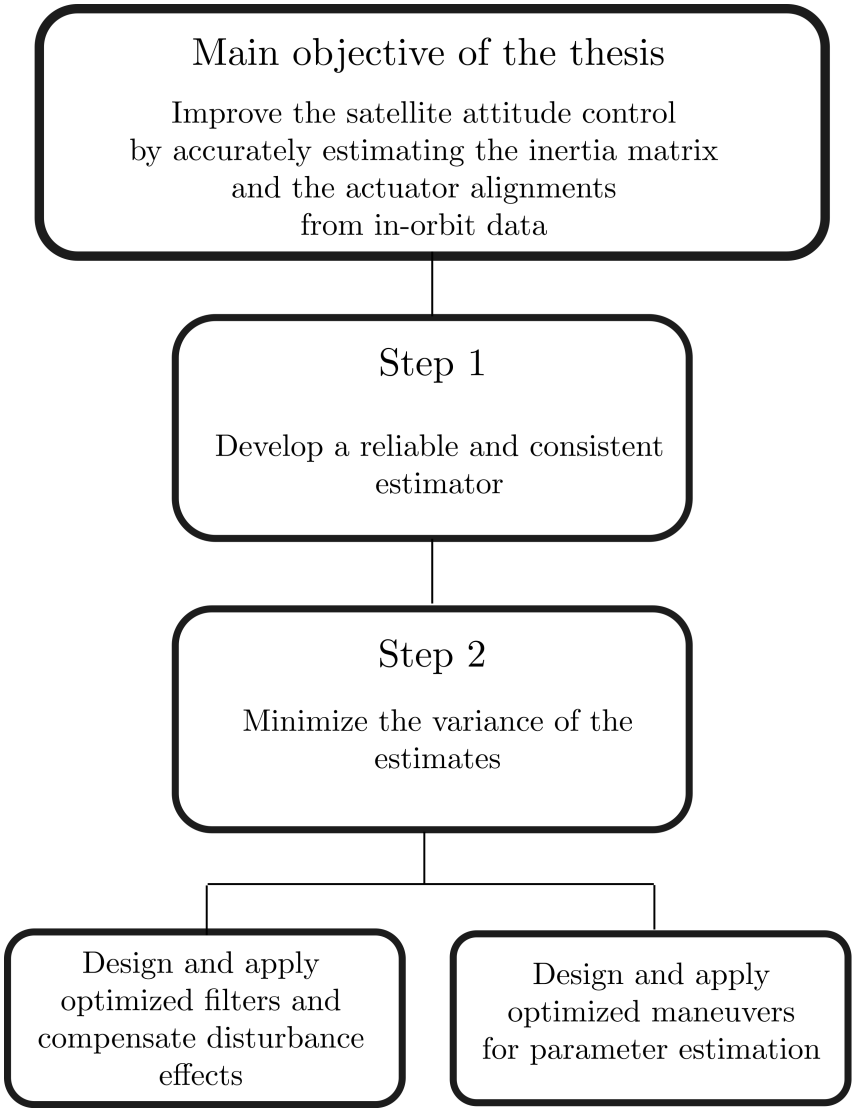


Figure 1.5: Objectives of the work.

- some output states like angular acceleration and angular velocity (for gyroless satellites) are not directly measured;
- due to the double integration behavior of the spacecraft (which makes the open-loop system unstable), the satellite is always operating in closed loop;
- a small, but not negligible, timing difference could be present among different sensor measures.

1.5 Outline of the Thesis

The sequence of the chapters in this thesis does not follow the logic flow described in Section 1.3. It is instead more related to the chronology of the activities in this doctoral work. First a reliable and consistent method has been searched considering a relatively simple estimation problem. Then several difficulties and further objectives have been gradually added. Finally, after satisfying estimation methods have been obtained, the experiment design has been taken into consideration, with the objective to further improve the accuracy of the estimates.

The organization of the chapters of the thesis is presented below.

- Chapter 2 begins with a brief description of the equations which govern the satellite attitude kinematics and dynamics. Then the satellite system is described with a focus on the components of the attitude determination and control system: sensors and actuators.

- Chapter 3 illustrates the estimation problem. After a description of the satellite model for parameter estimation, an overview of the possible estimation methods is presented, with a focus on the main method chosen for this work, the instrumental variable (IV) method.

- In Chapter 4, the inertia estimation problem for a gyro-equipped satellite is considered. This is the most common scenario considered in the literature. The main objective consists in showing how the instrumental variable can also be adapted and applied to this estimation problem. Moreover, the effect of the gyroscope drift on the estimates has been analyzed and a solution has been proposed to estimate and remove the drift effect. Part of this work was published in [Nainer et al., 2018].

- In Chapter 5, the instrumental variable method is extended to gyroless satellites, therefore the algorithm relies solely on attitude measurements. Additionally, a detailed noise analysis is performed that allows for an effective choice of the prefilter. For the numerical simulation, all the data for this chapter have been generated by a high-fidelity simulator from CNES, where all possible sources of disturbance are present. The simulation conditions are therefore as close as possible to the real environment. Finally, the method is tested with real data recorded during an experiment on the Picard satellite. This work, with the exception of the Picard results, was published in [Nainer et al., 2019].

- In Chapter 6, a more general case is studied: estimating both the inertia and actuator alignment parameters while still considering a gyroless satellite. Moreover, in the simulated data, the star tracker measurements are not perfectly synchronized with the actuator speed readings. This particular problem is not dealt by the satellite parameter estimation literature, although it affects the accuracy of the estimates. The IV method is further extended, for the simultaneous estimation of both parameters, as well as to take into consideration aforementioned delay. Finally, a more elegant solution is proposed for the prefilter design. The proposed algorithm is tested again using data from the CNES high-fidelity simulator. This work has been submitted in [Nainer et al., 2020b]

- Chapter 7 shifts the focus to the experiment design for satellite inertia estimation. The purpose of the experiment design is to maximize the information content of the data by means of optimized maneuvers. After the parameterization of the signal, maneuvers are designed by the minimization of a Fisher information-based cost function. The effectiveness of the optimized maneuvers has been tested through simulations. Moreover, also real experiments have been carried out on a small platform, that well emulates the satellite

behavior, which has been tested on a zero-G flight. This work has been submitted and accepted in [Nainer et al., 2020a].

- Finally, in Chapter 8 an overview of the contributions related to the satellite parameter estimation problem is given, as well as some guides to correctly use the proposed methods. Moreover, possible future works, especially regarding the extension of the proposed methods, are suggested.

2

Satellite System

Contents

2.1 Attitude Representation and Satellite Rotational Kinematics 13
2.1.1 Reference Systems 14
2.1.2 Attitude Representation 15
2.2 Satellite System and Dynamics 17
2.2.1 Closed-loop System 17
2.2.2 Satellite Attitude Dynamics 17
2.2.3 Disturbance Torque 19
2.3 Sensors 19
2.3.1 Star Tracker 20
2.3.2 Gyroscope 21
2.3.3 Reaction Wheel Sensor 21
2.3.4 Sensor Delay 22
2.4 Attitude Control Actuators 22
2.4.1 Reaction Wheels 23
2.4.2 Momentum Wheels 24
2.4.3 Control Moment Gyroscopes 24
2.4.4 Magnetorquers 24

The initial step of the parameter estimation process consists in the satellite modeling. In this research we are interested in the estimation of the parameters which affect the attitude control, therefore only the rotational dynamics is considered, while the translational dynamics is not part of the thesis identification work. In the first section, the satellite rotational kinematics is introduced, giving the bases for the second section where the satellite system and its rotational dynamics are described. In the last two sections the main sensors and actuators are illustrated, with a focus on their mathematical model for the measurements that is later used in this thesis.

2.1 Attitude Representation and Satellite Rotational Kinematics

This section introduces the equations and concepts of rotational attitude kinematics. These equations form the basis for spacecraft dynamical systems. A brief review of the main concepts is presented in this chapter, with a particular emphasis on the quaterion representation.

2.1.1 Reference Systems

The orientation of a spacecraft is usually referred to as "attitude". It is described by the orientation of a body fixed frame with respect to an inertially fixed reference frame. The attitude is described by the attitude coordinates, and consists of at least three parameters. Before introducing the attitude representation it is necessary to define reference frames and sign conventions [*TTVS* 2019]. All reference frames are right handed and with mutually orthogonal axes. At least two reference frames need to be defined: a body-fixed reference frame, that is rigidly attached to the satellite body, and a world reference frame that is fixed in the inertial space.

Body frame: A satellite-fixed coordinate system depends on the mechanical architecture of the satellite. Its origin may be a characteristic point of the satellite (usually the launcher interface) and its axes are oriented according to the principal sides of the bus. Secondary coordinate systems tied to the satellite can be defined, in particular at anchor points of moving parts on the main body of the satellite (e.g. anchor points of the solar panels) or at attitude sensors (e.g. star tracker). Finally, instrument coordinate systems can be defined, that are aligned according to the characteristic directions of the instrument (e.g. optical axes of an imaging instrument, boresight of an antenna, measurement axes of a gyroscope, ...).

Reference frames: The most commonly used reference frames are:

- *Local orbital coordinate system*: Orbital coordinate systems are tied to the orbit of the satellite and its position on that orbit. Two types of orbital coordinate systems can be defined: a coordinate system that refers to the geocentric axis, where the Z axis points to the Earth center, and one that refers to the satellite velocity, where the X axis is parallel to the velocity (these two reference systems coincide in case of circular orbit). In both cases, the Y axis is parallel to the angular momentum (Figure 2.1).

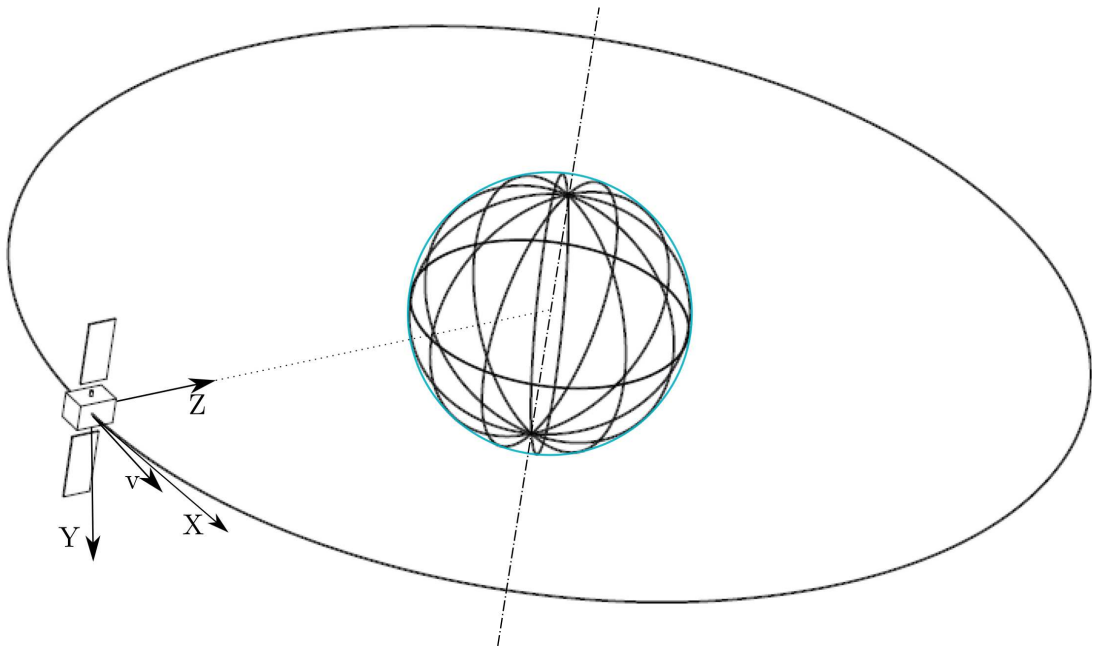


Figure 2.1: Geocentric local orbital coordinate system.

- *Inertial coordinate systems*: The inertial coordinate system coincides with the International

Celestial Reference System recommended by the IAU Working Group on Nomenclature for Fundamental Astronomy. It is aligned close to the mean equator and dynamical equinox of the J2000 epoch, in which the new star catalogues are expressed.

- *Solar coordinate system*: In certain fields, such as thermal and energy, the direction of the Sun plays a predominant role. This is why the Earth-Sun direction is used in some coordinate systems.

2.1.2 Attitude Representation

The attitude of a rigid body is defined as the transformation between a body frame and an inertial frame. There are several different ways to represent this transformation. The advantages and disadvantages of some important and commonly used representations are summarized below:

- **Rotation matrices**: This is the most general form for representing the attitude of a body. Since the 3×3 rotation matrix uses nine parameters to represent three angular degrees of freedom, there are six independent constraints on the matrix elements. Each column (and row) is a unit vector, which gives 3 constraints and the columns (and rows) are orthogonal to each other, yielding another 3 constraints. This parametrization is very redundant compared to all other parameterizations. Rotation matrices are therefore computationally more expensive than other representations. However, they have the advantage that there are no singularities or ambiguities.

- **Euler angles**: The Euler angles representation is based on the fact that three coordinate rotations in sequence can describe any rotation. Not all combination of three axes produce valid rotations sequences (two consecutive rotations about the same axis is not a valid sequence). Therefore only 12 combinations of coordinate rotation can be used to define an attitude and the Euler angles representation must specify the sequence of rotation axes (the sequences of rotation about three different axes are also called Cardan angles). This representation is very intuitive and easy to interpret physically. However, it suffers from a singularity also known as gimbal lock: some particular orientations can be obtained with different combinations of angles. The exact orientation at which gimbal lock occurs depends on the order of rotations used. This causes another problem: for some particular orientation one of the Euler angles may change suddenly in response to small change in rotation.

- **Unit quaternions**: Due to the parametrization by 4 variables instead of the 3 of the Euler angles, quaternions do not have singularity problems. Also, any rotation sequence can be represented by a continuous quaternion trajectory and do not suffer any discontinuity like Euler angles. They are also computationally more efficient than Euler angles. However, they possess the ambiguity of dual covering, i.e., the quaternions q and $-q$ represent the same transformation. It is also easier to interpolate between rotations and to chain rigid transformations.

Other representations exist, even if less common. In this work quaternions are used for almost all the computation, while the Euler's angles are sporadically used for plots. Some concepts and definitions used in quaternion algebra are described below.

Quaternions

Quaternions can be considered as a 4-components extended complex number of the form:

$$q = q_0 + q_1 i + q_2 j + q_3 k\,, \tag{2.1}$$

where the complex units i, j and k respect the following rules:

$$i^2 = j^2 = k^2 = -1\,, \qquad ij = k\,, \qquad jk = i\,, \qquad ki = j\,. \tag{2.2}$$

A quaternion is usually represented by a a vector q which contains its components:

$$q = [q_0\, q_1\, q_2\, q_3]^T = \begin{bmatrix} q_0 \\ q_{1:3} \end{bmatrix}\,, \tag{2.3}$$

where q_0 and $q_{1:3}$ are the real and imaginary parts, respectively. In some texts a different convention is used, where the real part is put in the fourth element of the vector.

The conjugate of a quaternion q is defined as:

$$\bar{q} = \begin{bmatrix} q_0 \\ -q_{1:3} \end{bmatrix}\,, \tag{2.4}$$

and the norm is defined as:

$$||q|| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}\,. \tag{2.5}$$

The product of two quaternions q and p is denoted by $q \otimes p$ and can be calculated using the rules (2.2) as:

$$q \otimes p = \begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} = Q(q)p\,. \tag{2.6}$$

It must be noted that, in general, $q \otimes p \neq p \otimes q$.

A unit quaternion ($|q| = 1$) can be used to represent a rotation in space. This can be easily understood considering the Euler's rotation theorem:

Any rotation or sequence of rotations of a rigid body or coordinate system about a fixed point is equivalent to a single rotation by a given angle about a fixed axis (called Euler axis) that runs through the fixed point.

Any attitude of a rigid body can be defined by an axis with unit vector v, and a rotation about that axis, α. The corresponding rotation quaternion can be represented in terms of these two parameters as:

$$q = \begin{bmatrix} cos(\alpha/2) \\ v_{1:3} sin(\alpha/2) \end{bmatrix}\,. \tag{2.7}$$

Quaternion derivatives are often used to compute the angular velocity or to integrate the system dynamics. The body's angular velocity, $\omega \in \mathbb{R}^{3\times1}$, and angular acceleration, $\dot{\omega} \in \mathbb{R}^{3\times1}$, (expressed in the body reference frame) can be directly computed from the attitude quaternion as

$$\omega = 2W(q)\dot{q}\,, \qquad \dot{\omega} = 2W(q)\ddot{q}\,, \tag{2.8}$$

with the function $W(\cdot)$ defined as follows

$$W(q) = \begin{bmatrix} -q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix}\,. \tag{2.9}$$

2.2 Satellite System and Dynamics

The first step of the parameter estimation process consists in the satellite modeling. In this research we are interested in the estimation of the parameters which affect the attitude control, therefore just the rotational dynamics is considered.

2.2.1 Closed-loop System

Due to the double integration behavior of the rotational dynamics, the satellites are always controlled in closed loop (Figure 2.2), while a feedforward action is sometimes added to improve the attitude tracking and spacecraft overall agility. Filters, like extended Kalman filter (EKF), are present in the feedback path, to improve the measurement signal-to-noise ratio, to ensure good function in case of missing measurements and to estimate states, which are not measured, necessary for the control law (e.g. angular velocity in case of gyroless satellites) [Crassidis et al., 2007; Wertz, 2012].

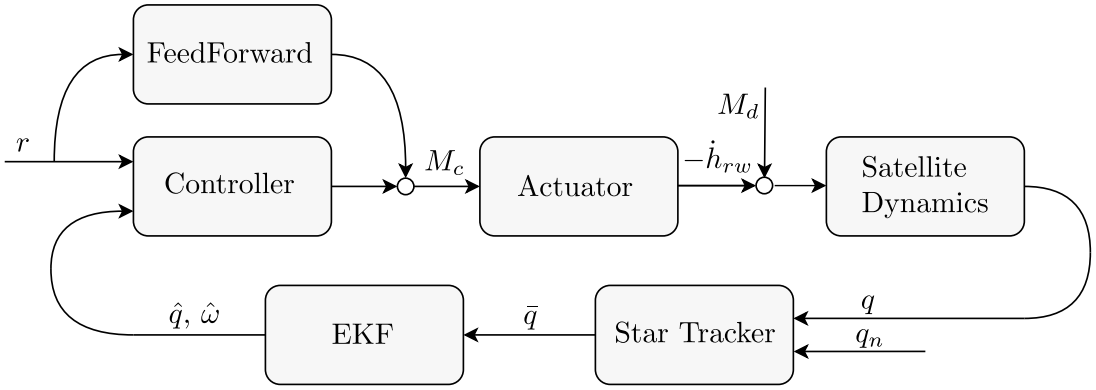


Figure 2.2: Example of closed-loop satellite system.

2.2.2 Satellite Attitude Dynamics

For the rotational dynamics, the satellite is considered as a rigid body. The flexible modes of the satellite are not significant enough to justify a more complex model (the control law already avoids the spacecraft excitation around critical frequencies), and the small mismatch can be easily included into a disturbance torque term.

The rotational dynamics of a rigid body, expressed in the body reference frame, is described by the Euler's equations [Sidi, 1997]

$$M(t) = J\,\dot{\omega}(t) + \omega(t) \wedge J\,\omega(t)\,, \tag{2.10}$$

where $M(t) \in \mathbb{R}^{3\times1}$ is the sum of the applied torques, $J \in \mathbb{R}^{3\times3}$ is the inertia matrix, $\omega(t) = [\omega_x(t)\,,\,\omega_y(t)\,,\,\omega_z(t)]^T \in \mathbb{R}^{3\times1}$ is the body angular velocity and $\wedge$ is the cross product operator. The inertia J is a symmetric positive definite matrix with the following form

$$J = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{12} & J_{22} & J_{23} \\ J_{13} & J_{23} & J_{33} \end{bmatrix}\,, \tag{2.11}$$

therefore there are only 6 independent inertia parameters.

In the following, we will use the matrix form of the cross product multiplication and (2.10) can be rewritten as

$$M(t) = J\,\dot{\omega}(t) + \omega(t)^{\times} J\,\omega(t)\,, \tag{2.12}$$

where $\omega(t)^{\times}$ is defined as follows

$$\omega(t)^{\times} = \begin{bmatrix} 0 & -\omega_z(t) & \omega_y(t) \\ \omega_z(t) & 0 & -\omega_x(t) \\ -\omega_y(t) & \omega_x(t) & 0 \end{bmatrix}. \tag{2.13}$$

Depending on the type of satellite, the commanded torque $M(t)$ can be generated by different types of actuators. In some rare cases, thrusters can be used to generate the required torque. However, even if high torques can be achieved, this method is usually avoided since the amount of fuel in the satellite is limited and it should be preserved to increase the duration of the mission. The most common way to control the attitude of a satellite is by means of reaction wheels (RWs) or control moment gyros (CMGs) [Larson and Wertz, 1992]. These types of actuators exchange their angular momentum with the spacecraft body to generate an equivalent torque. In this case, the rotational dynamics equation becomes

$$-\dot{h}_{rw}(t) - \omega(t)^{\times} h_{rw}(t) + M_d(t) = J\,\dot{\omega}(t) + \omega(t)^{\times} J\,\omega(t)\,, \tag{2.14}$$

or

$$-\dot{h}_{C}(t) - \omega(t)^{\times} h_{C}(t) + M_d(t) = J\,\dot{\omega}(t) + \omega(t)^{\times} J\,\omega(t)\,, \tag{2.15}$$

where $h_{rw}(t)$ and $h_C(t)$ are the reaction wheels total angular momentum and CMGs total angular momentum, respectively. The sum of all the disturbance torques affecting the satellite during the maneuver are collected into the term $M_d(t)$. Since the dynamics is very similar for the two cases, we will consider a satellite equipped with reaction wheels, unless differently stated. For the reaction wheel case, the total angular momentum is given by

$$h_{rw}(t) = \sum_{i=1}^{n} J_{rw,i}\Omega_i(t)a_i\,, \tag{2.16}$$

where n is the number of reaction wheels (typically $n = 4$), Ω_i is the i-th wheel rotation speed, $J_{rw,i}$ is the i-th reaction wheel scalar inertia around its spinning axis, and $a_i \in \mathbb{R}^{3\times 1}$ is the reaction wheel axis orientation expressed as a unit vector.

Due to the gyroscopic terms the overall satellite dynamics is nonlinear. However, during a satellite maneuver, the gyroscopic term is typically a couple of order of magnitude smaller than the main linear term. As an example to prove this statement, we can consider the Pleiades-HR (High Resolution) satellite from CNES [Damilano, 2001]. It is a medium size satellite, given its mass of around $900\,kg$, however, thanks to its 4 powerful control moment gyros it can be considered an agile satellite. Considering its maximum achieved angular velocity of $3.4\,^{\circ}/s$ [Gleyzes et al., 2012], and that its diagonal inertia terms are between 500 and $700\,kg\,m^2$, the gyroscopic term $(\omega(t)^{\times} J\,\omega(t))$ may reach values up to around $0.4\,Nm$. However, even in this extreme scenario, the gyroscopic term is still significantly smaller than the maximum achievable torque, $45\,Nm$. Since during the maneuvers relatively high input torques are applied, we consider the nonlinearity for the typical satellite dynamics to be "mild".

2.2.3 Disturbance Torque

The disturbance torque M_d depends on several factors (Table 2.1) [Larson and Wertz, 1992]: gravity gradient torque, aerodynamic torque, solar radiation pressure, magnetic torque, internal satellite structure oscillation (e.g. solar panel oscillation). The magnitudes of these components depend on the satellite orbit and altitude. The magnitude of these torques in space is small when compared to terrestrial standards. However, even very small torques become significant when there is no friction to oppose them and when the orientation has to be very accurate. Terms, like the gravity gradient torque, can be easily estimated, while others can only be taken into consideration by the parameter estimation process in order to avoid poor parameter estimates.

Disturbance	**Type**	**Influenced primarily by**	**Order of magnitude**
Gravity gradient	Constant for Earth oriented satellites, cyclic for inertially oriented satellites	Spececraft inertia matrix and orbit altitude	$4 \times 10^{-5} Nm$
Solar radiation	Cyclic for Earth oriented satellites, constant for solar oriented satellites	Spacecraft geometry and surface reflectivity	$5 \times 10^{-6} Nm$
Magnetic field	Cyclic	Orbit altitude, residual satellite magnetic dipole, orbit inclination	$5 \times 10^{-5} Nm$
Aerodynamic	Constant for Earth oriented satellites, cyclic for inertially oriented satellites	Orbit altitude and spacecraft geometry	$3 \times 10^{-6} Nm$

Table 2.1: Main disturbance torques affecting a spacecraft. The order of magnitude is referred to a microsatellite in LEO.

For a low-orbit satellite, the overall disturbance torque can be well approximated as

$$\begin{aligned} M_{d,x} &\simeq M_{0,x} + M_{1,x} sin(\omega_0 t + \phi_{1,x}) + M_{2,x} sin(2\omega_0 t + \phi_{2,x}) \,, \\ M_{d,y} &\simeq M_{0,y} + M_{1,y} sin(\omega_0 t + \phi_{1,y}) + M_{2,y} sin(2\omega_0 t + \phi_{2,y}) \,, \\ M_{d,z} &\simeq M_{0,z} + M_{1,z} sin(\omega_0 t + \phi_{1,z}) + M_{2,z} sin(2\omega_0 t + \phi_{2,z}) \,, \end{aligned} \tag{2.17}$$

where ω_0 is the satellite orbital rate. The other scalar constants in these equations can be determined empirically.

2.3 Sensors

The attitude determination system is one of the most important component for any spacecraft. It has the function of estimating the orientation of the spacecraft with respect to some reference system. The satellite attitude estimation, during mission modes, relies almost always on star tracker measurements and on rate gyroscope measurements. Nonetheless, there are also acquisition modes where sun sensors and magnetometers are used. Theoretically, only one of these two sensors (star tracker and gyroscope) could be used for attitude and angular rate determination. However, the calculation of the angular rate from the star tracker is limited by the low dynamic of this sensor, while the calculation of the attitude by the integration of the angular rate is

seriously compromised by the gyroscope drift. The fusion of measurements from star trackers and gyroscopes is therefore a common approach for spacecraft attitude determination.

2.3.1 Star Tracker

Star trackers are the most common attitude sensors. A star tracker is an optical device that measures the orientation of the satellite with respect to distant (fixed) stars using photocells or a camera. Since the positions of many stars have been accurately measured by astronomers, a star tracker is a very accurate device. In order to do this, the star tracker must obtain an image of the stars, measure their apparent position in the reference frame of the spacecraft, identify the stars and compare their position with their known absolute position from a star catalog. A star tracker includes a processor to identify stars by comparing the observed pattern with the stored pattern of stars in the sky. Sometimes more than one optical head is used in order to avoid the possibility to be completely blinded by the sun or the Moon (see Figure 2.3).

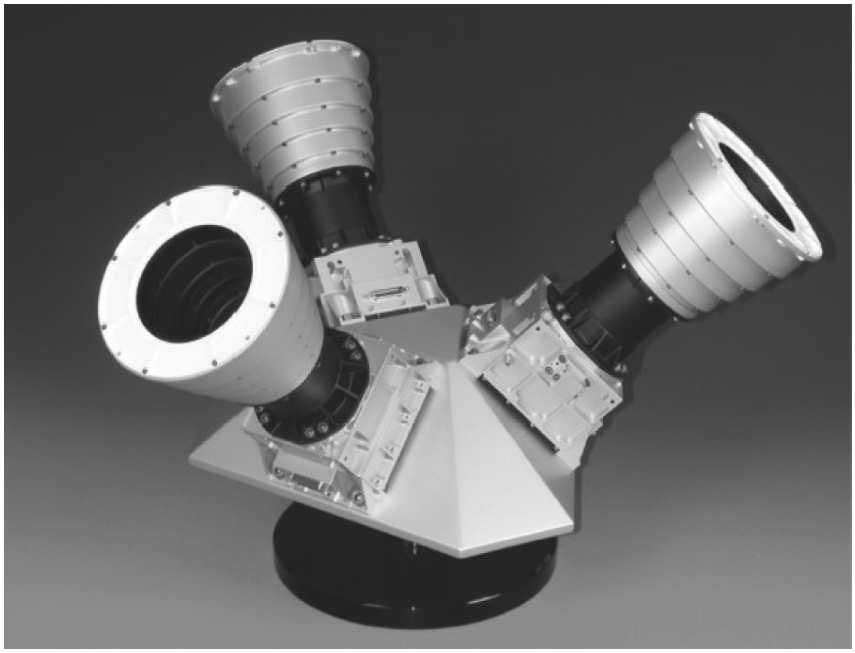


Figure 2.3: Star tracker with triple optical head (source: Spacecraft System Engineering [Fortescue et al., 2011]).

Since a star tracker uses a fixed inertial frame as reference, its measurements do not drift with time. The attitude measurement equation can be expressed as follows[3]

$$\overline{q}(t_k) = q_n(t_k) \otimes q_s \otimes q(t_k)\,, \tag{2.18}$$

where $\overline{q}(t_k)$ is the measured quaternion representing the satellite attitude, q_s is the quaternion transformation from body to sensor frame, and $q_n(t_k)$ is the quaternion measurement noise that can be approximated as

$$q_n(t_k) \simeq [1\,,\, e_x/2\,,\, e_y/2\,,\, e_z/2]^T\,, \tag{2.19}$$

[3]From now on, with the symbol $\overline{q}$ we mean the measured quaternion and not the conjugate of the quaternion q.

with

$$e_x \sim N(0,\, \sigma_x^2)\,, \qquad e_y \sim N(0,\, \sigma_y^2)\,, \qquad e_z \sim N(0,\, \sigma_z^2)\,, \tag{2.20}$$

where the notation $e \sim N(0, \sigma^2)$ means that the error e has a normal distribution with zero mean and variance σ^2. Moreover, the three components e_x, e_y, e_z are assumed to be mutually uncorrelated. Since the star tracker may be affected by some small bias/misalignment (b_x, b_y, b_z), a more complex noise term can be defined

$$q_n(t_k) \simeq \left[1\,,\, \frac{e_x + b_x}{2}\,,\, \frac{e_y + b_y}{2}\,,\, \frac{e_z + b_z}{2}\right]. \tag{2.21}$$

This is the noise model that is used by the CNES high fidelity simulator (see Chapter 5 and 6).

2.3.2 Gyroscope

Sensors which measure angular rates with respect to an inertial frame of reference are called gyroscopes. There are three main types of gyroscope: mechanical gyroscopes, micro-electro-mechanical system (MEMS) gyroscopes, and optical gyroscopes. Traditional mechanical gyroscopes utilize a rotating momentum wheel attached to a gimbal structure. However, rotating wheel gyroscopes have many disadvantages, primarily concerning bearing friction and wear. Another critical point of the mechanical gyroscopes is the so called gimbal lock: in case of alignment of two gimbals, one degree of freedom is lost and the device becomes locked. Vibrating gyroscopes, such as the Hemispherical Resonator Gyroscope (HRG) present an effective solution to the bearing problems by eliminating rotating parts. Finally, the most accurate (and expensive) gyroscopes are the optical gyroscopes, as the Fiber-Optic Gyroscope (FOG), which exploit the Sagnac effect. An example of optical gyroscope is shown in Figure 2.4.

The simplest gyroscope noise model consists in considering the angular velocity affected by a white noise $\tilde{\omega}$

$$\overline{\omega} = \omega + \tilde{\omega}\,. \tag{2.22}$$

A more realistic model considers also the drift and gain error terms [*TTVS* 2019]

$$\overline{\omega} = (1 + f_G)\omega + \omega_b + \tilde{\omega}\,, \tag{2.23}$$

where f_G is a constant scalar value, and ω_b is modeled as a random walk term:

$$\dot{\omega}_b = \omega_d\,, \quad \omega_d \sim N(0, \sigma_d^2)\,. \tag{2.24}$$

2.3.3 Reaction Wheel Sensor

In order to compute the reaction wheel angular momentum, the reaction wheel speed Ω_i must be measured. Usually, these actuators are equipped with Hall-effect sensors or optical encoders (the latter are used in MicroCarb satellite). The encoder measures a wheel angular position, from which the angular rate is computed. Even if the encoder quantization error has a uniform distribution, it is often approximated with an additive white noise $\tilde{\Omega}_i$

$$\overline{\Omega}_i(t_k) = \Omega_i(t_k) + \tilde{\Omega}_i(t_k)\,. \tag{2.25}$$

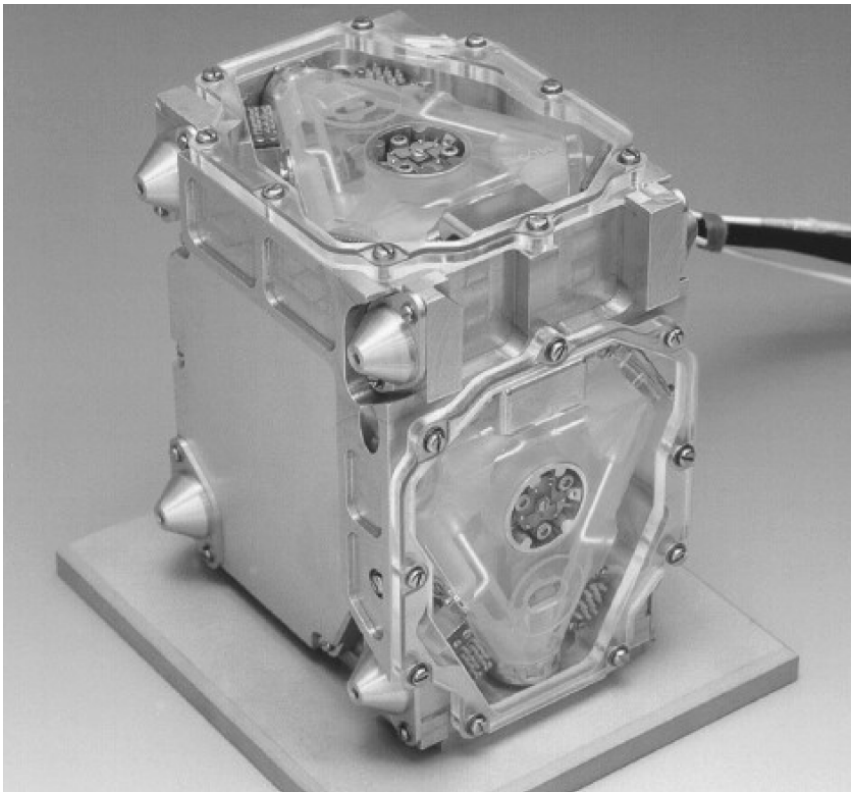


Figure 2.4: Double prism optical gyroscope (source: Spacecraft System Engineering [Fortescue et al., 2011]).

2.3.4 Sensor Delay

When the input-output data are obtained, it is usually assumed that the data are correctly synchronized. However, this is not always necessarily true. In fact, we may have a small (less than a sample step) delay t_δ between the input-output data (e.g. reaction wheel speed and star tracker measurements may be not perfectly synchronized). Even if small, this delay may have a non-negligible effect on the parameter estimation and it should be taken into account by the estimation algorithm.

2.4 Attitude Control Actuators

There are two main methods for attitude control: active methods, which control the satellite by using torque actuators, and passive methods which use the reaction of the satellite to some physical effect to ensure its stability (e.g. gravity gradient stabilization) or which use properties of the angular momentum (spin stabilization). Passive methods can only provide coarse attitude control, suitable for simple satellites, which only require basic alignment. Therefore, in order to achieve accurate and fast satellite pointing, active control techniques are present in most of the spacecrafts.

The torques acting on a satellite can be identified as external or internal to the spacecraft. The former will affect its total angular momentum, whereas the latter will affect only the distribution of momentum between its moving parts.

One main issue of the internal torquers is their saturation limit. A reaction wheel, for

example, has a maximum speed that can be achieved and this corresponds to a maximum angular momentum that can be stored. If a disturbance torque has a non-zero mean, then the wheel speed will increase until saturation to counteract the disturbance. It then becomes necessary to apply an external torque to bring down the wheel speed (desaturation); this must be achieved with an external type actuator (a thruster or a magnetorquer).

The main types of actuators [Larson and Wertz, 1992] are summarized in Table 2.2. Other actuators exist, such as sailor sails and aerodynamic control surfaces, however their use is very limited. In this thesis we consider satellite controlled by inertial actuators: reaction wheels or control moment gyros. Satellites are also equipped with magnetorquers, however, for the majority of the simulations in this work, they are kept disabled.

Actuator type	**Typical torque magnitude**	**Remarks**
Thrusters	Up to 10000 Nm	External torque On-off operation only Requires fuel
Magnetorquers	up to 0.18 Nm	External torque No torque about field direction Torque is altitude and latitude sensitive
Reaction wheels	0.01 - 1Nm	Internal torque Continuous fine pointing capability Non-linearity at zero speed
Control moment gyroscopes	25 - 500 Nm	Internal torque Complex mechanism Steering complexity due to Singularities

Table 2.2: Main types of actuators for attitude control.

2.4.1 Reaction Wheels

Reaction wheels (RWs) are internal torquers, suitable for attitude control but not for controlling the total momentum. These devices are precision-engineered wheels that rotate about a fixed axis, with a built-in motor. The basic design comprises a bearing unit, a flywheel mass and a DC motor in a vacuum housing. The stator is controlled by the drive electronics and the speed is usually measured by an encoder. Reaction wheels provide a reaction torque which results from rotational acceleration or deceleration of the flywheel (also a small gyroscopic effect is present, which should be considered in the torque calculation). Reaction wheels have a nominal zero speed, and may be rotated in either direction in response to the control torques required by the attitude controller. However, it should be noted that at low or zero angular rate, the wheels have a non-linear behavior due to bearing sticking friction, which could produce an irregular motion on the spacecraft in this region. This problem is often solved by setting the nominal operating speed of the wheels above zero rate, at a few rpm (revolution per minute). For three-axis control, three orthogonal reaction wheels are the minimum requirement. A redundant fourth wheel is normally added in order to avoid a single-point failure. In this case a pyramidal arrangement configuration is often used (Figure 2.5). There is a maximum angular momentum that a reaction wheel can store, which depends on the flywheel inertia and on its maximum speed. Before reaching this limit, a desaturation becomes necessary, which can be done by means of an external type torquer.

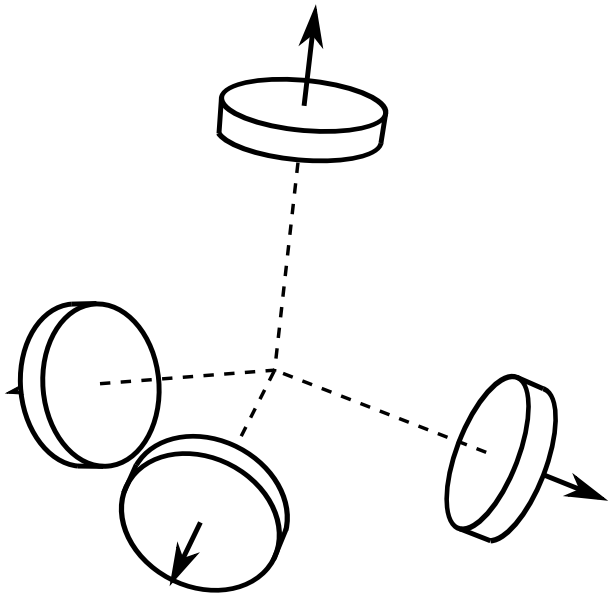


Figure 2.5: Common pyramidal arrangement configuration for reaction wheels.

2.4.2 Momentum Wheels

A momentum wheel is structurally very similar to a reaction wheel, but there is an important difference: the rotation speed of the flywheel is much higher and so, accordingly, its angular momentum. Since the total angular momentum of a spacecraft equipped with momentum wheels is kept different from zero, an important stabilization effect is obtained (also known as gyroscopic stiffness). Momentum wheels are usually smaller and achieve a lower torque with respect to reaction wheels.

2.4.3 Control Moment Gyroscopes

A control moment Gyroscope (CMG) is essentially a gimbal mounted momentum wheel. Torques are generated by commanding gimbal rotations by means of torque motors and thereby changing the spin axis orientation. Depending on the number of gimbals, there are single, 2 or 3 degree of freedom CMGs. Using three gimbals, all three components of torque may be generated by a single wheel, however they are less efficient that single-gimbal CMG. The torques achievable by these actuators may be up to 100 times higher than the ones of a reaction wheel for the same power consumption. Compared to reaction wheels, CMGs have a higher mechanical complexity (mainly due to the steering mechanism) and require much more volume and mass. They are preferably used in satellites that require high agility or in large size spacecrafts.

2.4.4 Magnetorquers

Magnetic torquers (or magnetorquers) are very popular due to their low mass, high reliability and efficiency. The downside is that they cannot produce the large torques required for fast maneuvers with the power available. Another limit of magnetic torquers is that they cannot produce a torque component about the local field direction. The torque strength can be controlled by means of the current flowing through the electromagnets. Since the torque generated by the magnetorquers depends of the Earth magnetic field strength and direction, a magnetometer

is also necessary. Magnetorquers are commonly used in satellites orbiting at LEO since their effectiveness decreases at the higher altitudes (the strength of the Earth's field weakens with height). Since magnetorquers generate an external torque, they are often used to desaturate the reaction wheels or CMGs.

3

Overview of the Satellite Parameter Estimation Problem

Contents

3.1 Satellite Parameter Estimation Overview 27
3.1.1 Satellite Modeling for Parameter Estimation 28
3.1.2 Parameter Estimation: Use of Inverse Model and other Considerations . 29
3.2 Overview of the State of the Art 31
3.3 Parameter Estimation Methods 33
3.3.1 The Least Squares Estimation 34
3.3.2 The Instrumental Variable Method 35
3.3.3 The Total Least Squares Method 37
3.3.4 The Kalman Filter as Parameter Estimation Method 39
3.3.5 Optimal Instrumental Variable Estimation 40
3.4 Next Steps towards the Satellite Parameter Estimation 43

This chapter deals with the general description of the satellite parameter estimation problem, while also providing some overview on the main estimation methods.

In the first section, after an analysis on the identifiable parameters, the most common satellite estimation models are presented, including the one that is used in the remainder of this work. Moreover, in the same section, the main difficulties of this estimation problem are highlighted. The second section gives an overview on the state of the art on the satellite parameter estimation problem. In the third section, several parameter estimation methods are briefly summarized with a deeper focus on the instrumental variable (IV) since it is the base for the methods proposed in Chapter 4-5-6. Finally, after having justified the choice of the IV method, the last section summarizes the main steps toward the satellite parameter estimation.

3.1 Satellite Parameter Estimation Overview

In order to have a reliable model of the satellite, the main parameters of the rotational dynamics need to be accurately known. By looking at (2.14) and (2.16) (spacecraft with reaction wheels as main actuators), the main satellite parameters are: the inertia matrix J, the reaction wheel inertias $J_{rw,i}$ and the reaction wheel alignments a_i. The first question that needs to be answered

is whether all these parameters can be estimated together (identifiability problem). Unfortunately, there is a scaling relation between the actuator inertia $J_{rw,i}$ and the satellite inertia J, making it impossible to estimate both simultaneously. In fact, by applying the same scale factor to the satellite and reaction wheel inertia the dynamic equation would have the same solution. Therefore, one of these two set of parameters need be considered known and kept fixed. Typically, the reaction wheel inertias can be accurately estimated on ground and their values are not expected to change after launch, thus we can simply consider all the inertias $J_{rw,i}$ known. Moreover, as explained in Section 1.4, reliable estimation of the satellite inertia and of the actuators alignments is only possible when the satellite is in orbit.

Given these assumptions, the overall estimation problem consists in performing an in-orbit estimation of the satellite inertia J and actuator alignments a_i from telemetry data. The identification process requires two main steps: first, a sufficiently exciting maneuver needs to be designed and executed in order to have enough information content in the data, secondly, an estimation method is applied on the input-output data to estimate the parameter characterizing the satellite rotational dynamics.

3.1.1 Satellite Modeling for Parameter Estimation

Before designing an estimation method, a suitable model structure of the satellite needs to be chosen. For simplicity, we first consider the inertia parameter estimation, and only in a second step extend the problem to the actuator alignments. From (2.10), we can write the output $\dot{\omega}(t)$ as function of the input torque $M(t)$

$$\dot{\omega}(t) = J^{-1}(M(t) - \omega(t)^{\times} J\,\omega(t))\,. \tag{3.1}$$

This is called "direct" or "forward" model, since it computes the output state ω as a function of the input $M(t)$. This representation is useful to simulate the model or to estimate the satellite states. However, this form has a major disadvantage when used for the identification since the parameter J appears nonlinearly in the equation. Nonetheless, estimation methods like the extended Kalman filter (EKF) can be applied to this nonlinear model, and good results have been obtained in the literature (see Section 3.2). However, the nonlinearity makes the use of other methods impractical (e.g. least squares), or it makes some methods unreliable due to highly nonconvex cost functions (e.g. prediction error method).

By keeping the system in its "inverse" form (2.10 or 2.14), the equation can be rearranged in order to linearly extract the 6 independent inertia parameters as follow

$$-\dot{h}_{rw}(t) - \omega(t)^{\times} h_{rw}(t) + M_d(t) = \Big(\Gamma(\dot{\omega}(t)) + \omega(t)^{\times}\Gamma(\omega(t))\Big)\,\theta_J\,, \tag{3.2}$$

where

$$\Gamma(\omega(t)) = \begin{bmatrix} \omega_x(t) & 0 & 0 & 0 & \omega_z(t) & \omega_y(t) \\ 0 & \omega_y(t) & 0 & \omega_z(t) & 0 & \omega_x(t) \\ 0 & 0 & \omega_z(t) & \omega_y(t) & \omega_x(t) & 0 \end{bmatrix}, \tag{3.3}$$

and

$$\theta_J = [J_{11}\,,\, J_{22}\,,\, J_{33}\,,\, J_{23}\,,\, J_{13}\,,\, J_{12}]^T. \tag{3.4}$$

Another way to extract linearly the inertia parameters consists in using the equation based on the conservation of the angular momentum in the inertial reference frame:

$$Q_r(t_k)\Big(J\,\omega(t_k) + h_{rw}(t_k)\Big) = Q_r(t_j)\Big(J\,\omega(t_j) + h_{rw}(t_j)\Big) + \int_{t_j}^{t_k} M_d(t)\,dt\,, \tag{3.5}$$

where t_j and t_k are two different time-instants, and $Q_r(t)$ is the rotation matrix that transforms the coordinates from the body reference frame to the inertial reference frame. In (3.5), the inertia J appears linearly and it can be easily extracted

$$Q(t_k)h_{rw}(t_k) - Q(t_j)h_{rw}(t_j) - \int_{t_j}^{t_k} M_d(t) = \Big(Q(t_j)\Gamma(\omega(t_j)) - Q(t_k)\Gamma(\omega(t_k))\Big)\,\theta_J\,. \tag{3.6}$$

The representation (3.6) corresponds to (3.2), with a change of reference frame and with a linear integration filter (LIF) applied to both side of the equation. The LIF has a low-pass effect behaviour [Garnier et al., 2003] and its cut-off frequency depends on the time difference Δt between the two chosen time instants t_j and t_k. A larger Δt will results in a stronger low-pass effect, conversely, a small Δt will have a weak low-pass effect. This time choice is a user parameter that has a significant impact on the estimation, since with a large Δt the model will be highly sensitive to disturbance torques ($M_d(t)$), while a small Δt makes the model more sensitive to measurement errors.

Even if the use of (3.6) is rather popular, in this work we use the model (3.2) directly coming from the Euler's equations and keep the inverse form of the satellite dynamics. This representation allows for additional freedom in the choice of filters since we do not impose the use of a LIF.

3.1.2 Parameter Estimation: Use of Inverse Model and other Considerations

In order to analyze and develop the estimation method, some further considerations must be made:

- The satellite is controlled in closed loop, as shown in Figure 2.2, therefore the presence of the output noise may have some effects on the input [Ljung, 1999; Young, 2011; Forssell, 1999]. A similar estimation problem in the robotic field has been treated as a closed-loop identification problem [Janot et al., 2013; Brunot et al., 2018]. However, the presence of the Kalman filter (Figure 2.2) in the loop removes most of the so-called "circulatory noise", thus the feedback signal contains no significant trace of the noise process(es) [Veen et al., 2013].

  In fact, the typical issue in closed-loop identification is due to the feedback which correlates the input signal ($\dot{h}_{rw}$ for the satellite case) with the output noise ($\tilde{\omega}$). The presence of a filter in the feedback highly reduces the noise on the signal received by the controller, and consequently almost no noise effect is present on the system input[4].

  It seems therefore permissible to treat the system as operating in open loop, considering the actuator input torque $-\dot{h}_{rw}(t)$ uncorrelated with the noise on the attitude quaternions ($q_n(t)$) and on the angular velocity ($\tilde{\omega}(t)$)[5].

[4] Simulations have been performed and this assumption has been verified for this application.

[5] However, it should be noted that all the proposed methods in this thesis work also in the classical closed-loop settings (with no filter in the feedback loop).

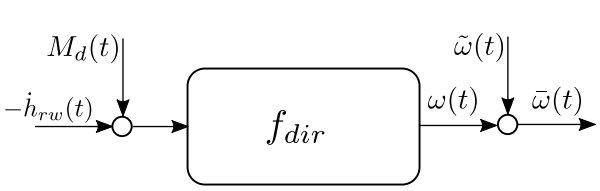


Figure 3.1: Satellite system, in "open-loop".

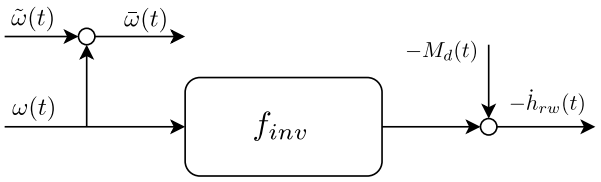


Figure 3.2: Satellite model in inverse form (EIV model).

- The open-loop system is represented in Figure 3.1. However, we are interested in the model written in its inverse form (Figure 3.2) since it allows to linearly extract the parameters to be estimated (as shown in the previous section). The system equation can be written in the following form:

$$g\Big(h_{rw}(t), \dot{h}_{rw}(t), \omega(t)\Big) = f\Big(\omega(t), \dot{\omega}(t)\Big)\,\theta - M_d(t)\,, \tag{3.7}$$

  or

$$g\Big(h_{rw}(t), \dot{h}_{rw}(t), q(t), \dot{q}(t)\Big) = f\Big(q(t), \dot{q}(t), \ddot{q}(t)\Big)\,\theta - M_d(t)\,, \tag{3.8}$$

  depending whether a gyroscope or a star tracker is used for the output data, and where θ is a vector containing all the parameters.

  The model in linear regression form[6], as in (3.7) and (3.8), allows the use of several different estimation methods.

  However, once the sensor noise is considered, the model presents several similarities with the errors-in-variables model [Söderström, 2007], as it can be seen in Figure 3.2, and, unfortunately, most methods do not yield consistent estimates in these conditions.

  Nonetheless, the satellite measurement noise term (whether it is on the angular velocity or on the attitude) can be brought to the right side of the model, as shown in Figure 3.3, where the input of the system is the measured (and therefore noisy) angular velocity $\overline{\omega}(t)$. In Figure 3.3, $f_n(\cdot)$ is a function that cancels out the effect of the noise $\tilde{\omega}$ passing through $f(\cdot)$. It must be noted that since $f(\cdot)$ is nonlinear, $f_n(\cdot) \neq f(\cdot)$. The corresponding equation in linear regression form, for the case of a gyro equipped satellite, is

$$g(h_{rw}(t), \dot{h}_{rw}(t), \overline{\omega}(t)) = f(\overline{\omega}(t), \dot{\overline{\omega}}(t))\,\theta - M_d(t) + \nu(t)\,, \tag{3.9}$$

  where the term ν collects the effects of the output sensor noise.

  This allows a better analysis of the effect of the disturbance and measurement noises. The filter design of Chapter 4 and 5 is based on such analysis.

- Finally, the signals are assumed to be uniformly sampled with a sampling interval T_s, obtaining N equation of the form

$$g(t_k) = f(t_k)\theta + \nu(t_k) - M_d(t_k)\,. \tag{3.10}$$

[6]It must be noted that in this work a slightly unconventional notation has been used, since the regressor $f(t) \in \mathbb{R}^{3\times 6}$ and not $\mathbb{R}^{6\times 3}$. The transpose operator is therefore not necessary in this equation.

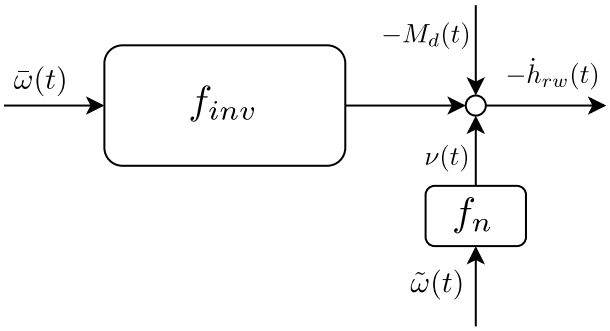


Figure 3.3: Satellite model in inverse form with noise term brought to the "output".

3.2 Overview of the State of the Art

Even if not widely developed as the attitude determination topic, there are several works in the literature dealing with the satellite parameter estimation. The majority of these works deal with the inertia estimation, being the inertia usually the parameter known with the worse accuracy. Some works consider also the actuator alignment estimation, either alone or combined with the inertia estimation. In a few works also other parameters are estimated, as the dipole moment of the spacecraft.

The most common approach in the literature consists in the use of least squares method. In [Lee and Wertz, 2002] the authors describe a test and an estimation algorithm applied to the Cassini probe during its flight to Saturn. A simple slewing maneuver has been commanded, and a least square algorithm, based on the angular momentum conservation has been applied. The algorithm uses the angular rate and the commanded torque.

The least squares, again in its batch form, is also used in [Tanygin and Williams, 1997], where, in addition to the inertia matrix, also the position of the center of mass is estimated. To avoid dealing with the non-linearity of the Euler's equation (in its direct form) with respect to the parameter, the equation is projected to the angular rate. Unlike all the other works, the coasting maneuvers are exploited to obtain the data for the estimation, instead of specific imposed trajectories.

In [Keim et al., 2006] the authors work with the Euler's equations. In this case the angular rate is not directly available and it is obtained from the attitude quaternions. A second order filter is applied to the quaternions in order to avoid performing the differentiation. To ensure the positiveness of the estimated inertia matrix, instead of using a least squares method, the authors solved a constrained optimization problem.

Online implementations of the least squares are also proposed, as in [Wilson et al., 2002] and [Manchester and Peck, 2017]. In the latter the estimation problem is formulated as a semidefinite programming with the constraint of positivedefiniteness of the inertia matrix and the triangular inequality applied to its diagonal elements to ensure a physically valid result. Both the batch and the recursive estimations have been considered. For the latter a QR decomposition is exploited.

Another online algorithm is proposed in [Bergmann et al., 1987], where a second-order filter (in a form similar to a Kalman filter) is developed to extract the inertia matrix, the center-of-mass location vector and the inverse of the mass. However the gyroscopic term has been neglected.

In [Psiaki, 2005] the algorithm's estimation equation is based on an integrated version of Euler's equations expressed in inertial coordinates. A wide set of attitude related parameters is estimated, which include moments and products of inertia, scale factors, alignments, and biases

for all reaction wheels and magnetic torque rods. The estimation problem includes the statistics of unmodeled torques and sensor errors, and it incorporates a scalar quadratic constraint to overcome the unobservability of the parameters overall scaling. The solution algorithm employs an iterative Newton method.

Adaptive attitude control algorithms is an alternative approach to estimate the inertia matrix, like in [Ahmed et al., 1998] and [Thienel et al., 2008]. In [Ahmed et al., 1998] an adaptive feedback control algorithm has been developed to provide tracking of commanded spacecraft motion. The algorithm assumes no initial knowledge of the inertia matrix and the estimation is achieved adaptively during the tracking of a maneuver which ensures a persistency of excitation.

In the literature several works can be found based on Extended (EKF) or Unscented (UKF) Kalman filters: in [Norman et al., 2011] the inertia and actuator alignment estimation problem is considered and the state dynamics is modeled as a discrete-time system, allowing for the implementation of several standard sub-optimal filtering approaches, such as an Extended Kalman Filter (EKF) or Extended Square-Root Information Filter (ESRIF). In order to make the actuator alignment observable, both kinetic energy and angular momentum measurement equations are included in the development of the filter.

In [VanDyke et al., 2004] and [Sekhavat et al., 2009] an UKF has been preferred. The latter work is particularly interesting, since it also presents an approach for synthesizing a set of optimized maneuvers. An unscented dual state-parameter Kalman filter is used to estimate the inertial parameters of a control moment gyro actuated spacecraft. The commanded trajectories are designed to minimize the condition number of an excitation matrix derived from the system dynamics in order to improve the convergence rate of the parameter estimator and to reduce the sensitivity of the estimation process to external disturbances and/or measurement noise.

A gyroless satellite is considered in [Yoon et al., 2017], where an extended Kalman Filter is used to estimate attitude and body rate as well as the attitude parameters, which include sensor and actuator alignments, spacecraft body moment of inertia and many others. In this work the gravity gradient torque and the magnetic torque are considered in the estimation, while the sum of all other disturbance torques are modeled as a drift. Since both the inertia matrix and the moment of inertia of the RWs are estimated, there is an observability issue that has been solved by imposing a value to J_{11}. A similar observability problem occurs trying to estimate both the star tracker alignment and the actuator alignment.

A least correlation method is proposed in [Jun et al., 2010] The paper, does not mention explicitly satellites, but the topic is exactly what is needed: identification of the inertia matrix of a rotating body based on an errors-in-variables model. This paper, as the previous paper of the same authors [Jun and Bernstein, 2006], presents an interesting analysis of the noise model which, in part, has been taken up for this work. The proposed estimation method, which is called extended least correlation, is similar to an instrumental variables method with delayed regressor as instrument. It is shown that the proposed method outperforms the least squares estimator in case of errors-in-variables models.

There are also works that deal specifically with the actuator alignments, such as [Peck, 2001] and [Fosbury and Nebelecky, 2009]. In [Peck, 2001] the author considers the actuator alignment estimation as dual to attitude determination, considering both problems as the problem of finding the representation of one set of basis vectors relative to another. A least squares method is used to calculate small misalignments of the actuators with respect to their nominal orientation. However, in order to obtain good estimates, the inertia matrix must be very accurately known. For this reason, in [Weiss et al., 2012], the authors design a null-space excitation in order to partially avoid this limitation. In this paper a spacecraft controlled by an overactuated reaction wheel array is considered. The redundancy of the actuators allows for the exploitation of the null

space in order to increase the magnitude of the excitation, while keeping a limited amplitude of the satellite maneuver.

Among all these works, the vast majority considers the availability of angular velocity measurements with sufficient accuracy. Only in [Sekhavat et al., 2009] and [Yoon et al., 2017] a gyroless satellite is considered. However, in many emerging small satellites (such as micro and nano-satellites), for size and cost constraints, only small and imprecise MEMS gyroscopes are used or in some cases no gyroscopes are present at all. These small satellites often must rely only on the star tracker measurements for both attitude and rate estimation. Considering the rapid increase of launches of micro and especially of nano-satellites, the parameter estimation based on the attitude measurements will become increasingly interesting. At the moment this topic can be still considered relatively new, with [Yoon et al., 2017] being the main paper on the subject.

3.3 Parameter Estimation Methods

The parameter estimation problem can be considered as identifying a model among the model set $\mathcal{M}(\theta)$ from data generated by the true system $\mathcal{S}$. However, in realistic scenarios, such the one under consideration, we have

$$\mathcal{S} \notin \mathcal{M}(\theta)\,, \tag{3.11}$$

since the model class (or structure) is typically defined using simplified physics to avoid undue complexity. In fact, for the satellite rotational dynamics, a rigid body model is considered and the internal structural response is neglected. Moreover, the external disturbance torques cannot be accurately modeled. This means that there is no parameter vector θ_0 such that $\mathcal{S} = \mathcal{M}(\theta_0)$.

However, for analysis purposes, the assumption

$$\mathcal{S} \in \mathcal{M}(\theta)\,, \tag{3.12}$$

is considered to be fulfilled, thus it is assumed that a solution to the estimation problem can be found ($\exists \theta_0 : \mathcal{S} = \mathcal{M}(\theta_0)$).

The typical approach to solve this estimation problem is to define a cost function, which is a measure of the fitting between the model and the measured data. The solution is given by the parameter θ which minimizes this cost function.

Let us consider a model in linear regression form, similar to (3.10)[7]

$$g(t_k) = f(t_k)\theta + \nu(t_k)\,, \tag{3.13}$$

where $\nu(t_k)$ represents an additive stationary noise term, and that (3.12) is fulfilled. Given this model (3.13), we need to implement and adapt a method that can accurately estimate the set of spacecraft parameters. Given N measurements taken during the experiment, the overdetermined linear system which is obtained can be solved by the minimization of some cost function.

Hereafter, we briefly present the main possible approaches to the estimation problem, with a more detailed description of the instrumental variable (IV) method since the proposed estimation approaches in this thesis work are based on it.

[7] For simplicity, we neglect the disturbance torque M_d.

3.3.1 The Least Squares Estimation

The least squares solution of a problem like (3.13) is obtained by finding the value of the parameter vector θ which minimizes the sum of the squares of the predition error [Ljung, 1999]:

$$\hat{\theta}_{LS} = \arg\min_{\theta} \sum_{k=1}^{N} \left|\left| g(t_k) - f(t_k)\,\theta \right|\right|^2 . \tag{3.14}$$

The least squares solution, obtained by setting the gradient of the cost to zero, is given by:

$$\hat{\theta}_{LS} = \left(\sum_{k=1}^{N} f^T(t_k) f(t_k) \right)^{-1} \sum_{k=1}^{N} f^T(t_k) g(t_k) \tag{3.15}$$

This solution can be written in a more compact form defining the matrices:

$$f = \begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_N) \end{bmatrix}, \qquad g = \begin{bmatrix} g(t_1) \\ g(t_2) \\ \vdots \\ g(t_N) \end{bmatrix}, \tag{3.16}$$

$$\hat{\theta}_{LS} = \left(f^T f \right)^{-1} f^T g\,. \tag{3.17}$$

For the least squares estimator to be consistent, that is, for $\hat{\theta}_{LS}$ to converge to the true parameter θ_0, the following assumption must be made:

A_1) The matrix $\sum_{k=1}^{N} f^T(t_k) f(t_k)$ is non-singular. This condition depends on the input properties, which is called "persistency of excitation". Anyway, this assumption is not restrictive in the sense that it is not difficult to find an input which satisfies this condition;

A_2) The error term $\nu(t_k)$ is independent of $f(t_k)$, that is $\overline{E}(f^T(t_k)\nu(t_k)) = 0$, where $\overline{E}(\cdot) = \lim_{N\to\infty} \sum_{k=1}^{N} E(\cdot)$, and where $E(\cdot)$ is the expectation operator;

A_3) The model (3.13) is linear in the parameter and time-invariant;

A_4) The error term $\nu(t_k)$ is a zero-mean stationary process.

It can be easily demonstrated that if the noise terms ν belong to a normal distribution, the least squares estimator is also the maximum likelihood estimator [Ljung, 1999].

If assumptions $A1$ to $A4$ are satisfied, it can be shown that this solution (3.15) yields consistent estimates of the true parameter θ_0. Suppose the true model is given by

$$g(t_k) = f(t_k)\theta_0 + \nu(t_k)\,. \tag{3.18}$$

Replacing (3.18) in (3.15) and taking the expectation we obtain:

$$\begin{aligned}\overline{E}[\hat{\theta}_{LS}] &= \lim_{N\to\infty} E\left[\left(\frac{1}{N}\sum_{k=1}^{N} f^T(t_k)f(t_k)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N} f^T(t_k)(f(t_k)\theta_0+\nu(t_k))\right)\right] \\ &= \theta_0 + \lim_{N\to\infty} E\left[\left(\frac{1}{N}\sum_{k=1}^{N} f^T(t_k)f(t_k)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N} f^T(t_k)\nu(t_k)\right)\right] \\ &= \theta_0 + \left(\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} f^T(t_k)f(t_k)\right)^{-1}\overline{E}(f^T\nu)\,.\end{aligned} \tag{3.19}$$

In order to have consistent estimates, the last term of (3.19) must be zero. Even if the noise $\nu(t_k)$ is supposed to be zero mean, a necessary condition is that the regressor $f(t_k)$ and the noise $\nu(t_k)$ are not correlated. In some cases, such as for the satellite model under consideration, this assumption is not true and f and ν are correlated.

The Weighted Least Squares Method

A generalization of the standard least squares method is the weighted least squares. Two cases can be considered: first, different measurements could have different variance. Ordinary least squares is a maximum likelihood estimator when the components of $\nu(t_k)$ in (3.13) are gaussian and independent and identically distributed (i.i.d.). This situation, of constant noise variance, is called homoskedasticity. Often however the magnitude of the noise is not constant and the data are not equally reliable (heteroskedasticity). In case of heteroskedasticity, the ordinary least squares method is no longer the maximum likelihood estimator, and so it is no longer efficient.

The second case, which is more interesting for our application, occurs in MIMO systems, where different components of the output have different variance. In this case, instead of minimizing the residual sum of squares, a weighted sum of squares is minimized:

$$\hat{\theta}_{WLS} = \arg\min_{\theta}\sum_{k=1}^{N}\left|\left|(g(t_k)-f(t_k)\theta)^T W(g(t_k)-f(t_k)\theta)\right|\right|^2, \tag{3.20}$$

where $W = diag(w_i)$, $w_i = 1/\sigma_i$.

The solution of (3.20) is similar to (3.15):

$$\hat{\theta}_{WLS} = \left(\sum_{k=1}^{N} f^T(t_k)W f(t_k)\right)^{-1}\sum_{k=1}^{N} f^T(t_k)W g(t_k)\,. \tag{3.21}$$

3.3.2 The Instrumental Variable Method

The instrumental variable (IV) method has been studied for several years since it is able to overcome some of the limitations of the least squares approach while retaining its simplicity [Young, 2011; Söderström and Stoica, 1983]. Even the least correlation method [Jun and Bernstein, 2006], that can reliably estimate EIV models, presents many similarities with the IV method.

The IV method is a variation of the LS method which has the ability to remain consistent for a class of noise signals which make the LS estimates biased. Both methods are based on a prediction model structure which is linearly parameterized and both have a similar structure of the solution.

Different forms of IV algorithms have been developed, with the main objective of achieving unbiased estimates with the smallest variance, like the refined IV [Young, 2011].

Given a model written as in (3.13) the basic instrumental variable method consists in the minimization of the following cost function

$$\hat{\theta}_{IV} = \arg\min_{\theta} \sum_{k=1}^{N} \left|\left| Z^T(t_k) \Big(g(t_k) - f(t_k)\,\theta \Big) \right|\right|^2 , \tag{3.22}$$

where Z is the "instrument", that is a matrix or a vector of the same size of f, to be chosen, which needs to respect the following properties

$$\begin{cases} \overline{E}(Z^T(t_k) f(t_k)) \text{ is non singular} \\ \overline{E}(Z^T(t_k)\nu(t_k)) = 0\,. \end{cases} \tag{3.23}$$

The analytical solution of (3.22) is closely related to the one of the least squares [Söderström and Stoica, 1983]

$$\hat{\theta}_{IV} = \left[\sum_{k=1}^{N} Z^T(t_k) f(t_k) \right]^{-1} \left[\sum_{k=1}^{N} Z^T(t_k) g(t_k) \right] , \tag{3.24}$$

or in matrix form as

$$\hat{\theta}_{IV} = (Z^T f)^{-1} (Z^T g)\,. \tag{3.25}$$

where

$$f = \begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_N) \end{bmatrix} , \qquad g = \begin{bmatrix} g(t_1) \\ g(t_2) \\ \vdots \\ g(t_N) \end{bmatrix} , \qquad Z = \begin{bmatrix} Z(t_1) \\ Z(t_2) \\ \vdots \\ Z(t_N) \end{bmatrix} . \tag{3.26}$$

The assumptions for the IV estimator to be consistent, which include (3.23), are:

A_1) The matrix $\overline{E}(Z^T(t_k) f(t_k))$ is non-singular. This depends on the input properties, which must be "sufficiently exciting";

A_2) The error term $\nu(t_k)$ is independent of $Z(t_k)$, that is $\overline{E}(Z^T(t_k)\nu(t_k)) = 0$;

A_3) The model (3.13) is linear in the parameter and time-invariant;

A_4) The error term $\nu(t_k)$ is a zero mean stationary process.

The consistency of the IV estimator, again considering the true model (3.18), can be shown by using the two requirements (3.23):

$$\begin{aligned}\overline{E}[\hat{\theta}_{IV}] &= \lim_{N\to\infty} E\left[\left(\frac{1}{N}\sum_{k=1}^{N} Z^T(t_k)f(t_k)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N} Z^T(t_k)(f(t_k)\theta_0+\nu(t_k))\right)\right]\\ &= \theta_0 + \lim_{N\to\infty} E\left[\left(\frac{1}{N}\sum_{k=1}^{N} Z^T(t_k)f(t_k)\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N} Z^T(t_k)\nu(t_k)\right)\right] \qquad (3.27)\\ &= \theta_0 + \left(\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} Z^T(t_k)f(t_k)\right)^{-1}\overline{E}(Z^T\nu) = \theta_0\,.\end{aligned}$$

As for the least squares method, a weighted version of the instrumental variable exists, which can be used to take care of heteroskedastic measures.

One important step of the IV method consists in the construction of a suitable instrument. The properties in (3.23) are a necessary condition in order to obtain consistent estimates but, in order to minimize the variance of the estimates, the instrument must be as correlated as possible with the regressor f. Ideally, the optimal instrument is the noise-free version of the regressor f [Söderström and Stoica, 1983; Young, 2011]. There are two common basic choices for the instrument (e.g. [Söderström and Stoica, 1983]):

- use a delayed regressor as instrument;
- use an estimated noise-free regressor as instrument. Typically, an "auxiliary model" (together with an input signal) is used to generate the noise-free state estimates.

In case a "short" correlation is expected between the noise ν and the regressor f, it is possible that a delay of a few samples makes the second property of (3.23) satisfied[8], while still keeping a high correlation between instrument and regressor. In this particular case the delayed regressor as instrument can be the preferred choice, due to its simplicity. However, in other cases, the use of a delayed regressor as instrument may significantly increase the variance of the IV estimates.

The IV method will be applied to different estimation problems in Chapter 4, 5 and 6. For each of this problems, a differently tailored IV algorithm has been applied (different regressor equations, different filtering and prefiltering, different augmented model and different choice of the instrument).

3.3.3 The Total Least Squares Method

The instrumental variable method is not the only solution to the least squares limitations. One possible alternative method is the total least squares (TLS), which is a natural generalization of the least squares approximation method when the data in both f and g are perturbed [Golub and Van Loan, 1980; Van Huffel and Vandewalle, 1991; Markovsky and Van Huffel, 2007].

To better understand the total least squares method, we have first to consider the ordinary least squares from a different point of view: the LS method corresponds to correct the left-hand side g of (3.13) as little as possible in the Frobenius norm sense, so that the corrected system of equations

$$\hat{g} = f\theta\,, \qquad (3.28)$$

[8] $\overline{E}(f(t_{k+\tau})\nu(t_k)) = 0 \quad \forall \quad \tau >$ minimum delay.

has an exact solution, where $\hat{g} = g + \delta g$. The definition of the total least squares method is motivated by the attempt to remove the asymmetry of the least squares method: g is corrected while f is not. In the case of the EIV problem, it is reasonable to treat f and g symmetrically. The total least squares problem can be defined as:

$$\hat{\theta}_{TLS} = \arg \min_{\theta, \delta f \delta g} \left|\left| [\delta g \quad \delta f] \right|\right|_F \tag{3.29}$$

subject to

$$g + \delta g = (f + \delta f)\theta \tag{3.30}$$

The least squares approximation is statistically motivated as a maximum likelihood estimator in a linear regression model under some assumptions. Similarly, the total least squares approximation is a maximum likelihood estimator in the errors-in-variables model. The classical solution of the TLS problem makes use of the singular value decomposition. Equation (3.30) is equivalent to

$$\begin{bmatrix} f + \delta f & g + \delta g \end{bmatrix} \begin{bmatrix} \theta \\ -1 \end{bmatrix} = 0\,. \tag{3.31}$$

Since the matrix $[f \quad g]$ has rank $n + 1$, the only way to have a non-trivial solution is to select δf and δg in order to reduce the rank to n.

Using the singular value decomposition, the matrix $[f \quad g]$, which has rank $n + 1$, can be expressed as:

$$[f \quad g] = U \Sigma V^T\,, \tag{3.32}$$

where $\Sigma = diag(\sigma_1 \cdots \sigma_{n+1})$. The Eckart-Young-Mirsky theorem states that the rank n matrix $[f + \delta f \quad g + \delta g]$ with the minimum Frobenius norm of the perturbation $||\delta f \quad \delta g||_F$ is given by:

$$U \hat{\Sigma} V^T\,, \tag{3.33}$$

where $\hat{\Sigma} = diag(\sigma_1 \cdots \sigma_n, 0)$[9]. The TLS estimation is therefore the solution of

$$U \hat{\Sigma} V^T \begin{bmatrix} \theta_{TLS} \\ -1 \end{bmatrix} = 0\,. \tag{3.34}$$

The matrix V can be decomposed as:

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}\,, \tag{3.35}$$

where V_{11} has dimension $n \times n$ and V_{12} $n \times 1$. Since the matrix V is orthonormal,

$$V^T \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{3.36}$$

and

$$\hat{\Sigma} V^T \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix} = 0\,. \tag{3.37}$$

[9]The singular values are supposed to be arranged in decreasing order.

Since the last term of (3.34) is -1, the TLS solution is obtained by scaling the vector $[V_{12} \quad V_{22}]^T$ by the factor $-V_{22}$:

$$\hat{\theta}_{TLS} = -V_{12}/V_{22} \tag{3.38}$$

The TLS estimator assumes that all noises in the data have the same variance. Since this hypothesis is often not true, some variants of this method have been developed, like the LS-TLS method, which considers the case when some of the data are unaffected by noise [Van Huffel and Vandewalle, 1991].

Both TLS and IV have very similar performance as shown in [Van Huffel and Vandewalle, 1989; Söderström and Mahata, 2002], therefore, in ideal conditions, it is difficult to suggest one over the other. The main advantage of the IV approach is that its implementation remains simple and consistent without requiring a priori knowledge of the noise statistics, while the TLS should be carefully used in more complex scenarios, such the one considered in this thesis work. In particular, one of the main reasons that made the IV choice prevail over the TLS is the difficult tuning of the TLS algorithm given the different noise standard deviations.

3.3.4 The Kalman Filter as Parameter Estimation Method

The Kalman filter has been initially developed with the purpose to estimate the state of a dynamic system given known parameters and noisy measurements [Haykin, 2001; Brown and Hwang, 1992], and it is widely used for aerospace applications [Crassidis and Junkins, 2011]. However, it may be also used to estimate parameters given a known state and noisy measurements or to simultaneously estimate both the state and parameters of a system [Ljung, 1979; Klein and Morelli, 2006].

In its simplest form, the Kalman filter addresses the general problem of estimating the state x of a discrete-time process that is governed by the linear stochastic difference equation:

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_k + w_{k-1} \\ y_k &= Hx_k + v_k \,, \end{aligned} \tag{3.39}$$

where x_k is the state, u_k is the input, y_k is the output and the random variables w_k and v_k represent respectively the process and measurement noise. They are assumed to be independent of each other, white, and with normal probability distributions:

$$\begin{aligned} w_k &\sim N(0, Q) \\ v_k &\sim N(0, R) \,. \end{aligned} \tag{3.40}$$

The Kalman filter uses a form of feedback control: the filter estimates the process state at some time and then obtains feedback from the noisy measurements and corrects the estimate. The equations of the Kalman filter fall therefore into two groups: time update equations (or predictor equations) and measurement update equations (or corrector equations):

predictor equations:

$$\begin{aligned} \hat{x}_k^- &= A\hat{x}_{k-1} + Bu_k \\ P_k^- &= AP_{k-1}A^T + Q \end{aligned} \tag{3.41}$$

corrector equations:

$$\begin{aligned} K_k &= P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \,. \\ P_k &= (I - K_k H)P_k^- \end{aligned} \tag{3.42}$$

In order to exploit the Kalman filter for parameter estimation, a joint state and parameter identification technique must be used. A common approach to estimate model parameters is to augment the state vector with the unknown parameters and then estimate the augmented state through a Kalman filter.

The Kalman filter can be extended to nonlinear models (like for the satellite case under consideration) by means of linearization (extended Kalman filter), and thus it can be applied directly even to the direct (forward) model.

3.3.5 Optimal Instrumental Variable Estimation

Up to now we have mainly focused on the consistency of the estimator. The instrumental variable method has been chosen for its capacity to remain consistent in a wide range of conditions, where some other methods fail. However, consistency is not the only desirable property of an estimator, also minimal variance of the estimates should be achieved. Optimal IV methods (in the sense of minimum variance of the estimates), that include the use of prefilters, have been long studied, considering open-loop linear systems [Young, 2011; Söderström and Stoica, 1983], closed-loop linear systems [Hof, 1998; Gilson and Hof, 2005; Gilson et al., 2011], as well as continuous-time systems [Young, 1981; Chen et al., 2016; Garnier, 2015].

The Box-Jenkins Model and the Instrumental Variable

To introduce the optimal filtering, a more general form of system model must be defined, which is usually referred to as Box-Jenkins model. Typically, it is often described in discrete-time. However, as commonly done for grey-box modeling, the satellite is modeled in continuous time. For this reason, a continuous-time Box-Jenkins model[10] is considered hereafter (Figure 3.4).

The system to be identified is assumed to be continuous time of finite order:

$$\mathcal{G}(\theta): \quad x(t) = \frac{B(s)}{A(s)} u(t)\,, \tag{3.43}$$

where $x(t)$ is the output, $u(t)$ is the input and $A(s)$ and $B(s)$ are polynomials in s:

$$\begin{aligned} A(s) &= s^{n_a} + a_1 s^{n_a-1} + ... + a_{n_a} \\ B(s) &= b_0 s^{n_b} + b_1 s^{n_a-1} + ... + b_{n_b} \end{aligned} \tag{3.44}$$

The system is parametrized by the coefficients of the polynomials $A(s)$ and $B(s)$:

$$\theta = [a_1, a_2 ... a_{n_a}, b_0, ... b_{n_b}]^T \in \mathbb{R}^{n_a+n_b+1} \tag{3.45}$$

A measuring noise $\nu(t)$ is added to the output $x(t)$. The noise model is defined by:

$$\mathcal{H}(\eta): \quad \nu(t) = \frac{D(s)}{C(s)} e(t)\,, \tag{3.46}$$

where

$$\begin{aligned} C(s) &= s^{n_c} + c_1 s^{n_c-1} + ... + c_{n_c} \\ D(s) &= d_0 s^{n_d} + d_1 s^{n_d-1} + ... + d_{n_d} \end{aligned} \tag{3.47}$$

[10] Also a continuous-time hybrid Box-Jenkins model can be found in the literature, where the noise is modeled in discrete time, while a continuous-time approach is kept for the system dynamics.

and $e(t)$ is a Gaussian zero-mean white noise, i.e. $e(t) \sim N(0, \sigma^2)$. The associated noise model parameters are stacked columnwise in the parameter vector

$$\eta = [c_1, ...c_{n_c}, d_0, ...d_{n_d}]^T \in \mathbb{R}^{n_c+n_d+1}\,. \tag{3.48}$$

The whole model is defined by:

$$\mathcal{M}(\theta\,,\eta): \quad \begin{cases} x(t) & = \frac{B(s)}{A(s)}u(t) \\ \nu(t) & = \frac{D(s)}{C(s)}e(t) \\ y(t) & = x(t) + \nu(t) \end{cases} \tag{3.49}$$

This set of models, parametrized by (θ, η) corresponds to the set of candidate models in which we seek the model that explains the measured data under a given identification criterion.

The following assumption are made:

A_1) The polynomial $A(s)$ has all zeros in the negative half-plane (the system is stable);

A_2) The polynomials $A(s)$ and $B(s)$ are coprime;

A_3) The input $u(t)$ is persistently exciting of order $n_a + n_b$ and is independent of the noise $\nu(t_k)$;

A_4) The polynomials $C(s)$ and $D(s)$ are coprime;

A_5) the true model, $\mathcal{G}_0$, belongs to $\mathcal{G}(\theta)$[11].

The problem of an optimal filtering for the IV method has been studied in several works. In [Söderström and Stoica, 1983], a discrete model is considered, while a hybrid Box-Jenkins model can be found in [Young et al., 2008; Young, 2011]. A full continuous model can be found in older works (for example in [Young, 1981]). In all these works, the optimal IV prefilter for the Box-Jenkins model has the following form

$$F_{opt}(s) = \frac{C(s)}{D(s)\,A(s)}\,, \tag{3.50}$$

and the solution of the optimal[12] IV, for a continuous-time model, is

$$\hat{\theta}_{IV}^{opt} = \left[\sum_{k=1}^{N} Z_f^T(t_k)\phi_f(t_k)\right]^{-1}\left[\sum_{k=1}^{N} Z_f^T(t_k)y_f(t_k)\right], \tag{3.51}$$

with

$$\phi(t_k) = [-y^{(n_a-1)}(t_k)\,,\,\cdots\,,-y(t_k)\,,\,u^{(n_b)}(t_k)\,,\,\cdots\,,u(t_k)]\,, \tag{3.52}$$

where $Z(t_k)$ is a noise-free version of the regressor $\phi(t_k)$, and $y^{(n)}$, $u^{(n)}$ represent the n^{th}-time derivative of y and u.

[11] $\exists\theta_0 : \mathcal{G}_0 = \mathcal{G}(\theta_0)$.

[12] Optimal in the sense of yielding estimates with minimum variance.

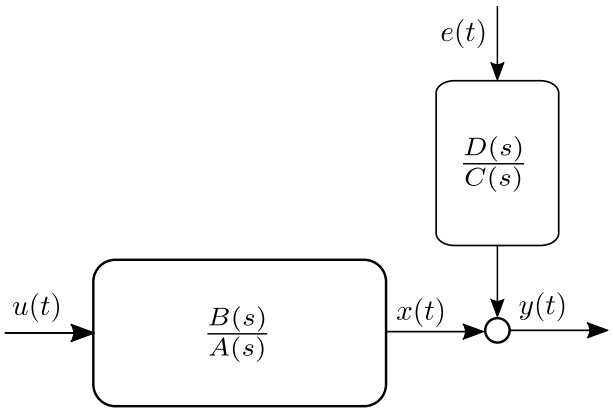


Figure 3.4: Box-Jenkins model.

This important result will be used in the following chapters in order to design the sub-optimal filter (we say "sub-optimal" since, due to the nonlinear terms of the satellite equation, this theory cannot be exactly applied).

In fact, even if the satellite model of Figure 3.2 looks quite different, by moving the effect of the measuring noise to the output, it can be represented in a form similar to the Box-Jenkins model (Figure 3.3). The main difference is due to the nonlinearity from the gyroscopic effect in (3.2). This nonlinearity, given the model inversion, does also affect the overall noise ν. However, as already shown, the satellite system can still be written in linear regression form with respect to the parameters θ, therefore no model approximation is necessary for the IV cost function. Another main difference with respect to Figure 3.4 is the presence of two noise sources on the output, but, as it will be shown in Chapter 5 and 6, an approximation can be found.

The Refined Instrumental Variable Method

Typically, the ARMA (autoregressive moving average) noise model $\frac{D(z^{-1})}{C(z^{-1})}$ (or the equivalent continuous version, CARMA, $\frac{D(s)}{C(s)}$) is not known, therefore in order to design the optimal prefilter, the noise model must be estimated. This is what is done in the refined instrumental variable (RIV) method [Young, 2008; Young, 2011]. The RIV algorithm is iterative and it can be briefly summarized as follows (here it is described considering a hybrid Box-Jenkins model):

1. Estimate the system model parameters ($A(s)$ and $B(s)$) with the simplified refined instrumental variable (SRIV) method (or SRIVC for the continuous-time model). This consists in setting $C(z^{-1}) = D(z^{-1}) = 1$, thus assuming that the output measurements are affected by a white noise. Therefore, the first IV estimation uses as prefilter simply $\frac{1}{A(s)}$;

2. With the previously estimated system parameters, form the estimate

$$\hat{x}(t) = \frac{B(s)}{A(s)} u(t) \tag{3.53}$$

   and calculate the residuals

$$\hat{\xi}_r(t) = y(t) - \frac{B(s)}{A(s)} u(t)\,; \tag{3.54}$$

3. From the residuals $\hat{\xi}_r(t)$ estimate the noise model $\frac{D(z^{-1})}{C(z^{-1})}$ by using an ARMA estimation algorithm;

4. Estimate the system model parameters using as prefilter

$$F = \frac{1}{A(s)} \frac{C(z^{-1})}{D(z^{-1})} . \tag{3.55}$$

Repeat steps 2-3-4 until parameter convergence is reached.

3.4 Next Steps towards the Satellite Parameter Estimation

In Section 3.3 several parameter estimation methods have been presented, and after a first initial analysis, the instrumental variable method has been chosen to be the base for this thesis work. The main reasons that brought to the IV approach choice are the following:

- it can be easily applied to models in linear regression form with respect to the parameters;
- like the least squares (LS) method, it relies on a convex cost function that has an analytical solution;
- it is able to overcome the limitation of the LS regarding the biased estimates due to the noise in the regressor;
- it is robust with respect to noise non-idealities;
- contrary to other methods, the IV estimator is not affected by the use of the inverse model for the linear system case [Ho and Enqvist, 2018].

Moreover, since the system presents only "mild" nonlinearities under the satellite operating conditions, a quasi-optimal filter can be designed to get close to minimum variance parameter estimates. Finally, as shown in Section 3.3, the IV approach can be considered to be relatively new in the aerospace domain making its implementation further interesting.

However, besides the IV choice, several additional and important tailoring operations and data processing steps are necessary to produce accurate results for our estimation problem. These steps and adaptation are included as parts of the specific estimation problems and are described in details in the next chapters:

- Estimate the state derivatives (see Chapter 4, 5 and 6),
- Estimate and remove the possible bias effects due to the nonlinear term of the noise (see Chapter 4, 5 and 6),
- Compensate for the drift of the gyroscope measurements (see Chapter 4),
- Estimate and remove the main deterministic components of the disturbance torque $M_d(t)$ (see Chapter 5 and 6),
- Estimate and compensate the possible delay between the input and output measurements (see Chapter 6),
- Include a prefiltering to reduce the effect of the measurement noise and of the unmodeled terms of the disturbance torque (see Chapter 3 and for more details Chapter 5 and 6),
- Improve the information content of the data by means of optimized satellite maneuvers (Chapter 7).

4

Inertia Matrix Estimation for Gyro-equipped Satellite

Contents

4.1 Satellite Model and Noise Analysis 46
4.2 Filtering for Derivative Estimation 47
4.3 Proposed Instrumental Variable for Inertia Estimation 48
4.4 Simulation Results 51
4.5 Solution for the Inertia Estimation with Biased Gyroscope 54
4.6 Simulation Results with Biased Gyroscope 56
4.7 Conclusions 57

In this chapter we consider the inertia estimation problem for a gyro-equipped satellite. This is the most common scenario considered in the literature, as described in Section 3.2. The main objective consists in showing how the instrumental variable can also be adapted and effectively applied to this estimation problem.

As shown in Section 3.1.1, by considering the model in its inverse form the inertia parameters can be linearly extracted:

$$M(t) + M_d(t) = \Big(\Gamma(\dot{\omega}(t)) + \omega(t)^{\times}\Gamma(\omega(t))\Big)\,\theta_J\,, \tag{4.1}$$

or in the specific case of reaction wheels as actuators

$$-\,\dot{h}_{rw}(t) - \omega(t)^{\times}h_{rw}(t) + M_d(t) = \Big(\Gamma(\dot{\omega}(t)) + \omega(t)^{\times}\Gamma(\omega(t))\Big)\,\theta_J\,. \tag{4.2}$$

with

$$\theta_J = [J_{11}\,,\, J_{22}\,,\, J_{33}\,,\, J_{23}\,,\, J_{13}\,,\, J_{12}]^T. \tag{4.3}$$

This is a major advantage for the estimation process, however, further problems need to be tackled. Firstly, in Section 4.1 we analyze the overall noise term since, given the model inversion, the measurement noise properties may change in this inverse model. Section 4.2, deals with the filter design. The designed filter is used to estimate the unmeasured state derivatives, while it has also a function of "prefilter" for the IV method (Section 3.3.5). Afterwards, in Section 4.3, the proposed IV approach is implemented to the satellite case, where a closed-loop auxiliary

model is used to generate the required intrument. The results, to demonstrate the performance of the proposed method, are presented in Section 4.4. The IV algorithm is tested via Monte Carlo simulations and its results are compared to the more common LS approach. Finally, in Section 4.5, the method is extended in order to deal with a biased gyroscope, and additional simulations further illustrate the algorithm performance (Section 4.6). Part of this work has been published in [Nainer et al., 2018].

4.1 Satellite Model and Noise Analysis

The analysis of the noise term in the inverse model is an important step for the development and implementation of an estimation method, like the instrumental variable. Firstly, the choice of the IV "prefilter" highly depends on the overall noise behavior. Secondly, especially given the nonlinear structure of the system, the model inversion may drastically change the noise properties. In fact, even if the output (angular velocity) is assumed to be affected by a zero-mean noise, the same cannot be necessarily true for the overall noise term in the inverse model (e.g. the zero-mean property may not hold anymore).

In this work, a grey-box study of the noise analysis is performed. This is approach, quite uncommon in the system identification community, is taken since an inverse model is used for the parameter estimation. As shown in Section 3.1.2 (Figure 3.3), the noise model is directly influenced by the satellite inverse model. Since the satellite model structure is known, under some initial assumptions on the sensor noise, the overall noise term (given also the system nonlinearities) can be better analyzed through a grey-box approach.

As a first step, the simplest noise model (2.22) is here considered (a more complete gyroscope model is considered in the next sections)

$$\overline{\omega} = \omega + \tilde{\omega}\,, \tag{4.4}$$

where the three components of $\tilde{\omega}$ are supposed independent on each other and independent on the angular velocity ω. They are also supposed to have the following normal distribution

$$\tilde{\omega}_x \sim N(0, \sigma_x^2)\,, \qquad \tilde{\omega}_y \sim N(0, \sigma_y^2)\,, \qquad \tilde{\omega}_z \sim N(0, \sigma_z^2)\,. \tag{4.5}$$

In this case we consider that the gyroscope measurement errors have the same standard deviation σ on the three axes. The case where the sensor noises have different standard deviations on each axis is analyzed in Chapter 5 and 6.

By substituting (4.4) into the satellite inverse model (4.2), the following equation is obtained

$$\underbrace{-\dot{h}_{rw} - (\overline{\omega} - \tilde{\omega})^{\times} h_{rw}}_{M_a} + M_d = \underbrace{\Big(\Gamma\big(\dot{\overline{\omega}} - \dot{\tilde{\omega}}\big) + (\overline{\omega} - \tilde{\omega})^{\times}\Gamma\big(\overline{\omega} - \tilde{\omega}\big)\Big)}_{\phi}\,\theta_J\,. \tag{4.6}$$

where the dependency on the time (t) has been omitted for sake of simplicity, and where M_a represents the overall torque generated by the actuators (including their gyroscopic effects).

Following the approach of [Jun et al., 2010], the regressor matrix ϕ can be split into three parts

$$\phi = \psi - \delta - \epsilon\,, \tag{4.7}$$

where

$$\psi = \Gamma(\dot{\overline{\omega}}) + \overline{\omega}^{\times}\Gamma(\overline{\omega})\,, \tag{4.8}$$

$$\delta = \Gamma(\dot{\tilde{\omega}}) + \tilde{\omega}^{\times}\Gamma(\tilde{\omega})\,, \tag{4.9}$$

$$\epsilon = \omega^{\times}\Gamma(\tilde{\omega}) + \tilde{\omega}^{\times}\Gamma(\omega)\,. \tag{4.10}$$

Therefore, the right-hand side of (4.6) can be rewritten as

$$\psi\,\theta_J - \delta\,\theta_J - \epsilon\,\theta_J\,. \tag{4.11}$$

The only term that can be computed is ψ, since it is only function of the measurements $\overline{\omega}$, while the other two terms, δ and ϵ are function of the noise free angular velocity ω and of the noise realization $\tilde{\omega}$. We can consider the two terms $\delta\,\theta_J$ and $\epsilon\,\theta_J$ as noise terms.

The full model becomes

$$-\dot{h}_{rw} - \overline{\omega}^{\times}h_{rw} = \psi\,\theta_J\, + \nu - M_d\,, \tag{4.12}$$

where $\nu = -\delta\,\theta_J - \epsilon\,\theta_J - \tilde{\omega}^{\times}h_{rw}$ contains the whole overall noise effects.

In order to apply an estimation algorithm, we first need to analyze the overall noise in the model equations, and thus study the noise property based on the measurement assumptions in (4.4). The term $\tilde{\omega}$ appears in quadratic form only in the second term of (4.9), which can be rewritten as

$$\begin{bmatrix} 0 & -\tilde{\omega}_y\tilde{\omega}_z & \tilde{\omega}_y\tilde{\omega}_z & (\tilde{\omega}_y^2 - \tilde{\omega}_z^2) & \tilde{\omega}_x\tilde{\omega}_y & -\tilde{\omega}_x\tilde{\omega}_z \\ \tilde{\omega}_x\tilde{\omega}_z & 0 & -\tilde{\omega}_x\tilde{\omega}_z & -\tilde{\omega}_x\tilde{\omega}_y & (\tilde{\omega}_z^2 - \tilde{\omega}_x^2) & \tilde{\omega}_y\tilde{\omega}_z \\ -\tilde{\omega}_x\tilde{\omega}_y & \tilde{\omega}_x\tilde{\omega}_y & 0 & \tilde{\omega}_x\tilde{\omega}_z & -\tilde{\omega}_y\tilde{\omega}_z & (\tilde{\omega}_x^2 - \tilde{\omega}_y^2) \end{bmatrix}. \tag{4.13}$$

Even if the noise $\tilde{\omega}$ is zero-mean, the overall noise term $-\delta\theta_J - \epsilon\theta_J$ may have non-zero expected value due to its nonlinear terms (4.13). The nonlinear noise terms (4.13) are overall zero-mean only if $\sigma_x = \sigma_y = \sigma_z$ (considering also the assumption that the three components of the noise are mutually independent). If at least one axis presents a different noise standard deviation, there will be a non-zero term that should be taken into account by the estimator. In all the other terms of the overall error, the measuring noise $\tilde{\omega}$ is simply multiplied by a noiseless signal and therefore it maintains the original zero-mean property.

Considering a zero-mean Gaussian white noise for $\tilde{\omega}$, the overall noise model can be further approximated as

$$\nu_i \simeq D_i(s)\,\tilde{\omega}_i = (d_{0,i} + d_{1,i}s)\,\tilde{\omega}_i \qquad \forall i = x\,,\, y\,,\, z\,, \tag{4.14}$$

where $D_i(s)$ is the noise transfer function that comes from a linearization of the term ν_i with respect to the noise. It has to be noticed that the transfer function is not proper due to the use of the inverse model.

4.2 Filtering for Derivative Estimation

The regressor matrix ψ in (4.12) contains both the measured angular velocity $\overline{\omega}$ and its derivative $\dot{\overline{\omega}}$. Since the gyroscope provides only $\overline{\omega}$, the time derivative should be computed. However, a numerical differentiation significantly amplifies the noise. There are different ways to overcome this drawback, depending on whether the system is linear or nonlinear. In this particular scenario it should be noted that, even if the regressor is nonlinear with respect to the states, the state

derivatives $\dot{\overline{\omega}}$ appear only in linear terms ($\Gamma(\dot{\overline{\omega}})\theta$). For this reason, techniques typically used for linear systems can be used as well.

In this case, a state-variable filter (SVF) is used as a way to avoid numerical differentiation, and moreover, as it is shown further below, it does also achieve the function of the prefilter for the IV method. The SVF is a low-pass filter and its general form is as follows

$$F(s) = \frac{1}{(\gamma\, s + 1)^n}\,, \tag{4.15}$$

where γ is the only user-parameter to be chosen in order to define the SVF bandwidth, while n is the order of the filter. For this particular application, the order is set to $n = 1$ since only the first derivative needs to be estimated. A low-pass filter attenuates the amplification of the noise at the high frequencies, so it contrasts the noise amplification due to numerical differentiations. However, an even more elegant solution consists in rewriting the filtered derivative state ($\dot{\overline{\omega}}_f$) as a function of the original measured angular velocity ($\overline{\omega}$) and its filtered equivalent ($\overline{\omega}_f$). In fact, for $n = 1$, the following holds

$$\dot{\overline{\omega}}_f = \frac{1}{\gamma\, s + 1}\,\dot{\overline{\omega}} = \frac{\overline{\omega}\, s}{\gamma\, s + 1} = \frac{1}{\gamma}\frac{\gamma\overline{\omega}\, s + \overline{\omega} - \overline{\omega}}{\gamma\, s + 1} = \frac{1}{\gamma}\frac{\overline{\omega}(\gamma\, s + 1) - \overline{\omega}}{\gamma\, s + 1} = \frac{\overline{\omega} - \overline{\omega}_f}{\gamma}\,. \tag{4.16}$$

In a similar way it can be demonstrated that for a second order SVF

$$\ddot{\overline{\omega}}_f = \frac{\overline{\omega} - 2\gamma\dot{\overline{\omega}}_f - \overline{\omega}_f}{\gamma^2}\,. \tag{4.17}$$

Since filters generate phase lags and distortions due to different signal amplification depending on the signal frequency, the only way to keep the equality unchanged is to apply the SVF on both sides of the equation (4.12).

$$\begin{aligned}\Big(-\dot{h}_{rw} - \overline{\omega}^{\times} h_{rw}\Big)_f &= \Big(\psi\Big)_f \theta_J + \Big(\nu - M_d\Big)_f \\ &= \left(\Gamma\Big(\frac{\overline{\omega} - \overline{\omega}_f}{\gamma}\Big) + \Big(\overline{\omega}^{\times}\Gamma(\overline{\omega})\Big)_f\right)\theta_J + \Big(\nu - M_d\Big)_f\,,\end{aligned} \tag{4.18}$$

where $(\cdot)_f$ means that the filter is applied. In this equation, the fact that $\dot{\omega}$ appears linearly in Γ has been exploited and the filter has been directly applied to the state ($\Gamma_f = \Gamma(\dot{\omega}_f)$). The same approach is also applied to h_{rw} to compute $\dot{h}_{rw,f}$.

In this case, the SVF does also perform as a prefilter for the IV method. In fact, considering that the inverse satellite model does not have any integrator, and that the noise model is well approximated as (4.14), the optimal IV prefilter design of Section 3.3.5 would yield as prefilter

$$F_i(s) = \frac{1}{D_i(s)} = \frac{1}{d_{1,i}\, s + d_{0,i}} \qquad \forall i = x\,,\, y\,,\, z\,, \tag{4.19}$$

that is exactly a scaled version of (4.15) with $n = 1$.

4.3 Proposed Instrumental Variable for Inertia Estimation

The instrumental variable method was introduced in Chapter 3. However, additional steps need to be tackled in order to adapt it to the inertia estimation problem. Given the model in linear

regression form, and the SVF for the state derivative estimation, the solution of the IV cost function becomes

$$\hat{\theta}_{IV-SVF} = \arg\min_{\theta_J} \sum_{k=1}^{N} \left|\left| Z_f^T(t_k) \Big(M_{a_f}\big(\overline{\omega}(t_k), h_{rw}(t_k)\big) - \psi_f\big(\overline{\omega}(t_k)\big)\, \theta_J \Big) \right|\right|^2 . \tag{4.20}$$

The choice of the instrument Z plays a fundamental role in the IV approach, since from it depends whether the estimation method yields accurate results and whether the variance is similar to its least square counterpart (considering the same prefilter). As already mentioned, to guarantee consistent estimates it is required that $\overline{E}(Z^T\nu) = 0$, while, to achieve estimates with low variance, a strong correlation between Z and ψ is needed.

Typically, for the IV method, the instrument is built from noise-free estimates of the output states (in this case $\hat{\omega}$). To do so, an auxiliary model is built based on the previous knowledge of the system parameters and the output signal is generated based on the input ($-\dot{h}_{rw}$ and its gyroscopic effect for the satellite case). This approach is very common in the literature, but, even if a noise-free input signal is available, it can be only used if the system is stable. Unfortunately, this is not the case for the satellite system given its double integrator behavior. This means that by simulating an auxiliary model in open-loop, its output $\hat{\omega}$ could drift from the true signal ω and, after some time, the correlation between Z and ψ may become significantly low. A possible solution consists in the generation of $\hat{\omega}$ by simulating the auxiliary model in closed-loop and therefore avoiding any possible drift of the estimated angular velocity. This approach presents many similarities with the closed-loop identification problem, even if it is used for different reasons.

In order to run the auxiliary model in closed-loop, additional information are required: firstly, the reference signal given to the controller should be available, secondly the controller model should be either known or separately estimated. In this work the controller is considered to be known, which is quite a realistic assumption. Moreover, due to the nonlinear structure of the controller, its estimation would be non-trivial. Since the auxiliary model uses a value of the parameters, the estimation algorithm includes an iterative process in which at each iteration new estimated parameters are used to update the auxiliary model.

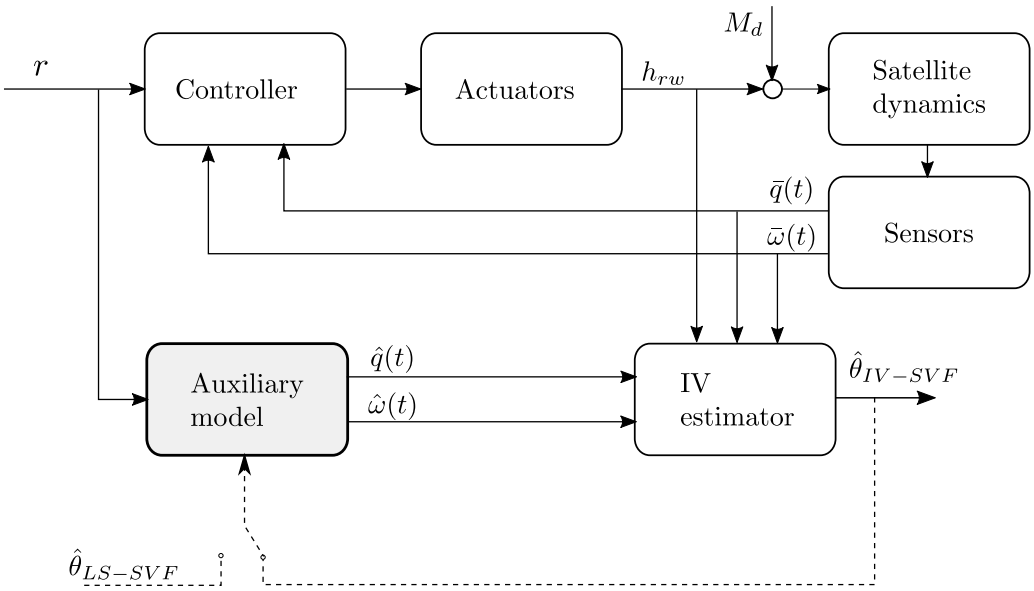


Figure 4.1: IV implementation scheme with the auxiliary model.

The overall estimation algorithm (Figure 4.1) can be synthesized as follows:

1 • Choose the cut-off frequency of the SVF filter (user-parameter γ) ;
2 • Prefilter the model (4.12) with the low-pass SVF, and computed the filtered state derivative as in (4.16) ;
3 • Initialize the inertia parameters, either from *a priori* information (e.g. from CAD software) or from a LS-SVF estimate:

$$\hat{\theta}_{LS-SVF} = \left[\sum_{k=1}^{N} \psi_f^T(t_k)\psi_f(t_k)\right]^{-1}\left[\sum_{k=1}^{N} \psi_f^T(t_k)M_{a_f}(t_k)\right]; \tag{4.21}$$

4 **do**
5 • estimate the noise-free output $\hat{\omega}$ from the reference signal $r(t)$ and the closed-loop auxiliary model based on the previous inertia estimate $\hat{\theta}$;
6 • build the (filtered) instrument Z_f

$$Z_f = \Gamma(\dot{\hat{\omega}}_f) + \left(\hat{\omega}^{\times}\Gamma(\hat{\omega})\right)_f; \tag{4.22}$$

7 • estimate the parameters with the IV method

$$\hat{\theta}_{IV-SVF} = \left[\sum_{k=1}^{N} Z_f^T(t_k)\psi_f(t_k)\right]^{-1}\left[\sum_{k=1}^{N} Z_f^T(t_k)M_{a_f}(t_k)\right], \tag{4.23}$$

8 • update the auxiliary model with the new estimated parameters.
9 **while** $(max\left(|\hat{\theta}(new) - \hat{\theta}(previous)|/\hat{\theta}(previous)\right) < threshold$;

It should be noted that very few iterations are necessary for this sub-optimal IV method. In fact, for the proposed numerical example, convergence was reached in 3-4 iterations.

The only user-parameter in this method is a constant γ that defines the SVF cut-off frequency. This parameter depends mainly on the effect of the gyroscope noise and of the disturbance torques during the satellite maneuver. A deeper analysis on the choice of the low-pass filters is done in Chapter 5 and 6, which considers both the noise sources and the disturbance torque for the filter design.

4.4 Simulation Results

The performance of the proposed method has been tested via numerical simulations. The true inertia value for the satellite was set as

$$J_0 = \begin{bmatrix} 31.3819 & -1.1136 & -0.2601 \\ -1.1136 & 21.1878 & -0.7783 \\ -0.2601 & -0.7783 & 35.7042 \end{bmatrix} [kg\,m^2]\,, \tag{4.24}$$

which are typical values of a microsatellite inertia. The spacecraft is controlled in closed-loop by a PD controller and the reaction wheel actuators present a second order dynamics (with a time constant of $1\,s$). For the FOG gyroscope, a zero-mean Gaussian noise is considered with $\sigma_x = \sigma_y = \sigma_z = 8.5 \cdot 10^{-5}\,rad/s$ as standard deviation. A random-walk term is also taken into consideration to simulate the drift behavior, generated by integrating a zero-mean Gaussian noise with $\sigma_b = 1.3 \cdot 10^{-6} rad/s^2$. The satellite was affected by a disturbance torque, with a varying amplitude oscillating between $\pm 3 \cdot 10^{-5}\,Nm$. The sampling frequency for all the measurements is set to $4\,Hz$, while a parameter $\gamma = 100$ (corresponding to a cut-off frequency of $0.01\,rad/s$) is used for the low-pass prefilter (SVF). The IV method has been initialized by the least squares algorithm, so no initial knowledge is required for the estimation algorithm. The satellite has been excited in all his three axes by a simple maneuver, as shown in Figure 4.2, for a total duration of the simulation of $650\,s$.

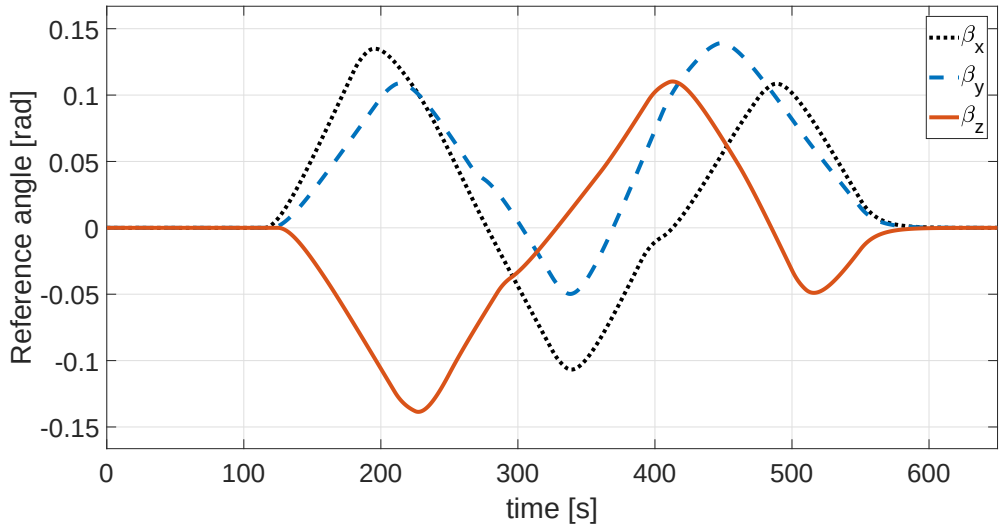


Figure 4.2: Angle reference profile, represented in Euler's angles (sequence z-y-x).

A realization of the input-output data used for the inertia estimation is shown in Figure 4.3

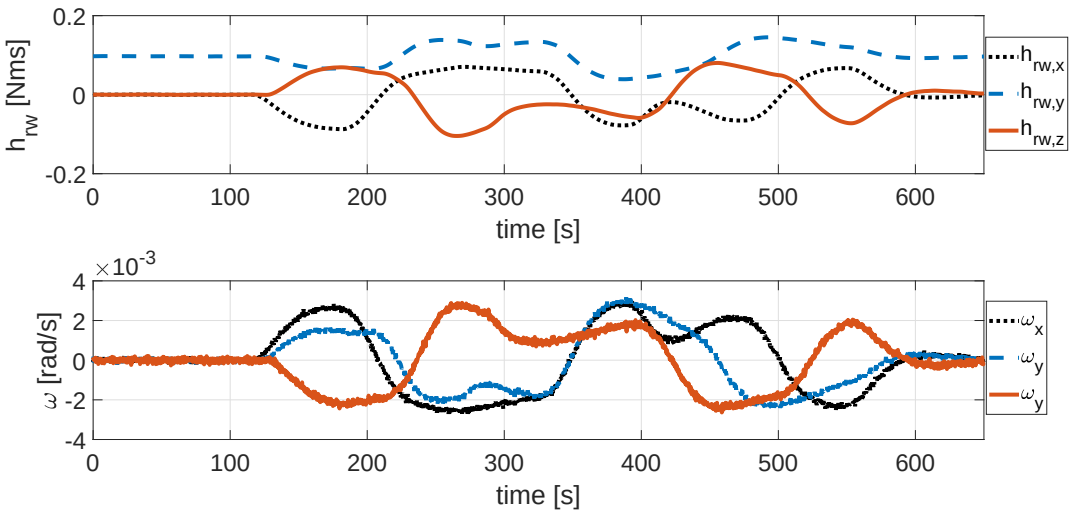


Figure 4.3: Input-Output data: reaction wheel total angular momentum h_{rw} at the top, and satellite measured angular velocity $\overline{\omega}$ at the bottom.

	J_{11} $[kg\,m^2]$		J_{22} $[kg\,m^2]$		J_{33} $[kg\,m^2]$		J_{23} $[kg\,m^2]$		J_{13} $[kg\,m^2]$		J_{12} $[kg\,m^2]$	
	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{\theta}_{LS}$	31.028	0.050	20.880	0.047	35.444	0.058	-0.642	0.036	-0.449	0.043	-0.839	0.042
$\hat{\theta}_{IV}$	31.385	0.051	21.185	0.050	35.705	0.059	-0.780	0.035	-0.254	0.043	-1.114	0.044
θ_0	31.382		21.188		35.704		-0.778		-0.260		-1.114	

Table 4.1: Parameter estimation results for the LS-SVF and IV-SVF methods.

The proposed method results from a Monte Carlo simulation of 100 runs are shown in Table 4.1 where the parameter estimates of the IV approach are compared with the ones from the LS based method. Figure 4.4, showing the inertia errors of the estimates in a box plot[13], illustrates more clearly the different performance between the SVF-LS and SVF-IV methods.

The similar variance between the LS and IV methods demonstrates how, for this application, effective instruments are obtained by using a closed-loop auxiliary model. Additionally, the uncorrelation between the instrument Z and the overall noise ν allows for consistent estimates from the IV method, while the least squares results present, as expected, an evident bias.

To test the sensitivity of the LS and IV methods with respect to the SVF cut-off frequency, 4 different Monte Carlo simulations have been performed, having each a different value for user-parameter γ. The results, for the first inertia parameter, J_{11}, are shown in Figure 4.5. It is evident how the choice of the filter highly affects the LS estimates, and for small values of γ the results are significantly biased (for $\gamma = 1$ the error was above $20\,kg\,m^2$). Instead, the IV method, yields accurate results over a wide range of γ values. Moreover, the variance remain small for a wide range of γ and only for $\gamma = 1$ it begins to significantly increase.

To better visualize the comparison, the mean squared error (MSE) has been computed for

[13]Box plots are often used in this work to show the results. The main box includes the data from the $25th$ to the $75th$ percentile. The circle (or line) inside the box represents the median. The two lines (whiskers) extend from the box to a distance corresponding to $2.7\,\sigma$ (standard deviations). Possible outliers are plotted as individual circles $\circ$ or crosses $+$.

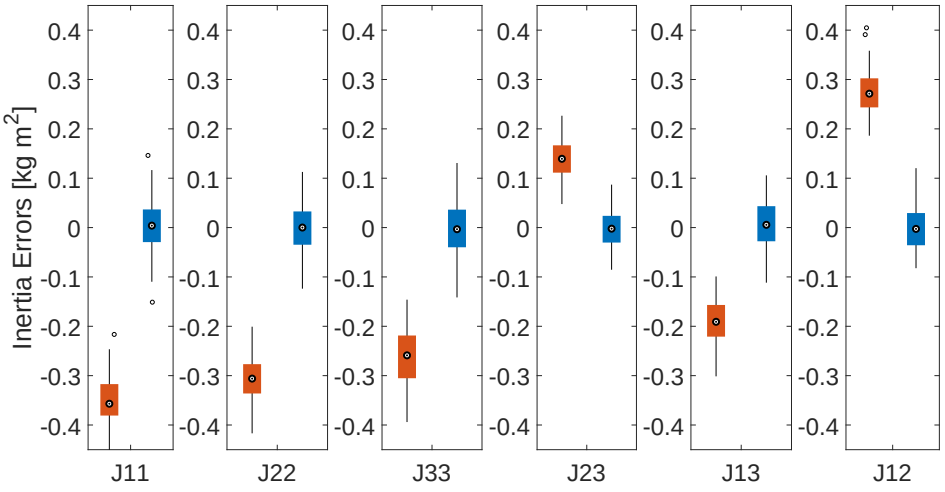


Figure 4.4: Box plots for the errors of the inertia parameter estimates. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF.

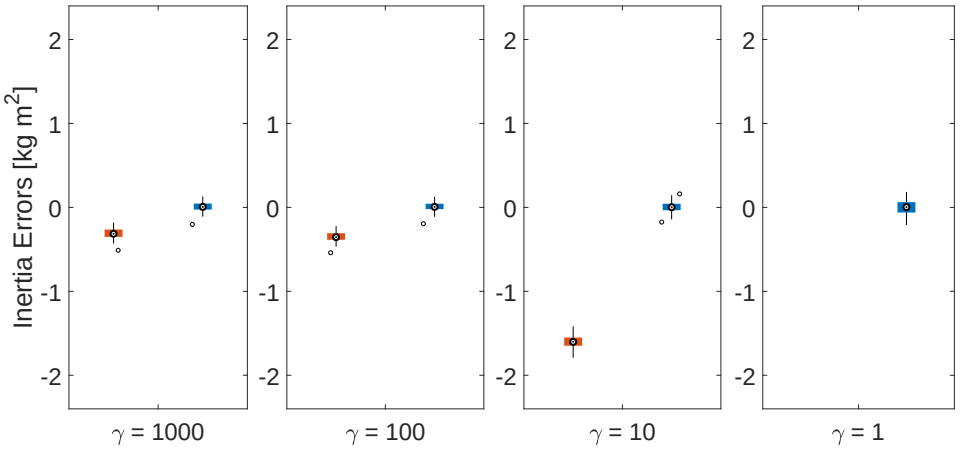


Figure 4.5: Box plots of the error in the $\hat{J}_{11}$ estimate for different values of γ. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF.

both the LS and IV results. The average value among the 6 inertia parameter errors is shown in Figure 4.6.

At a first glance, the results obtained from the least squares (Figure 4.6) seems to be still acceptable (from an application viewpoint) if a good prefilter (SVF) is used. This can be explained by the relatively good sensor measurements, thus resulting in an high signal to noise ratio.

A new Monte Carlo simulation has been performed, where the gyroscope has a standard deviation 4 times higher than what was previously considered: $\sigma_x = \sigma_y = \sigma_z = 34 \cdot 10^{-5}\, rad/s$. The results are shown in Figure 4.7. In this case, the IV results are, as expected, still unbiased, while, on the other hand, the LS performs poorly, with a large bias affecting the estimated parameters.

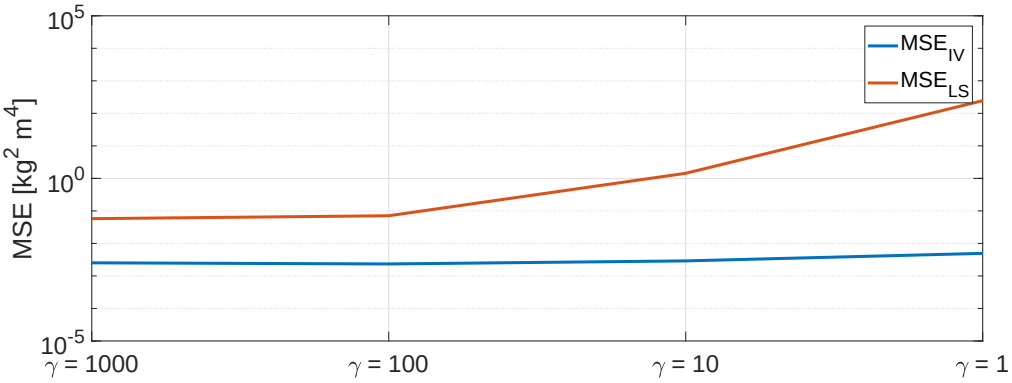


Figure 4.6: Log-plot of the average MSE for the LS-SVF (in red) and IV-SVF (in blue) methods for different values of γ.

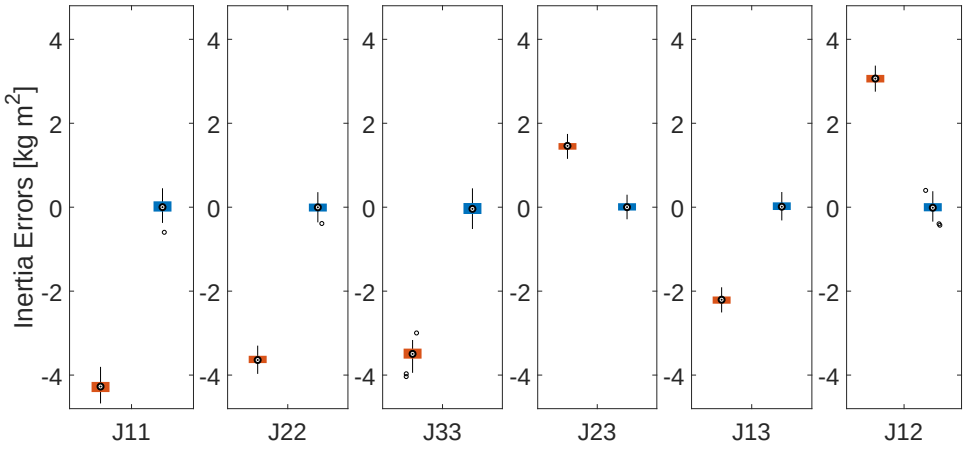


Figure 4.7: Box plots for the errors of the inertia parameter estimates with $\sigma = 34 \cdot 10^{-5}\, rad/s$. In red (on the left) the results for the LS-SVF, while in blue (on the right) the results for the IV-SVF.

These results well justify the choice of the IV method, since its performance is robust with respect of the user-parameter choice and it is able to yield unbiased estimates for the wide variety of test performed. Moreover, the use of the SVF allows for an accurate parameter estimates despite the unmeasured angular accelerations $\dot{\bar{\omega}}$.

Instead, the LS method is not as reliable considering its high sensitivity to the filter user-parameter and that it yields biased estimates. In particular, its use should be avoided for low signal to noise ratio, as shown by the simulation results, since the bias can become unacceptably large and significantly affect the quality of the estimates.

4.5 Solution for the Inertia Estimation with Biased Gyroscope

At first we considered the simple gyroscope model (2.22), (4.4), however, depending on the gyroscope quality, this model could be far from the truth. A more realistic gyro model (2.23) is

due to Farrenkopf:

$$\overline{\omega} = (1 + f_G)\omega + \omega_b + \tilde{\omega}\,, \tag{4.25}$$

where additional terms are present, which may affect the parameter estimation. The first is the gain factor f_G, that scales the overall angular velocity measures. The second is the drift term ω_b. Typically, for accurate gyroscopes $f_G \simeq 0$, therefore the gain error effect is very limited compared to the other noise sources. The drift term, instead, could have some important and detrimental effects on the measures.

For this work we consider the gain f_G to be close to zero and thus negligible, and therefore, the measurement model under consideration is

$$\overline{\omega} = \omega + \omega_b + \tilde{\omega}\,. \tag{4.26}$$

The drift term ω_b is often modeled as a random walk term ($\dot{\omega}_b = \omega_d\,, \quad \omega_d \sim N(0, \sigma_d^2)$), however, considering the relatively short duration of the experiment, its value remains almost constant during the maneuver. It seems therefore acceptable to approximate the drift as an unknown bias:

$$\overline{\omega} = \omega + b + \tilde{\omega}\,, \tag{4.27}$$

where the term b corresponds to the integrated random walk noise up to the experiment time.

A solution to the estimation problem with "biased" gyroscope consists in estimating and removing the effect of the bias generated from the drift term in the gyroscope model.

For linear systems a common technique consists in detrending the data, for instance by removing the average input signal value to the input data, and the output average to the output ([Ljung, 1999], Ch.14). However, this approach cannot be applied in this scenario given the nonlinearity of the satellite system.

As done previously, (4.27) is substituted into the satellite model (4.2)

$$-\dot{h}_{rw} - (\overline{\omega} - b - \tilde{\omega})^{\times} h_{rw} + M_d = \Big(\Gamma\big(\dot{\overline{\omega}} - \dot{b} - \dot{\tilde{\omega}}\big) + (\overline{\omega} - b - \tilde{\omega})^{\times}\Gamma\big(\overline{\omega} - b - \tilde{\omega}\big)\Big)\,\theta_J\,. \tag{4.28}$$

and the equation can be split as in (4.7)-(4.10) since the function $\Gamma(\cdot)$ is linear in the parameter. The term $\dot{b}$ is clearly zero, while $\tilde{\omega}^{\times}\Gamma(\tilde{\omega})$ has been already considered in the previous noise analysis. The constant term $b^{\times}\Gamma(b)$ can be removed by means of an augmented regressor. The crossed terms that contain a noise multiplied by a bias, as $b^{\times}\Gamma(\tilde{\omega})$, can be treated as an additive noise and they do not generate bias. The following terms remain and their effect must be considered:

$$\big(-b^{\times}\Gamma(\omega) - \omega^{\times}\Gamma(b)\big)\,\theta_J - b^{\times} h_{rw}\,. \tag{4.29}$$

Expanding the three terms in (4.29), and collecting the bias components, we obtain

$$-b^{\times}\Gamma(\omega)\,\theta_J = \begin{bmatrix} 0 & -(J_{13}\omega_x + J_{23}\omega_y + J_{33}\omega_z) & (J_{12}\omega_x + J_{22}\omega_y + J_{23}\omega_z) \\ (J_{13}\omega_x + J_{23}\omega_y + J_{33}\omega_z) & 0 & -(J_{11}\omega_x + J_{12}\omega_y + J_{13}\omega_z) \\ -(J_{12}\omega_x + J_{22}\omega_y + J_{23}\omega_z) & (J_{11}\omega_x J_{12}\omega_y + J_{13}\omega_z) & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}, \tag{4.30}$$

$$-\omega^{\times}\Gamma(b)\,\theta_J = \begin{bmatrix} (J_{12}\omega_z - J_{13}\omega_y) & (J_{22}\omega_z - J_{23}\omega_y) & (J_{23}\omega_z - J_{33}\omega_y) \\ (-J_{11}\omega_z + J_{13}\omega_x) & (-J_{12}\omega_z + J_{23}\omega_x) & (-J_{13}\omega_z + J_{33}\omega_x) \\ (J_{11}\omega_y - J_{12}\omega_x) & (J_{12}\omega_y - J_{22}\omega_x) & (J_{13}\omega_y - J_{23}\omega_x) \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}, \quad (4.31)$$

$$-b^{\times}h_{rw} = h_{rw}^{\times}b = \begin{bmatrix} 0 & -h_z & h_y \\ h_z & 0 & -h_x \\ -h_y & h_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}. \quad (4.32)$$

It must be noted that the terms (4.30) and (4.31) contain the unknown parameter J. This problem can be overcome by using the nominal value of the inertia, and the estimates can be further refined by substituting J with the estimated inertia $\hat{J}$.

Defining as B the sum of the matrices that multiply b in (4.30)-(4.32), the full model for the inertia estimation has an augmented regressor ψ_{aug} as follows

$$M_a = \psi_{aug}\theta_{aug} = \begin{bmatrix} \psi & I_3 & B \end{bmatrix} \begin{bmatrix} \theta_J \\ b_0 \\ b \end{bmatrix}. \quad (4.33)$$

thus 6 additional parameters, 3 ($b_0 \in \mathbb{R}^{3\times 1}$) to compensate for the constant effect due to the biased gyroscope measurements ($b^{\times}\Gamma(b)$), and 3 ($b \in \mathbb{R}^{3\times 1}$) to compensate for the varying effect depending on ω (4.29).

As before, the low-pass prefilter (SVF) is applied on both side of the equations, thus the IV solution is

$$\hat{\theta}^{aug}_{IV-SVF} = \left[\sum_{k=1}^{N} Z^T_{aug,f}(t_k)\psi_{aug,f}(t_k)\right]^{-1}\left[\sum_{k=1}^{N} Z^T_{aug,f}(t_k)M_{a_f}(t_k)\right], \quad (4.34)$$

4.6 Simulation Results with Biased Gyroscope

In order to test the effectiveness of the proposed IV method with an augmented regressor, a new Monte Carlo simulation of 100 runs have been performed. The same reference signal used in Section 4.4 is used for this simulation (Figure 4.2), however, a significant initial bias has been added to the drift term ω_b

$$\omega_b(t=0) = [9\,,\,-8\,,\,11]^T \cdot 10^{-4} rad/s\,, \quad (4.35)$$

while the zero-mean gaussian noise and the random-walk term were kept as in the previous simulations ($\sigma_x = \sigma_y = \sigma_z = 8.5 \cdot 10^{-5}\, rad/s$, $\sigma_b = 1.3 \cdot 10^{-6} rad/s^2$).

The first proposed IV method (4.23) has been compared with its augmented version (4.34), and for sake of clarity they have been renamed as IV_1 and IV_2, respectively. The results, from a Monte Carlo simulation of 100 runs, are shown in Figure 4.8 where the estimation error for each parameter is plotted, as well as in Table 4.2. For the "standard" IV method (in red) the results present a significant bias, while the augmented version (in blue) still yields consistent estimates despite the large amount of bias in the angular velocity measurements. Additionally, both methods present similar variances, thus the additional estimated parameters do not worsen the precision of the new proposed algorithm.

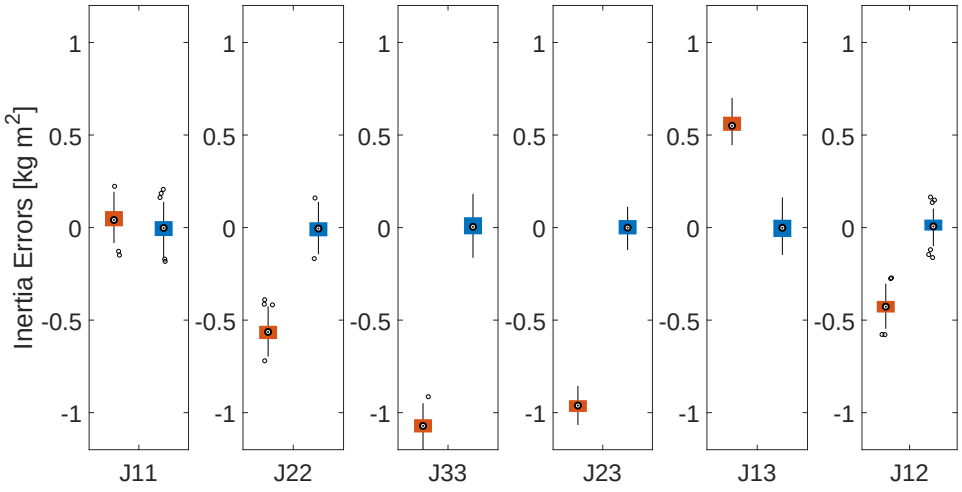


Figure 4.8: Box plots for the errors of the inertia parameter estimates. In red (on the left) the results for the IV-SVF (IV_1), while in blue (on the right) the results for the augmented IV-SVF (IV_2).

	J_{11} $[kg\,m^2]$		J_{22} $[kg\,m^2]$		J_{33} $[kg\,m^2]$		J_{23} $[kg\,m^2]$		J_{13} $[kg\,m^2]$		J_{12} $[kg\,m^2]$	
	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{\theta}_{IV_1}$	31.428	0.066	20.623	0.060	34.631	0.062	-1.739	0.046	0.298	0.057	-1.541	0.054
$\hat{\theta}_{IV_2}$	31.380	0.072	21.181	0.060	35.713	0.076	-0.780	0.052	-0.261	0.064	-1.104	0.054
θ_0	31.382		21.188		35.704		-0.778		-0.260		-1.114	

Table 4.2: Parameter estimation results for the IV-SVF method without ($\hat{\theta}_{IV_1}$) and with ($\hat{\theta}_{IV_2}$) gyroscope bias compensation.

4.7 Conclusions

In this chapter an instrumental variable method has been developed and applied for the inertia matrix estimation from telemetry data. Due to the lack of angular acceleration measurements, a SVF prefilter has been applied in order to avoid performing the derivatives. The proposed IV based algorithm exploits a closed-loop auxiliary model to generate the output noise-free estimates required by the instrument. Since the auxiliary model operates in closed loop and it uses the same reference signal as the real system, drifts between real and estimated output is avoided.

A comparison with a least squares estimation has been made. Due to the similarities with a typical errors-in-variables model, it has been shown how the least squares approach cannot guarantee unbiased and consistent estimates, while the IV counterpart is able to overcome the LS limitations and obtain accurate inertia estimates. Moreover, the results show also how the IV is significantly more robust with respect to the choice of the SVF cut-off frequency. It should also be noted that, given the same prefilter, the LS method has theoretically lower variance than the IV method. The almost identical standard deviations in the estimates for the two approaches demonstrates therefore the effectiveness of the instrument generated from the closed-loop auxiliary model.

Finally, the case of biased gyroscope has been considered, and it has been shown how, with an

appropriate modification of the IV method, consistent parameter estimates can still be obtained. Even if this approach has a higher number of parameter to be estimated, the variance increase is relatively small compared to the results of Section 4.4 making its use interesting if a high gyroscope measurement bias is expected.

There are, however, satellites that do not use gyroscopes, but rely mainly on the star tracker sensor for both attitude and angular velocity estimation. For these gyroless satellites, the method proposed in this chapter cannot be applied. This is the main focus of the following chapter, where the IV approach is extended in order to work directly on attitude measurements. In this new scenario, the prefilter design is more critical given that also a second order time-derivative is required. Moreover, the effects of the disturbance torques are better analyzed and compensated in the inertia estimation algorithm.

5

Inertia Matrix Estimation for Gyroless Satellite

Contents

5.1 Satellite Closed-Loop System 60
5.2 Satellite Model and Noise Analysis 61
5.2.1 Satellite Model 61
5.2.2 Noise Analysis 61
5.3 Derivative Estimates 64
5.4 Prefilter Design for Instrumental Variable 65
5.5 Augmented Model and Algorithm Implementation 67
5.6 Numerical Results from High-Fidelity Simulator 69
5.7 Inertia Estimation Using Real Data 73
5.8 Conclusions 74

In Chapter 4 it has been shown how to obtain consistent estimates of the inertia matrix by using the instrumental variable method. For the algorithm, it was assumed that angular velocity measurements were available, as done by the majority of works in the literature. However, small satellites (Myriade, CubeSat, ...) are emerging due to their lower costs. These types of satellites have less sensor availability compared to bigger spacecraft, making the attitude determination as well as parameter estimation more challenging. Usually in larger satellites, both star tracker and accurate gyroscopes (e.g. fiber-optic gyroscopes (FOG)) are used for attitude and rate estimation. However, in smaller satellites, for size or cost constraints, small MEMS gyroscopes are used (as well as the star tracker for attitude measurements). These gyroscopes present much worse accuracy than FOGs, due to much higher random walk noise and to a higher sensitivity to the environment conditions. These types of satellites typically rely only on star tracker measurements for both attitude and rate estimation. Studies on attitude and rate estimation in gyroless conditions have been made, however, as shown in Section 3.2, very few works focused on the satellite parameter estimation from star tracker measurements only.

In this chapter the instrumental variable method has been tailored to work on gyroless satellites, therefore relying solely on attitude measurements. Additionally, a detailed noise analysis is performed that allows for an effective choice of the prefilter.

For the numerical experiment, all the data that have been processed and presented in this chapter have been generated by a high-fidelity simulator from CNES, where all possible sources

of disturbance are present. The simulation conditions are therefore as close as possible to the real environment. Finally, the method is tested with real data recorded during a set of maneuvers on the Picard satellite. This work, with the exception of the Picard's results, has been published in [Nainer et al., 2019].

5.1 Satellite Closed-Loop System

For this study we consider the Microcarb satellite. Microcarb is a micro-satellite that will be launched in 2021 to map sources and sinks of carbon dioxide. Since this spacecraft is still in development, only Microcarb simulated data have been used. The real telemetry data are obtained from the Picard satellite, an older but similar micro-satellite.

Given the open-loop satellite dynamics unstability (it contains two integrators from torque to attitude), all the experiments must be carried out under a closed-loop control configuration. The overall system is shown in Figure 5.1. Since the spacecraft is gyroless, and to improve the signal to noise ratio, the feedback includes an extended Kalman filter (EKF) that estimates the satellite angular rate (required by the controller) from the attitude measurements (e.g. [Crassidis et al., 2007]). Since the extended Kalman filter is based on an approximate model and because we aim to have a fully independent estimation method, we prefer not to use the angular rate obtained this way. The satellite controller structure [Genin and Viaud, 2018] is as follows

$$K(s) = (K_p + K_d \cdot s)\,\frac{1}{s}\,\frac{1}{1+\tau_1 \cdot s}\,\frac{1+\tau_2 \cdot s}{1+\tau_3 \cdot s}\,, \tag{5.1}$$

and it includes the following feedforward control law

$$K_{ff} = G_{ff} \cdot J \cdot \dot{\omega}_{target}(t + \Delta T)\,, \tag{5.2}$$

where $\dot{\omega}_{target}$ is the target acceleration, G_{ff} and ΔT are two tuning parameters.

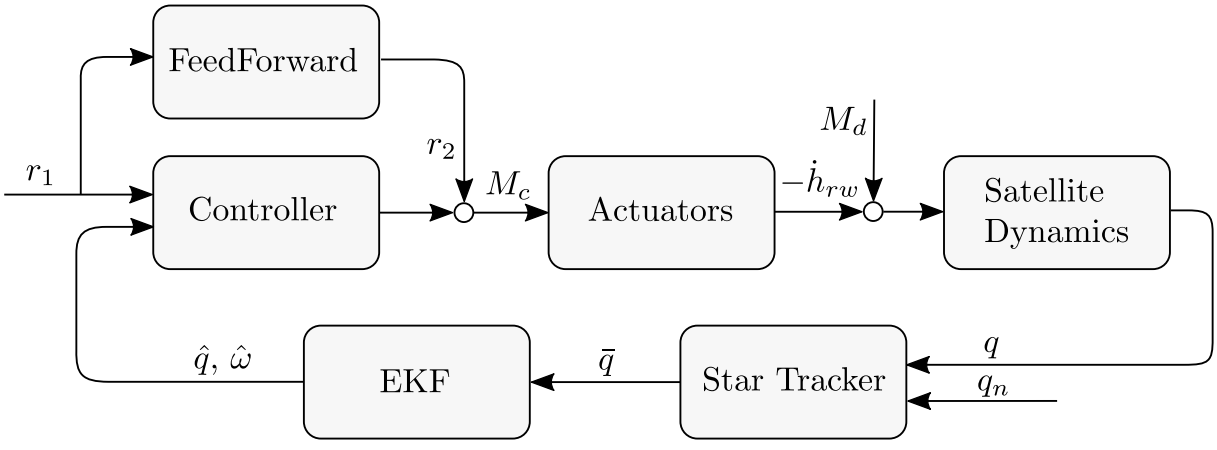


Figure 5.1: Closed-loop satellite system. The equivalent torque generated by the reaction wheels is represented by $-\dot{h}_{rw}$, while r_1 and r_2 are the reference and feedforward signals, respectively.

5.2 Satellite Model and Noise Analysis

5.2.1 Satellite Model

The satellite model is kept in its inverse form like in the gyro-equipped case. As shown in (3.2) in Chapter 3 and later used in Chapter 4, we can write

$$-\dot{h}_{rw}(t) - \omega(t)^{\times} h_{rw} + M_d(t) = J\,\dot{\omega}(t) + \omega(t)^{\times} J\,\omega(t)\,, \tag{5.3}$$

and

$$-\dot{h}_{rw}(t) - \omega(t)^{\times} h_{rw} + M_d(t) = \Big(\Gamma(\dot{\omega}(t)) + \omega(t)^{\times}\Gamma(\omega(t))\Big)\,\theta_J\,, \tag{5.4}$$

for the equivalent system in linear regression form. However, under the new assumptions, we do not have access to the angular velocity measurements $\overline{\omega}$. Therefore, the angular velocity must be computed from the attitude measurements provided by the star tracker.

As shown in Chapter 2, the attitude expressed in quaternions can be directly related to the angular velocity and its derivative (in the body reference frame)

$$\omega = 2W(q)\dot{q}\,, \qquad \dot{\omega} = 2W(q)\ddot{q}\,, \tag{5.5}$$

where $W(\cdot)$ is defined as follows

$$W(q) = \begin{bmatrix} q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix}. \tag{5.6}$$

5.2.2 Noise Analysis

Since the attitude is also affected by measurement noise, the noise effect on the reconstructed angular velocity should be analyzed. We first consider the case where the star tracker frame is aligned with the body reference frame, thus we have the following measurement equation

$$\overline{q}(t_k) = q_n(t_k) \otimes q(t_k)\,, \tag{5.7}$$

where, for small errors, as in Chapter 2

$$q_n(t_k) \simeq [1\,,\, e_x/2\,,\, e_y/2\,,\, e_z/2]^T\,, \tag{5.8}$$

with

$$e_x \sim N(0,\, \sigma_x^2)\,, \qquad e_y \sim N(0,\, \sigma_y^2)\,, \qquad e_z \sim N(0,\, \sigma_z^2)\,, \tag{5.9}$$

where e_x, e_y, e_z are assumed to be mutually uncorrelated. By expanding (5.7) and omitting the time t_k for sake of simplicity, we obtain

$$\overline{q} = \begin{bmatrix} q_0 - \frac{e_x q_1}{2} - \frac{e_y q_2}{2} - \frac{e_z q_3}{2} \\ q_1 + \frac{e_x q_0}{2} - \frac{e_y q_3}{2} + \frac{e_z q_2}{2} \\ q_2 + \frac{e_x q_3}{2} + \frac{e_y q_0}{2} - \frac{e_z q_1}{2} \\ q_3 - \frac{e_x q_2}{2} + \frac{e_y q_1}{2} + \frac{e_z q_0}{2} \end{bmatrix}, \tag{5.10}$$

therefore the measured quaternion can be rewritten in additive form as

$$\overline{q} = q + \tilde{q}\,, \tag{5.11}$$

where

$$\tilde{q} = \begin{bmatrix} -\frac{e_x q_1}{2} - \frac{e_y q_2}{2} - \frac{e_z q_3}{2} \\ \frac{e_x q_0}{2} - \frac{e_y q_3}{2} + \frac{e_z q_2}{2} \\ \frac{e_x q_3}{2} + \frac{e_y q_0}{2} - \frac{e_z q_1}{2} \\ -\frac{e_x q_2}{2} + \frac{e_y q_1}{2} + \frac{e_z q_0}{2} . \end{bmatrix} \tag{5.12}$$

From (5.5) the angular velocity $\overline{\omega}$ can be computed from the measured quaternions (and their derivatives) as

$$\overline{\omega} = 2W(\overline{q})\dot{\overline{q}} = 2W(q)\dot{q} + 2W(\tilde{q})\dot{q} + 2W(q)\dot{\tilde{q}} + 2W(\tilde{q})\dot{\tilde{q}}\,, \tag{5.13}$$

with

$$\dot{\tilde{q}} = \frac{1}{2}\begin{bmatrix} -\dot{e}_x q_1 - e_x\dot{q}_1 - \dot{e}_y q_2 - e_y\dot{q}_2 - \dot{e}_z q_3 - e_z\dot{q}_3 \\ \dot{e}_x q_0 + e_x\dot{q}_0 - \dot{e}_y q_3 - e_y\dot{q}_3 + \dot{e}_z q_2 + e_z\dot{q}_2 \\ \dot{e}_x q_3 + e_x\dot{q}_3 + \dot{e}_y q_0 + e_y\dot{q}_0 - \dot{e}_z q_1 - e_z\dot{q}_1 \\ -\dot{e}_x q_2 - e_x\dot{q}_2 + \dot{e}_y q_1 + e_y\dot{q}_1 + \dot{e}_z q_0 + e_z\dot{q}_0 \end{bmatrix} . \tag{5.14}$$

and where $\dot{e}_i$ represents the numerical differentiation of e_i, and therefore having as variance $\sigma^2_{\dot{e}_i} = 2\sigma^2_{e_i}/(\Delta T)^2$ (or $\sigma^2_{\dot{e}_i} = 2\sigma^2_{e_i}/(2\Delta T)^2$ in case of the central differentiation) where ΔT is the sampling period.

Equation (5.14) can be approximated as

$$\dot{\tilde{q}} \simeq \frac{1}{2}\begin{bmatrix} -\dot{e}_x q_1 - \dot{e}_y q_2 - \dot{e}_z q_3 \\ \dot{e}_x q_0 - \dot{e}_y q_3 + \dot{e}_z q_2 \\ \dot{e}_x q_3 + \dot{e}_y q_0 - \dot{e}_z q_1 \\ -\dot{e}_x q_2 + \dot{e}_y q_1 + \dot{e}_z q_0 \end{bmatrix} , \tag{5.15}$$

assuming

$$\|\dot{e}\|\,\|q\| \gg \|e\|\,\|\dot{q}\|\,, \tag{5.16}$$

which is a very realistic condition in our satellite scenario: the quaternion is an unit vector, while its derivative is usually small for satellite maneuvers; moreover the noise on the quaternions is smaller than its derivative.

Since $2W(q)\dot{q} = \omega$, the noise on $\overline{\omega}$ is as follows

$$\tilde{\omega} = \overline{\omega} - \omega = 2W(\tilde{q})\dot{q} + 2W(q)\dot{\tilde{q}} + 2W(\tilde{q})\dot{\tilde{q}}\,. \tag{5.17}$$

From assumption (5.16), and considering that $\|q\| \gg \|\tilde{q}\|$, the overall noise on the angular velocity can be approximated as

$$\tilde{\omega} \simeq 2W(q)\dot{\tilde{q}} = \begin{bmatrix} \dot{e}_x(q_0^2 + q_1^2 + q_2^2 + q_3^2) \\ \dot{e}_y(q_0^2 + q_1^2 + q_2^2 + q_3^2) \\ \dot{e}_z(q_0^2 + q_1^2 + q_2^2 + q_3^2) \end{bmatrix} = \begin{bmatrix} \dot{e}_x \\ \dot{e}_y \\ \dot{e}_z \end{bmatrix} , \tag{5.18}$$

since $q(t)$ is a unit quaternion.

From this result, the overall noise on the angular rate is directly related to the noise on the quaternions, and if the noise e is zero-mean and with components mutually uncorrelated, the same can be said for $\tilde{\omega}$.

In the more general case where the star tracker is not aligned with the body reference frame, an additional rotation must be performed. Therefore the angular rate ω in the satellite frame can be written as

$$\omega = R\,\omega_{st}\,, \tag{5.19}$$

where ω_{st} is the angular rate in the star tracker reference frame and R is the rotation matrix transforming from the star tracker to the body reference frame. If the three components of the noise e have the same standard deviation σ_e then, even after the rotation, the noise term $\tilde{\omega}$ remains zero-mean with its three components mutually uncorrelated.

This can be demonstrated as follows:

Suppose that ω is a zero mean vector such that $E(\omega_x^2) = E(\omega_y^2) = E(\omega_z^2) = \sigma^2$ and with the three components uncorrelated.

- The zero mean property is clearly maintained after a rotation since $E(R\,\omega) = R\,E(\omega) = 0$.
- To demonstrate that the three components of ω remain uncorrelated, we first apply a rotation about the x-axis with an angle α to the vector ω:

$$\omega' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha) & sin(\alpha) \\ 0 & -sin(\alpha) & cos(\alpha) \end{bmatrix} \omega = \begin{bmatrix} \omega_x \\ \omega_y cos(\alpha) + \omega_z sin(\alpha) \\ -\omega_y sin(\alpha) + \omega_z cos(\alpha) \end{bmatrix} . \tag{5.20}$$

  The first component of ω' is clearly independent of the other two, therefore $E(\omega'_x\,\omega'_y) = E(\omega'_x\,\omega'_z) = 0$. Also the second and third component are independent since:

$$E(\omega'_y\,\omega'_z) = E(-\omega_y^2 sin(\alpha)cos(\alpha) + \omega_z^2 sin(\alpha)cos(\alpha) + \omega_y\omega_z cos^2(\alpha) + \omega_y\omega_z sin^2(\alpha)) = 0\,, \tag{5.21}$$

  since ω_y and ω_z are uncorrelated and they have the same variance.

  The components of the vector ω' have also the same variance of the original vector ω. For the first component this statement is trivial. For the other two we have:

$$\begin{aligned} E(\omega'_y\,\omega'_y) &= E(\omega_y^2 cos^2(\alpha) + \omega_z^2 sin^2(\alpha) + 2\omega_y\omega_z sin(\alpha)cos(\alpha)) = \sigma^2 \\ E(\omega'_z\,\omega'_z) &= E(\omega_y^2 sin^2(\alpha) + \omega_z^2 cos^2(\alpha) - 2\omega_y\omega_z sin(\alpha)cos(\alpha)) = \sigma^2\,. \end{aligned} \tag{5.22}$$

  Applying a second rotation about the y-axis and a third rotation about the z-axis the situation does not change: the three components remain uncorrelated and their variance does not change.

However, if the three noise components have different variance, then there will be a correlation between $\tilde{\omega}_x$, $\tilde{\omega}_y$ and $\tilde{\omega}_z$.

The overall effect of the noise in the satellite model is then simply computed by substituting $\overline{\omega} = \omega + \tilde{\omega}$ where $\tilde{\omega} = \dot{e}$ as in (5.18). Once the substitution is performed, similarly to Chapter 4, the following is obtained

$$\underbrace{-\dot{h}_{rw} - (\overline{\omega} - \tilde{\omega})^{\times} h_{rw}}_{M_a} + M_d = \underbrace{\left(\Gamma\left(\dot{\overline{\omega}} - \dot{\tilde{\omega}}\right) + (\overline{\omega} - \tilde{\omega})^{\times}\Gamma\left(\overline{\omega} - \tilde{\omega}\right)\right)}_{\phi}\,\theta_J\,. \tag{5.23}$$

By following the same steps as in (4.7)-(4.10), (5.23) can be rewritten as

$$-\,\dot{h}_{rw} - \overline{\omega}^{\times} h_{rw} = \psi(\dot{\overline{\omega}}, \overline{\omega})\,\theta_J + \nu - M_d\,, \tag{5.24}$$

where the overall noise is contained in ν:

$$\nu = -\delta\theta_J - \epsilon\theta_J - \tilde{\omega}^{\times} h_{rw}\,. \tag{5.25}$$

with

$$\begin{aligned}\psi &= \Gamma(\dot{\bar{\omega}}) + \bar{\omega}^{\times}\Gamma(\bar{\omega})\,,\\ \delta &= \Gamma(\dot{\tilde{\omega}}) + \tilde{\omega}^{\times}\Gamma(\tilde{\omega})\,,\\ \epsilon &= \omega^{\times}\Gamma(\tilde{\omega}) + \tilde{\omega}^{\times}\Gamma(\omega)\,.\end{aligned} \tag{5.26}$$

As before, noise term $\tilde{\omega}$ appears in a quadratic form only in the second term of δ as

$$\begin{bmatrix} 0 & -\tilde{\omega}_y\tilde{\omega}_z & \tilde{\omega}_y\tilde{\omega}_z & (\tilde{\omega}_y^2-\tilde{\omega}_z^2) & \tilde{\omega}_x\tilde{\omega}_y & -\tilde{\omega}_x\tilde{\omega}_z \\ \tilde{\omega}_x\tilde{\omega}_z & 0 & -\tilde{\omega}_x\tilde{\omega}_z & -\tilde{\omega}_x\tilde{\omega}_y & (\tilde{\omega}_z^2-\tilde{\omega}_x^2) & \tilde{\omega}_y\tilde{\omega}_z \\ -\tilde{\omega}_1\tilde{\omega}_2 & \tilde{\omega}_1\tilde{\omega}_2 & 0 & \tilde{\omega}_x\tilde{\omega}_z & -\tilde{\omega}_y\tilde{\omega}_z & (\tilde{\omega}_x^2-\tilde{\omega}_y^2) \end{bmatrix}. \tag{5.27}$$

As in the previous chapter, we can conclude that a necessary (and sufficient) condition to have a zero-mean overall noise effect is that the three components $\tilde{\omega}_1$, $\tilde{\omega}_2$, $\tilde{\omega}_3$ must be mutually uncorrelated and have the same variance. Since star trackers have usually different variance on different axes, the assumption of equal variance is no more satisfied.

Therefore, for the gyroless case, the non-zero expected value of the overall noise should be taken into consideration to correctly estimate the inertia parameters.

From now on, we will consider the regressor $\psi^T(\cdot)$ directly as function of the quaternions and their derivatives

$$\psi\Big(\bar{q}(t),\dot{\bar{q}}(t),\ddot{\bar{q}}(t)\Big) = \Gamma\Big(2W(\bar{q}(t))\ddot{\bar{q}}(t)\Big) + \Big(2W(\bar{q}(t))\dot{\bar{q}}(t)\Big)^{\times}\Gamma\Big(2W(\bar{q}(t))\dot{\bar{q}}(t)\Big), \tag{5.28}$$

and therefore the overall model in linear regression form becomes

$$-\dot{h}_{rw}(t) - \Big(2W(\bar{q}(t))\dot{\bar{q}}(t)\Big)^{\times} h_{rw}(t) = \psi\Big(\bar{q}(t),\dot{\bar{q}}(t),\ddot{\bar{q}}(t)\Big)\,\theta_J + \nu(t) - M_d(t)\,. \tag{5.29}$$

5.3 Derivative Estimates

The model (5.29) is function of the attitude quaternion $q(t)$, as well as of its time derivatives $\dot{q}(t)$, $\ddot{q}(t)$. However the star tracker provides only the angular position $\bar{q}$. The simplest solution would be to perform a numerical differentiation, having as main drawback a significant noise amplification. A way to limit the noise amplification consists in filtering the signal before performing the differentiation. The SVF approach of Chapter 4 cannot be used in this case since all the state derivatives appear nonlinearly in the model. Indeed, while in (4.1) the derivative term appears linearly in $\Gamma(\dot{\omega})$, the same is not true for the derivative terms in (5.28) and (5.29).

Therefore, other solutions must be sought. In order to avoid distortion of the signal, the filter must not generate any phase lag, its magnitude should be as flat as possible in the passband, and it should have a cut-off frequency above the bandwidth of the satellite dynamics. Since we consider an offline estimation method, a smoothing operation based on a Butterworth low-pass filter is chosen, since it has a maximally flat magnitude in the pass-band. This smoother consists in applying the filter in both forward and backward directions in order to cancel any phase lag inherent to the filtering. Since the filtering could slightly change the mangnitude of the quaternions, a normalization is performed. Afterwards, the derivatives are computed through central finite differences. If instead an online estimation is required, then the Butterworth smoother could be replaced by a delayed state-variable filter (DSVF) [Tsang and Billings, 1994]. The DSVF is a filter that has a frequency response presenting constant unit gain and an almost linear phase with respect to the frequency in the passband of the filter. Usually, the filter is based

on a Butterworth filter since it presents a maximally flat amplitude characteristic and an almost linear phase characteristic for frequencies below the cut-off. Then, all-pass equalizers are used to further improve the linear phase characteristic in the pass band of the Butterworth filter. A more detailed description of the DSVF is given in Appendix B.

Other approaches are available in the literature, such as the integrative random walk smoother (IRWSM) [Young, 2011], the Wavelet decomposition-based smoothing [Coca and Billings, 1997], and polynomial smoothing like the Savitzky-Golay filter [Savitzky and Golay, 1964; Schafer, 2011]. However, their performance is not necessarily superior to the Butterworth low-pass smoothing, or for the IRWSM case, its implementation is not as trivial.

It should be noted that this procedure is mainly used to remove the high-frequency noise components without altering the noise-free signal. The main part of the data processing is performed afterwards by the so called "prefilter"[14], and it allows to obtain low-variance parameter estimates.

5.4 Prefilter Design for Instrumental Variable

The implementation of a basic IV method, thus solving (3.22), can only guarantee consistency of the estimates, whereas the variance could still be large making the parameter estimates not reliable.

The optimal instrumental variable method, described in Chapter 3, exploits a suitable filter in order to minimize the variance of the estimates. To apply this method, the various components of the Box-Jenkins model must be recognized in the system equation and, first of all, the noise model must be identified.

Equation (5.29) corresponds to the model of Figure 3.3 where the effect of the noise on the attitude has been brought to the output side and added to the disturbance torque. This model has many similarities with the Box-Jenkins model of Figure 3.4, and the idea to exploit the optimal IV filter (3.50) developed for this model would be very attractive. Unfortunately, apart from the nonlinearity of the model, in (5.29) two independent noise sources are present on the output. As far as the authors are aware, multiple noise sources have never been deeply studied and the prefilter design is not obvious [Mahata and Garnier, 2006]. Given the two independent noise sources, the overall noise cannot be obtained as the output of a single equivalent filter with a white noise input, therefore an approximation is sought: a linear shaping filter which approximates the power spectrum of the overall wideband noise is estimated and its inverse is used as a prefilter. Even if there is no theoretical demonstration of optimality, this approximation is justified by the intent to develop a filter which whitens the output noise. The intuitive reasoning behind this filter, is that, if the residual is highly colored, it means that there is still information in the data that has not been captured, therefore better estimate can still be achieved.

The satellite model (3.8) is nonlinear, therefore the theory does not provide any optimal filter for the IV method. Anyway, since the operating conditions of the satellite do not vary significantly, by performing a linearization of the model, we can design a sub-optimal filter that will be still able to achieve good performance [Brunot et al., 2018; Nainer et al., 2019].

As already mentioned, the model of Figure 3.4, presents several similarities with (5.29) at exception from the double noise source and the presence of some nonlinearities. Even before performing any linearization, we can observe that the term $\psi(\cdot)$ does not have any integrator inside, but only derivatives. Therefore for the filter construction, considering that for linear

[14]The prefilter is not applied directly to the measured states, but it is applied to regressor ψ and the input $\dot{h}_{rw} - \overline{\omega}^{\times} h_{rw}$.

systems the optimal IV filter is $\frac{C(s)}{A(s)D(s)}$ (Section 3.3.5), in this particular case, we can set $A(s) = 1$, and only the terms $\nu(t)$ and $M_d(t)$ affect the filter design.

From Section 5.2.2, the overall noise term is (5.25)

$$\nu = -\delta\theta_J - \epsilon\theta_J - \tilde{\omega}^{\times}\overline{h}_{rw}\,. \tag{5.30}$$

By linearizing the term ν around the operating conditions a linear approximation of ν is obtained (of the type $\frac{D(s)}{C(s)}$, where in this case $C(s) = 1$ given the absence of integrators in the noise model). Since the noise term bandwidth is not affected by its sign, the linearization point has been chosen as the average of the absolute values of h_{rw} and ω during the satellite maneuver

$$h^*_{rw} = \frac{1}{N}\sum_{k=1}^{N}\left|h_{rw}(t_k)\right|, \qquad \omega^* = \frac{1}{N}\sum_{k=1}^{N}\left|\overline{\omega}(t_k)\right|. \tag{5.31}$$

In order to linearize ν, we first compute the linearization with respect to $\tilde{\omega}$ and $\dot{\tilde{\omega}}$, and then, exploiting (5.18), we obtain a linearization with respect to e and its derivatives:

$$\nu_l = (\nabla_{\dot{\tilde{\omega}}}\nu)\,\dot{\tilde{\omega}} + (\nabla_{\tilde{\omega}}\nu)\,\tilde{\omega} = -J\dot{\tilde{\omega}} - (\omega^{\times}J - (J\omega)^{\times} - h^{\times}_{rw})\,\tilde{\omega} \tag{5.32}$$

This linearization has the following form[15]

$$\nu_l = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\tilde{\omega} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}\dot{\tilde{\omega}} \tag{5.33}$$

Since the noise $\tilde{\omega}$ can be accurately approximated by $\dot{e}$ (as shown in (5.18)), (5.33) can be rewritten as

$$\nu_l = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\dot{e} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}\ddot{e}\,. \tag{5.34}$$

To have three separate scalar noise models (and thus a three scalar filters), the matrices in (5.34) are diagonalized by calculating the diagonal terms as a mean square of the row values

$$a_i = \sqrt{a_{i1}^2 + a_{i2}^2 + a_{i2}^2}\,, \qquad b_i = \sqrt{b_{i1}^2 + b_{i2}^2 + b_{i2}^2} \qquad \text{for} \quad i = 1\cdots 3\,, \tag{5.35}$$

$$\nu_l \simeq \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix}\dot{e} + \begin{bmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix}\ddot{e}\,. \tag{5.36}$$

The noise model transfer function can be therefore written as

$$D_l(s) = D_1\,s + D_2\,s^2\,. \tag{5.37}$$

It should be noted that in this particular noise model there are no terms without differential operator s.

In case of a disturbance free satellite, the inverse of the noise model ($\frac{1}{D_l(s)}$), weighted by the different standard deviations of the three e components, would be the sub-optimal filter.

[15] we temporary use a_{ij} to describe the linearized noise model. It should not be confused with the actuator alignment a in Chapter 6.

For the disturbance, we consider a white noise filtered with very low-pass cut off, thus having a behavior similar to a bounded random walk signal. This model represents the residual term after having estimated and removed the constant components. Also the gravity gradient torque can be removed, since it can be accurately computed from the satellite attitude with respect to the Earth [Curtis, 2013]. Therefore, the following transfer function is used to describe the disturbance model

$$M_{d,s} = \frac{\gamma}{s+\gamma}\eta \qquad \text{with} \quad \eta_x \sim \eta_y \sim \eta_z \sim N(0, \sigma_\eta^2)\,, \tag{5.38}$$

where γ is an empirical value (e.g. $\gamma = 0.002\,rad/s$) to describe the low-pass behavior of the disturbance torque. Alternatively, the disturbance torque could be modeled by a simpler integral random walk noise $\dot{M}_d = \eta$, like in [Yoon et al., 2017].

Since these two noise terms (ν and M_d) are uncorrelated, the two transfer functions cannot be simply added. The correct way to combine these noise terms is to add the power spectra with their respective gains. However, in this particular case the two spectra are substantially disjoint: the noise effect produces a high pass spectrum, while the disturbance is strongly low-pass. Therefore, since there is no overlapping, it is permissible to add the two transfer functions, with their respective gains, instead of combining the two power spectra.

The noise term including the disturbance can be then approximated as

$$\nu_i = (b_i\, s^2 + a_i\, s)e_i + \frac{\gamma}{s+\gamma}\eta_i \qquad \forall\, i = x, y, z\,. \tag{5.39}$$

In order to "merge" the terms, the ratio between σ_e and σ_η must be computed. The value of σ_e can be easily obtained from the sensor datasheet, while for σ_η, a rough estimate can be obtained from the knowledge of the typical disturbance torque magnitude. The overall noise model can be then written as[16]

$$H(s) = b\, s^2 + a\, s + \frac{\sigma_\eta \gamma/\sigma_e}{s+\gamma} = \frac{d_3 s^3 + d_2 s^2 + d_1 s + d_0}{s+\gamma}\,, \tag{5.40}$$

and the prefilter $F(s)$ used in the IV is the reciprocal of the full noise transfer function $H(s)$ (5.40)

$$F(s) = (H(s))^{-1} = \frac{s+\gamma}{d_3 s^3 + d_2 s^2 + d_1 s + d_0} \tag{5.41}$$

An example of the $F(s)$ magnitude, for the satellite scenario under consideration, is shown in Figure 5.4.

This approach can be linked to the refined IV (Section 3.3.5). In fact, the grey-box analysis of the noise (and disturbance) model substitutes the ARMA noise model estimation in the more common RIV approach. Moreover, as in (3.55), the prefilter is designed as the inverse of the noise model, where, in this particular scenario, $A(s) = 1$. A more direct comparison between these two approaches is shown in Chapter 6.

5.5 Augmented Model and Algorithm Implementation

In Section 5.2 it has been shown how, depending on the star tracker orientation and noise standard deviation, the noise term ν may have a non-zero expected value. Moreover, the disturbance

[16] For ease of notation we consider here just one of the three components of the noise transfer function.

torque contains some slowly varying terms, which must be estimated and removed. One suggested solution to deal with these terms consists in augmenting the regressor (and therefore the instrument and the parameter vector as well)

$$\begin{array}{l} \psi_{aug}(t_k) = \begin{bmatrix} \psi(t_k) & I_3 \end{bmatrix} \\ Z_{aug}(t_k) = \begin{bmatrix} Z(t_k) & I_3 \end{bmatrix} \end{array} \qquad \forall k \in [1 \cdots N]\,, \tag{5.42}$$

where the identity matrix I_3 allows the estimation of the overall constant terms due to the measurement noise and the disturbance $M_d(t)$.

In the satellite scenario under consideration, the prefilter (Figure 5.4) has a very slow dynamics, therefore the initial conditions of the filter have a significant impact on the inertia estimates. Waiting for the end of the transient behavior may result in discarding a significant portion of the data. A better solution consists in estimating the filter initial conditions together with the satellite parameters. To do so, the (filtered) regressor and instrument are further augmented as

$$\begin{array}{l} \psi_{aug,f}(t_k) = \begin{bmatrix} \psi_f(t_k) & I_3 & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k) \end{bmatrix}\,, \\ Z_{aug,f}(t_k) = \begin{bmatrix} Z_f(t_k) & I_3 & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k) \end{bmatrix}\,, \end{array} \tag{5.43}$$

where $T_{ri}(t_k)$ are the transient behaviors corresponding to each one of the filter poles (in this case we considered a third order filter) which may have one of the following forms[17]

$$T_{ri}(t_k) = \begin{cases} I_3\, e^{\alpha t_k}\,; \\ I_3\, e^{\alpha t_k} sin(\omega_1 t_k)\,; \\ I_3\, e^{\alpha t_k} cos(\omega_1 t_k)\,; \\ I_3\, t_k e^{\alpha t_k}\,. \end{cases} \tag{5.44}$$

These four forms correspond to: real pole, couple of complex poles, and double real pole, respectively.

The fully extended parameter vector θ_{aug} contains now 12 parameters in addition to θ_J.

As in Chapter 4 (Section 4.3), we build the instrument from a closed-loop auxiliary model. The feedforward and the PD components of the controller are supposed to be known, and thus only the inertia parameters are required by the auxiliary model. Since the inertia has not yet been estimated, the auxiliary model is initialized with a preliminary estimate. Then, when a new estimate becomes available, the model is updated and a few iteration are made until convergence.

The IV based algorithm for the inertia estimation from star tracker can be synthesized as

[17] For sake of simplicity the diagonal terms of the matrices are considered the same. In a more general form, each component could have a different transient behavior.

follows:

- apply the Butterworth smoothing filter to the quaternion measurements $\overline{q}(t_k)$ and compute the time derivatives $\dot{\overline{q}}(t_k)$, $\ddot{\overline{q}}(t_k)$;
- initialize the inertia parameters, either from *a priori* information (e.g. from CAD software) or from a low-pass filtered LS (LS-F) estimate[a]

$$\hat{\theta}_{LS-F} = \left[\sum_{k=1}^{N} \psi_f^T(t_k)\psi_f(t_k)\right]^{-1}\left[\sum_{k=1}^{N} \psi_f^T(t_k)M_{a_f}(t_k)\right], \tag{5.45}$$

do

- estimate the noise-free output $\hat{\omega}$ from the reference signal $r(t)$ and the closed-loop auxiliary model based on the previous inertia estimate $\hat{\theta}$ (it should be noted that the instrument can be directly built from ω, so the noise-free estimate of q is not required);
- build the prefilter $F(s)$ based on the noise analysis that also depends on the parameter θ_J;
- build and filter by $F(s)$ the instrument instrument Z_f

$$Z_f = \Gamma(\dot{\hat{\omega}}_f) + \left(\hat{\omega}^{\times}\Gamma(\hat{\omega})\right)_f ; \tag{5.46}$$

- augment the filtered regressor $\psi_f(\cdot)$ and the instrument $Z_f(\cdot)$ as in (5.43):

$$\begin{aligned} \psi_{aug,f}(t_k) &= \begin{bmatrix}\psi_f(t_k) & \mathbf{I_3} & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k)\end{bmatrix}, \\ Z_{aug,f}(t_k) &= \begin{bmatrix}Z_f(t_k) & \mathbf{I_3} & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k)\end{bmatrix}, \end{aligned} \tag{5.47}$$

- estimate the parameters with the IV method

$$\hat{\theta}_{IV-F} = \left[\sum_{k=1}^{N} Z_{aug,f}^T(t_k)\psi_{aug,f}(t_k)\right]^{-1}\left[\sum_{k=1}^{N} Z_{aug,f}^T(t_k)M_{a_f}(t_k)\right], \tag{5.48}$$

- update the auxiliary model with the new estimated parameters.

while $(max\big(|\hat{\theta}(new) - \hat{\theta}(previous)|/\hat{\theta}(previous)\big) < threshold$;

[a] This is a LS method that uses a low-pass filter, similar to $F(s)$, as prefilter.

5.6 Numerical Results from High-Fidelity Simulator

The proposed estimation method has been tested through numerical simulations. The MicroCarb gyroless satellite, from the Myriade class, has been considered for this experiment. In order to generate the data, a high fidelity simulator from CNES has been used. The main parameters of the simulators are shown hereafter. The nominal satellite inertia matrix is

$$J_0 = \begin{bmatrix} 20.3852 & -3.7497 & -1.7515 \\ -3.7497 & 24.5764 & 0.7836 \\ -1.7515 & 0.7836 & 29.0328 \end{bmatrix} [kg\,m^2]. \tag{5.49}$$

The satellite is controlled by four reaction wheels. Since we assumed that the actuator alignments are known as well as their inertia, and that the noise on the reaction wheel speed readings is

negligible, we have direct access to the reaction wheel angular momentum h_{rw}. All the sensor sampling frequencies are set to $4\,Hz$. The noise standard deviation on the star tracker was set as $\sigma_x = \sigma_y = 11.7 \cdot 10^{-6}\,rad$ and $\sigma_z = 93 \cdot 10^{-6}\,rad$. The star tracker model had also bias and harmonic (at the orbital frequency) noise terms: the bias was set to $b_x = b_y = 58 \cdot 10^{-6} rad\,, b_z = 53 \cdot 10^{-6}\,rad$, while the harmonic amplitude was $\alpha_x = \alpha_y = 8 \cdot 10^{-6} rad\,, \alpha_z = 23 \cdot 10^{-6} rad$ (both in the star tracker frame).

The transformation matrix to move from the star tracker frame to the satellite frame was the following

$$Q_{sst}^{sat} = \begin{bmatrix} cos(\beta) & 0 & sin(\beta) \\ -sin(\beta)\,sin(\alpha) & -cos(\alpha) & sin(\alpha)\,cos(\beta) \\ sin(\beta)cos(\alpha) & -sin(\alpha) & -cos(\alpha)\,cos(\beta) \end{bmatrix}, \qquad \text{with } \alpha = 15^\circ \text{ and } \beta = 0^\circ\,. \quad (5.50)$$

All the main types of disturbance torques are considered[18]: aerodynamic torque, solar radiation pressure, magnetic torque, gravity gradient torque, as well as internal oscillation in the structure (mainly due to flexible elements like solar panels). Additionally, the satellite is equipped with a scanning mirror, that is generating additional disturbances. During the simulations, the solar panel are kept fixed in order to avoid any change of the satellite inertia during the estimation. The control law is described in (5.1) and (5.2). The reference guidance profile is shown in Figure 5.2, and it is kept the same for every simulation.

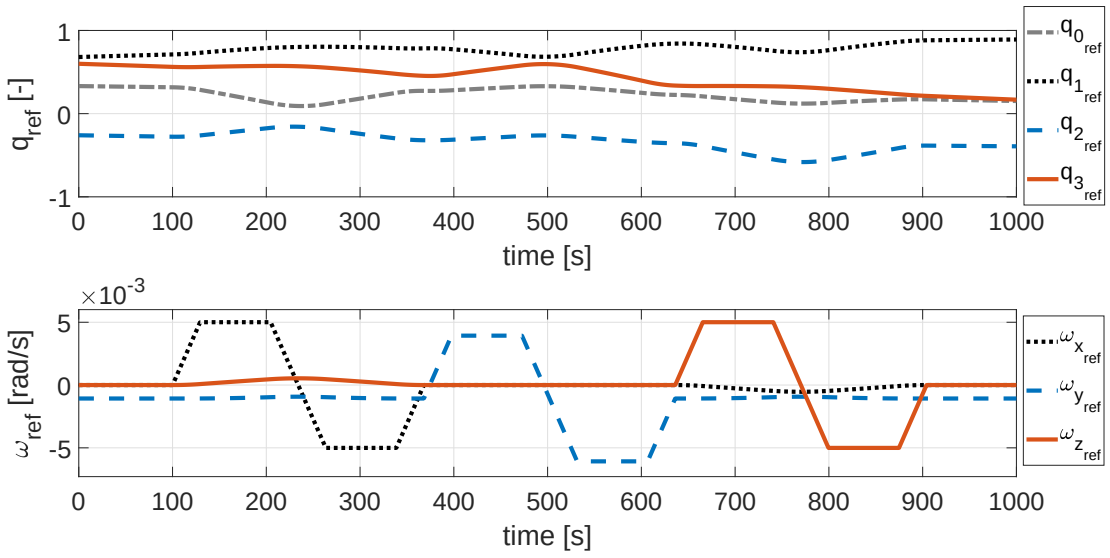


Figure 5.2: Attitude and angular rate reference profile.

Figure 5.3 shows the input-output of the satellite for a single simulation.
Two types of Monte Carlo simulations are performed:

- in the first type of simulation the star tracker follows the simple model (5.7);

[18]The satellite is operating in low Earth orbit (LEO).

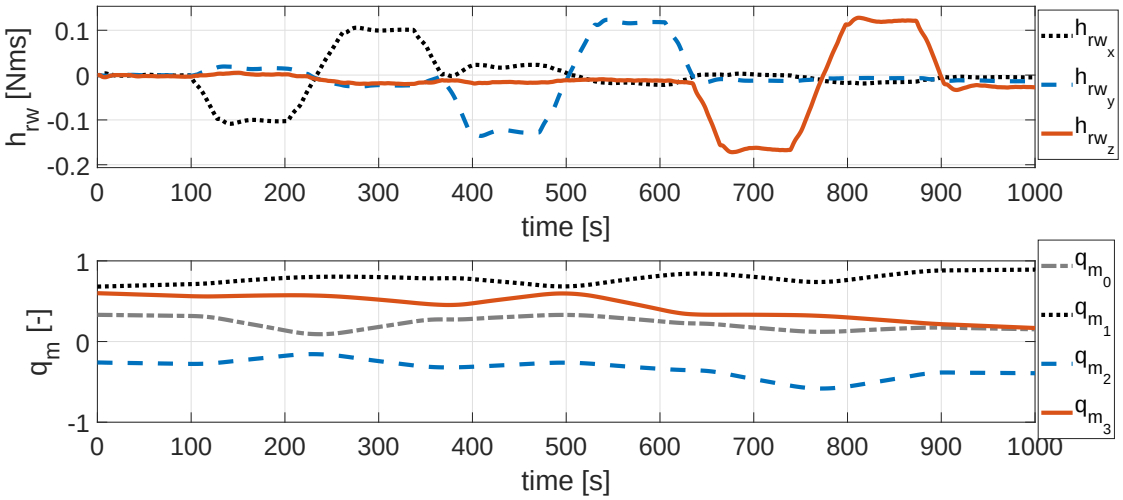


Figure 5.3: Input (reaction wheel total angular momentum) and output (in quaternions) of the satellite.

- the second considers a more realistic case, where also bias and harmonic noise terms are present.

In order to have a different disturbance torque profile, the initial satellite orbit position has been randomly changed in every simulation. The prefilter has been designed following Section 5.4 and its typical magnitude is shown in Figure 5.4.

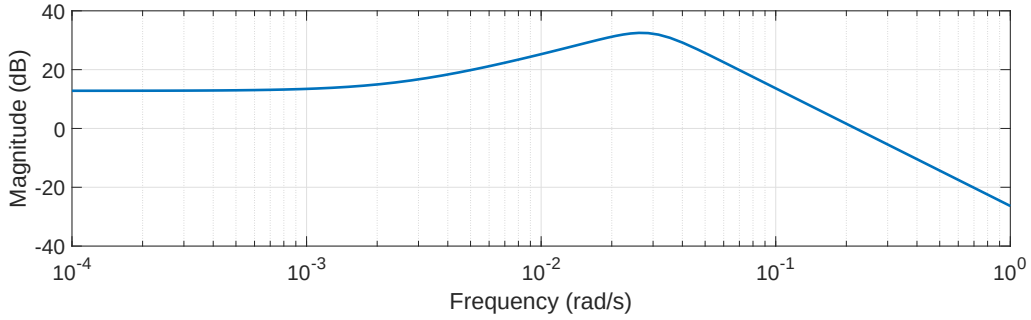


Figure 5.4: Magnitude Bode plot of the filter used in the IV method.

Table 5.1 shows the results from a Monte Carlo simulation of 100 runs. The IV estimates for both star tracker models present a low variance and a slight bias, however, the presence of bias in the quaternions does not seem to have a significant impact in the performance. The slight bias is mainly due to the disturbance torque, while, even in the biased star tracker case, the correlation between the noise term e and the noise free quaternions q is too small to generate any significant effect (as shown in the comparison of Table 5.1). The table includes also the results from a least squares algorithm with a second-order low pass filter, used as comparison, which shows a considerably higher variance and bias. The bias of the least squares results, from an application viewpoint, can be considered still low. This is mainly due to two factors: the use

of a well-tuned prefilter and the use of an accurate star tracker (low noise standard deviation). In fact, as it has been shown in Chapter 4, either a not well-tuned prefilter, or simply a lower signal to noise ratio can easily affect the quality of the LS estimates.

	J_{11} $[kg\,m^2]$		J_{22} $[kg\,m^2]$		J_{33} $[kg\,m^2]$		J_{23} $[kg\,m^2]$		J_{13} $[kg\,m^2]$		J_{12} $[kg\,m^2]$	
	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{\theta}_{LS}$	20.431	0.034	24.687	0.033	29.021	0.036	0.766	0.033	-1.742	0.031	-3.783	0.041
$\hat{\theta}_{IV_1}$	20.397	0.006	24.625	0.008	29.054	0.008	0.776	0.011	-1.742	0.009	-3.743	0.005
$\hat{\theta}_{IV_2}$	20.397	0.006	24.625	0.008	29.055	0.008	0.776	0.010	-1.741	0.009	-3.743	0.005
θ_0	20.385		24.576		29.033		0.784		-1.751		-3.768	

Table 5.1: Results from a Monte Carlo simulation of 100 runs using different satellite starting orbit positions. The true value is represented by θ_0. θ_{IV_1} is the inertia estimate for the unbiased star tracker, whereas θ_{IV_2} is for the biased case. In this simulation settings, the performance difference is negligible. Additionally, θ_{LS} shows the results from a least squares method with a second order low-pass filter.

The main user-parameter to be set is the ratio between σ_η and σ_e since its value determine the frequency for the intersection between the noise and disturbance transfer functions, thus it determines the cut-off of the filter. Figure 5.5 shows the mean square errors of the inertia estimates for different ratios, in order to demonstrate the algorithm robustness.

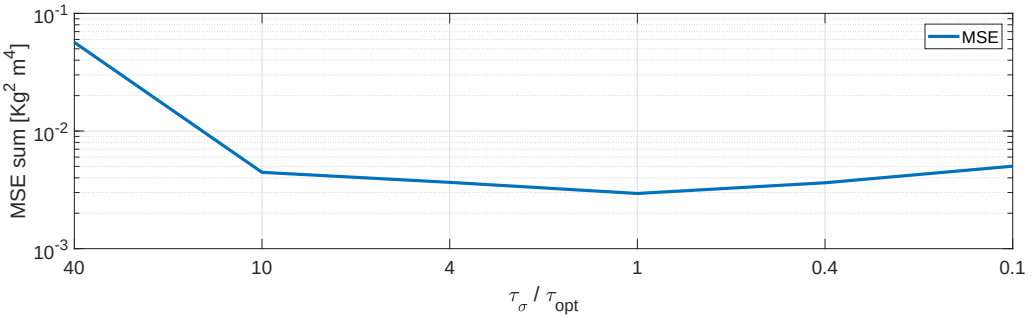


Figure 5.5: Sum of the mean square errors (MSE) of the 6 inertia parameter estimates for different values of the user-parameter τ_σ. The value of τ_σ is represented as its ratio with an empirical optimal value τ_{opt}.

5.7 Inertia Estimation Using Real Data

Picard is a CNES microsatellite from the Myriade series, that was launched in 2010 and remained operational till 2014. Before its end of life, an adaptive attitude control law has been tested [Pittet et al., 2015] and from that experiment real telemetry data have been recorded that can be used to further test the proposed parameter estimation algorithm.

The maneuver in the experiment was as follows

- at $t = 3600\,s$, 15° guidance step on the X axis,
- at $t = 14400\,s$, 15° guidance step on the Y axis,
- at $t = 25200\,s$, 15° guidance step on the Z axis,
- at $t = 36000\,s$, 3-axis guidance step back to Sun pointing guidance.

The attitude measurements have been provided by a star tracker, while the actuation is provided by reaction wheels and magnetic rods, thus the model is

$$M_T(t) - \dot{h}_{rw}(t) - \overline{\omega}(t)^{\times} h_{rw}(t) = \psi(\overline{\omega}, \dot{\overline{\omega}})\,\theta_J + \nu(t) - M_d(t)\,, \tag{5.51}$$

where $M_T(t)$ is the torque generated by the magnetorquer. Since (5.51) corresponds to (5.24), with the only exception of $M_T(t)$, the algorithm described in Section 5.5 can be directly applied to the Picard telemetry data. The input-output data are shown in Figure 5.6. During the maneuvers both the input and output are sampled at $1\,Hz$. Since some star tracker measurements were missing, an interpolation has been performed in order to reconstruct the missing data at evenly spaced instants.

For the algorithm implementation, the Butterworth smoother and the prefilter $F(s)$ (5.41) have been applied to the whole dataset. However, for the estimation only the data during the change of satellite attitude have been used since in the time intervals between the maneuvers the satellite was not excited as it can be seen in the quaternion measurements of Figure 5.6. Finally, since there is a large time difference among the maneuvers, the bias effects of the noise and in particular of disturbance torque have been separately estimated for each maneuvers. The solutions consisted in building the following augmented regressor (and corresponding instrument)

$$\begin{aligned}
\psi_{aug,f}(t_k) &= \begin{bmatrix}\psi_f(t_k) & I_3 & 0_3 & 0_3 & 0_3\end{bmatrix} \quad \text{if} \quad t_k \in \text{maneuver } n^{\circ}1\,,\\
\psi_{aug,f}(t_k) &= \begin{bmatrix}\psi_f(t_k) & 0_3 & I_3 & 0_3 & 0_3\end{bmatrix} \quad \text{if} \quad t_k \in \text{maneuver } n^{\circ}2\,,\\
\psi_{aug,f}(t_k) &= \begin{bmatrix}\psi_f(t_k) & 0_3 & 0_3 & I_3 & 0_3\end{bmatrix} \quad \text{if} \quad t_k \in \text{maneuver } n^{\circ}3\,,\\
\psi_{aug,f}(t_k) &= \begin{bmatrix}\psi_f(t_k) & 0_3 & 0_3 & 0_3 & I_3\end{bmatrix} \quad \text{if} \quad t_k \in \text{maneuver } n^{\circ}4\,,
\end{aligned} \tag{5.52}$$

where maneuver $n^{\circ}i$ represents set of data belonging to the i-th maneuver (change of attitude). It should be noted that the estimation of the filter initial conditions are not required in this case, considering the relatively long experiment duration with respect to the transient time.

The obtained estimated inertia parameters are shown in Table 5.2

	J_{11}	J_{22}	J_{33}	J_{23}	J_{13}	J_{12}
$\hat{\theta}_{IV}$ $[kg\,m^2]$	18.65	15.88	21.91	0.16	0.76	-2.15

Table 5.2: Inertia Estimate from the Picard data

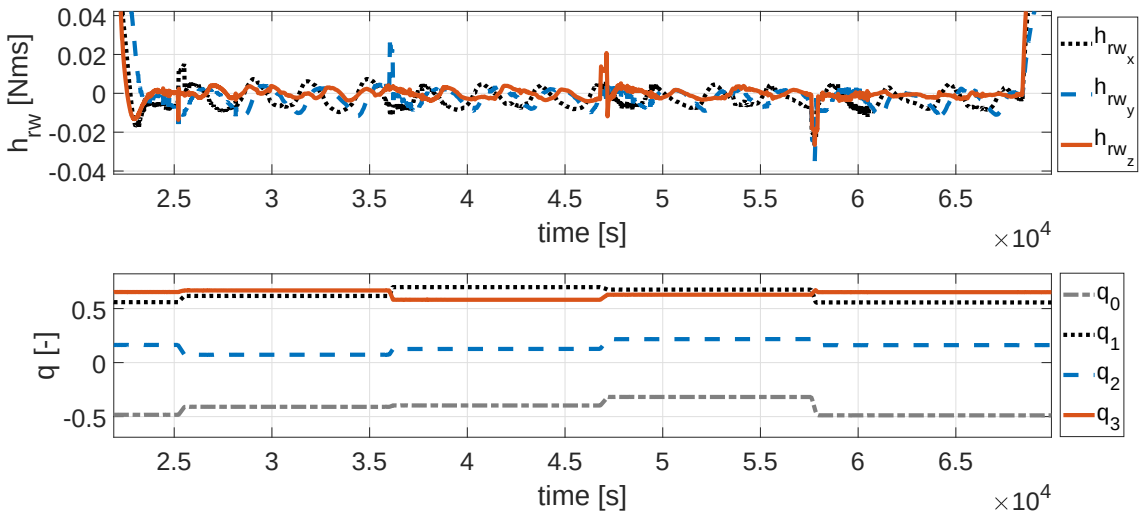


Figure 5.6: Input-output data of the Picard experiment.

In this real scenario we do not have exact values for comparison, so a torque cross-validation is performed to analyze the results. This consists of comparing the total torque $(M_T - \dot{h}_{rw} - \overline{\omega}^{\times} h_{rw})$ with its effect on the system $(\psi(\ddot{\overline{q}}, \dot{\overline{q}}, \overline{q})\,\hat{\theta})$. In order to avoid a large noise amplification, a smoother with cut-off at $0.2\,rad/s$ has been applied to the quaternions before computing $\overline{\omega}$ and $\dot{\overline{\omega}}$, and to h_{rw} for performing the torque cross-validation[19]. The results of this comparison are shown in Figure 5.7. A good fit among the signals can be observed, which indicates a good estimate of the inertia parameters, despite the little excitation given to the satellite in the experiment.

5.8 Conclusions

The IV method for inertia parameter estimation has been adapted to work directly from star tracker measurements. Firstly, a detailed grey-box analysis of the noise has been performed, allowing for the design of a high-performance prefilter for the estimation. Secondly, the method has been augmented in order to better deal with the disturbance torque as well as to deal with the initial transient of the filter if inaccurate filter initial conditions were set. The use of data from the CNES high-fidelity simulator allowed to perform a realistic validation of the proposed method. Since the simulator considers all the main disturbance torques, satellite internal oscillation as well as sensor non-idealities, this is probably the closest benchmark that can be achieved without relying on real-data. The results can be considered very satisfying given the high accuracy achieved. The good performance was also confirmed by the use of real telemetry data from the Picard satellite: the torque cross-validation showed a good match between the input torque and the equivalent torque resulted in the estimated spacecraft model.

So far, the proposed methods focused on the inertia matrix estimation problem. Even if the inertia parameters play a fundamental role in the satellite rotational dynamics, a correct knowledge of the actuator alignments is almost as important to guarantee an effective attitude control

[19]It should be noted that such low value for the smoother cut-off should never be used for the estimation since it is well below the bandwidth of the noise-free signal, thus it does create distortions. In this case it is used just to better visualize the torque cross-validation.

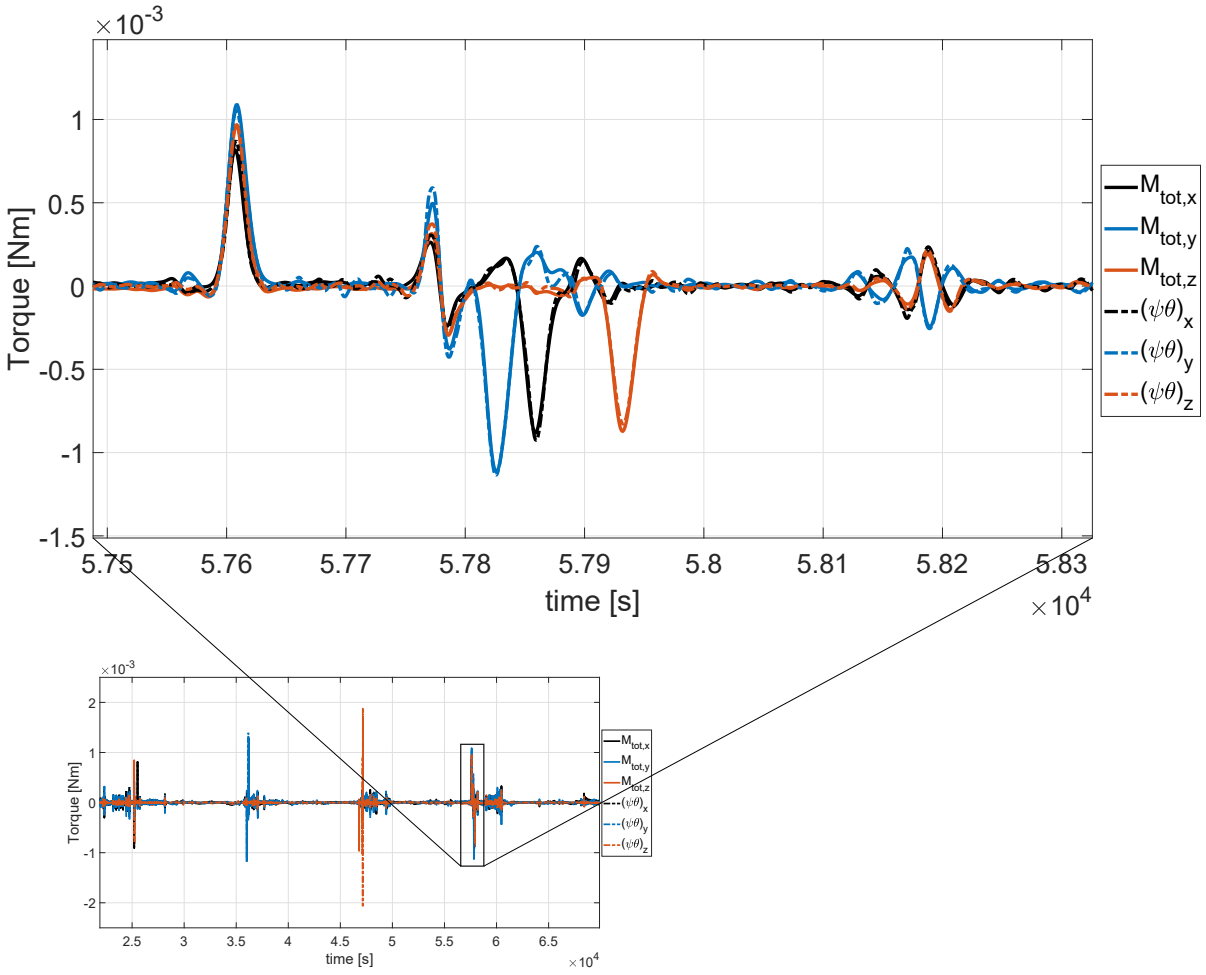


Figure 5.7: Torque cross-validation: measured (solid lines), estimated (dash-dot lines).

design. This is the topic of the next chapter, in which the IV is adapted to perform simultaneous inertia and actuator alignment estimation. Moreover, the star tracker measurements will be considered not perfectly synchronized with the reaction wheel speed readings, thus adding an additional challenge in the estimation problem.

6

Inertia and Actuator Alignment for Gyroless Satellite

Contents

6.1 Satellite Model for Parameter Estimation 77
6.2 Noise Term in the Linear Regression Model 80
6.3 Filter Design 82
6.4 Augmented Model and Delay Estimation 85
6.4.1 Augmented Regressor and Instrument 85
6.4.2 Modified Regressor for Sensor Delay Compensation 86
6.4.3 Proposed Instrumental Variable Based Algorithm 87
6.5 Numerical Simulations and Results 87
6.6 Conclusions 91

The mass distribution is the most important parameter that governs the satellite dynamics, therefore its correct estimation highly affects the attitude control performance, especially when a feedfoward action is used. However, the same can be said for the actuator alignments since they directly affect the generated control torque. For this reason, a simultaneously estimation of all these parameters is a problem of great interest.

In this chapter, a very general case is considered: estimating both the inertia and actuator alignment parameters while still considering a gyroless satellite (as in [Yoon et al., 2017]). It is also considered that the star tracker measurements are not perfectly synchronized with the actuator speed readings, causing a delay between the input and output signals in the telemetry data. This particular problem is not dealt by the satellite parameter estimation literature, although it affects the accuracy of the estimates.

The IV method is further developed for the simultaneous estimation of both parameters, as well as to take into consideration the aforementioned delay. Finally, a more elegant solution is proposed for the filter design compared to the one proposed in Chapter 5. This work, including part of the noise analysis of Chapter 5 has been submitted to a journal [Nainer et al., 2020b].

6.1 Satellite Model for Parameter Estimation

The closed-loop satellite system is the same as in Chapter 5 since it is based on the CNES high-fidelity simulator for the MicroCarb satellite, therefore the reader may refer to Section 5.1 for

further details.

The satellite modeling is still based on the Euler's equations, however, also the actuator alignment parameters must be linearly extracted in order to apply an IV method. As shown in Section 2.2.2, the satellite rotational dynamics is described as follows

$$-\dot{h}_{rw}(t) - \omega(t)^{\times} h_{rw}(t) + M_d(t) = J\,\dot{\omega}(t) + \omega(t)^{\times} J\,\omega(t)\,, \tag{6.1}$$

with

$$h_{rw}(t) = \sum_{i=1}^{n} J_{rw,i}\Omega_i(t)a_i\,. \tag{6.2}$$

As before (3.2), we can extract the inertia parameters

$$\dot{h}_{rw} - \omega^{\times} h_{rw} + M_d = \Big(\Gamma(\dot{\omega}) + \omega^{\times}\Gamma(\omega)\Big)\,\theta_J\,, \tag{6.3}$$

with $\theta_J = [J_{11}\,,\, J_{22}\,,\, J_{33}\,,\, J_{23}\,,\, J_{13}\,,\, J_{12}]^T$, and where the dependency on the time t_k has been omitted for sake of simplicity.

The actuator alignment parameters a_i, defined as unit vectors parallel to the reaction wheel axes, can be treated in a similar way as done for the inertia parameter vector θ_J:

- A first intuitive approach consists in extracting the alignment parameter vectors a_i, or the alignment error vectors (with respect to the previously known value $\overline{a_i}$) $e_{a,i} = a_i - \overline{a_i}$ from (6.2). In both cases unit norm constraints must be respected: $||a_i||_2 = 1$ or $||\overline{a}_i + e_{a,i}||_2 = 1$. In the case of alignment error estimation the equation is

$$h_{rw} = \overline{h}_{rw} + \chi(J_{rw}, \Omega)\, e_a = \overline{h}_{rw} + \begin{bmatrix} J_{rw,1}\Omega_1 I_3 & J_{rw,2}\Omega_2 I_3 & \cdots & J_{rw,n}\Omega_n I_3 \end{bmatrix} e_a\,, \tag{6.4}$$

  with $e_a = [e_{a_1},\, e_{a_2}\,\cdots,\, e_{a_n}]^T$ and where $\overline{h}_{rw}$ is the reaction wheel angular momentum computed with the nominal $\overline{a}_i$ directions.

  However, this approach would require to solve a quadratic programming problem with n quadratic equality constraints ($||\overline{a}_i + e_{a,i}||_2 = 1$). The addition of multiple quadratic constraints transforms the estimation process into a nonconvex optimization problem [Bar-On and Grasse, 1994; Bar-On and Grasse, 1997; Golub and Von Matt, 1991].

- In this work, an alternative solution is chosen. Since there is no guarantee to reach the global minimum of the previous cost function, a different model has been considered, which consists in representing the misalignment equation with two parameters instead of three as

$$a_i \simeq \overline{a}_i + \alpha_i u_{1,i} + \beta_i u_{2,i} \qquad \text{for} \qquad i = 1, \cdots n\,, \tag{6.5}$$

  where $u_{1,i}$ and $u_{2,i}$ are two unit vectors orthogonal to the direction $\overline{a}_i$ and to each other (their computation is described in Section 6.4). With this notation, we can linearly extract the misalignment parameters α_i and β_i while the quadratic constraint can be avoided[20]

$$h_{rw} = \overline{h}_{rw} + \rho(J_{rw}, \omega, \Omega, \dot{\Omega}, u_1, u_2)\theta_e \tag{6.6}$$

[20]Equation (6.5) does not yield an exactly unit norm result, but it will be shown that a solution can be found.

with

$$\rho(\cdot) = \left[J_{rw,1}\dot{\Omega}_1 u_{1,1} + \omega^{\times}(J_{rw,1}\Omega_1 u_{1,1}) \quad J_{rw,1}\dot{\Omega}_1 u_{2,1} + \omega^{\times}(J_{rw,1}\Omega_1 u_{2,1}) \quad \cdots \right. \\ \left. \cdots \quad J_{rw,n}\dot{\Omega}_n u_{2,n} + \omega^{\times}(J_{rw,n}\Omega_n u_{2,n})\right], \tag{6.7}$$

and where

$$\theta_e = [\alpha_1\,,\,\beta_1\,,\,\alpha_2\,,\,\beta_2\,,\,\cdots\,,\,\alpha_n\,,\,\beta_n]^T \tag{6.8}$$

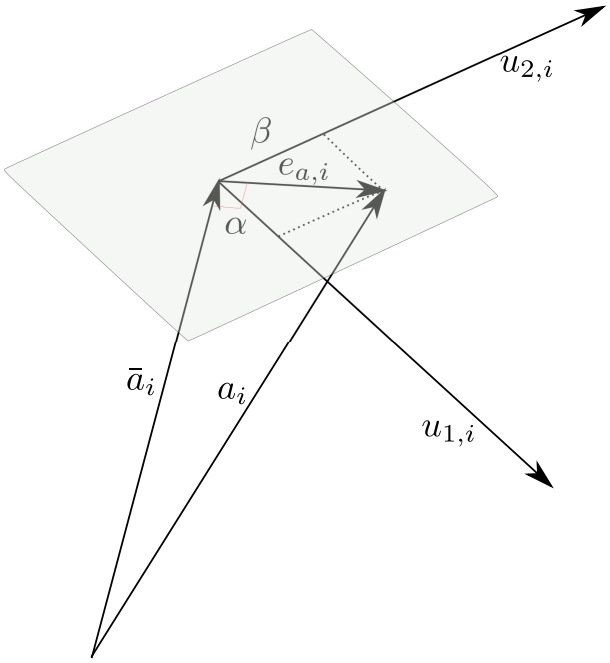


Figure 6.1: Geometrical representation of α and β.

Given this misalignment description (6.6), the full equation in linear regression form becomes the following

$$-\dot{\overline{h}}_{rw} - \omega^{\times}\overline{h}_{rw} + M_d = \left[\Gamma(\dot{\omega}) + \omega^{\times}\Gamma(\omega) \quad \rho(J_{rw}, \omega, \Omega, \dot{\Omega}, u_1, u_2)\right] \begin{bmatrix}\theta_J \\ \theta_e\end{bmatrix}. \tag{6.9}$$

Since only star tracker measurements are available, (6.9) must be written as a function of the quaternions q rather than the angular rate ω. For this purpose it is sufficient to substitute ω and $\dot{\omega}$ with $2W(q)\dot{q}$ and $2W(q)\ddot{q}$ respectively (see (2.8)), which leads to the linear regression estimation model

$$-\dot{\overline{h}}_{rw} - \left(2W(q)\dot{q}\right)^{\times}\overline{h}_{rw} + M_d = \\ = \left[\Gamma(2W(q)\ddot{q}) + \left(2W(q)\dot{q}\right)^{\times}\Gamma(\omega) \quad \rho(J_{rw}, q, \dot{q}, \Omega, \dot{\Omega}, u_1, u_2)\right] \begin{bmatrix}\theta_J \\ \theta_e\end{bmatrix}. \tag{6.10}$$

It should be noted, that even if the terms α and β are small, (6.5) yields a non-unit vector (Figure 6.1). An Euclidean normalization is therefore necessary after the estimation of these parameters. In order to improve the accuracy of the estimates, the estimation process is iterated until α and β become negligible, as it will be described in more detail in Section 6.4.

6.2 Noise Term in the Linear Regression Model

For the time derivative estimation of the quaternions, $\dot{\overline{q}}$ and $\ddot{\overline{q}}$, as in Section 5.3, a Butterworth based smoother is used in order to reduce the noise amplification. One of the reasons behind this choice is the maximally flat magnitude in the pass-band of the Butterworth filter. However, as already mentioned in Section 5.3, the smoother cut-off is limited by the bandwidth of the system noise-free input-output signals and therefore an additional prefilter must be designed in order to guarantee estimates with low variance.

The overall noise term for the new model must be analyzed in order to design a suitable filter. Since the encoders used for the reaction wheel speed measurements are very accurate, we can consider them noise free without loss of accuracy.

From (5.7)-(5.18), it has been shown how by computing the angular velocity $\overline{\omega}$ from the measured quaternions $\overline{q}$, the overall noise on the angular velocity can be accurately approximated as

$$\tilde{\omega} = \left[\dot{e}_x \, , \, \dot{e}_y \, , \, \dot{e}_z\right] . \tag{6.11}$$

Therefore, as in Chapter 5, we can directly analyze the noise in (6.9), instead of (6.10), replacing the noise-free angular rate ω with $\overline{\omega} - \tilde{\omega}$.

$$-\dot{\overline{h}}_{rw} - (\overline{\omega} - \tilde{\omega})^{\times} \overline{h}_{rw} + M_d = \Big[\underbrace{\Gamma(\dot{\overline{\omega}} - \dot{\tilde{\omega}}) + (\overline{\omega} - \tilde{\omega})^{\times} \Gamma(\overline{\omega} - \tilde{\omega})}_{\phi_J} \quad \underbrace{\rho(J_{rw}, (\overline{\omega} - \tilde{\omega}), \Omega, \dot{\Omega}, u_1, u_2)}_{\phi_a} \Big] \begin{bmatrix} \theta_J \\ \theta_e \end{bmatrix} . \tag{6.12}$$

The study of the noise on the first term ϕ_J of the full regression matrix in (6.12) relative to the inertia parameter remains unchanged with respect to the one done in (4.7)-(4.13). Therefore, similarly as in (4.7), we can substitute the noise-free term by the measured equivalent and two additional noise terms:

$$\phi_J \, \theta_J = \psi \, \theta_J - \delta \theta_J - \epsilon \theta_J \, . \tag{6.13}$$

with

$$\begin{aligned} \psi &= \Gamma(\dot{\overline{\omega}}) + \overline{\omega}^{\times} \Gamma(\overline{\omega}) \, , \\ \delta &= \Gamma(\dot{\tilde{\omega}}) + \tilde{\omega}^{\times} \Gamma(\tilde{\omega}) \, , \\ \epsilon &= \omega^{\times} \Gamma(\tilde{\omega}) + \tilde{\omega}^{\times} \Gamma(\omega) \, , \end{aligned} \tag{6.14}$$

and where $-\delta\theta_J - \epsilon\theta_J$ represents the "noisy" terms. The same analysis done in Section 5.2.2 remains valid regarding $-\delta\theta_J - \epsilon\theta_J$, thus the zero-mean property is not respected if $\tilde{\omega}$ has different standard deviation on its three components.

For the second term ρ of the full regression matrix in (6.12), under the assumption of negligible noise on the reaction wheel speeds, the analysis is straightforward. Given

$$\phi_a \theta_e = \Big(\rho(J_{rw}, (\overline{\omega} - \tilde{\omega}), \dot{\Omega}, u_1, u_2)\Big)\theta_e \,, \tag{6.15}$$

the overall noise is

$$\begin{bmatrix} \tilde{\omega}^{\times}(J_{rw,1}\Omega_1 u_{1,1}) & \tilde{\omega}^{\times}(J_{rw,1}\Omega_1 u_{2,1}) & \cdots & \tilde{\omega}^{\times}(J_{rw,n}\Omega_n u_{2,n}) \end{bmatrix} \theta_e \,, \tag{6.16}$$

which maintains the zero-mean property of $\tilde{\omega}$.

Finally, we need to consider the noise effect in the left-hand side of (6.12). In this part we only have the following noise and disturbance terms

$$\tilde{\omega}^{\times}\overline{h}_{rw} + M_d \,. \tag{6.17}$$

Therefore (6.12) can be rewritten as

$$-\dot{\overline{h}}_{rw} - \overline{\omega}^{\times}\overline{h}_{rw} = \begin{bmatrix} \psi(\overline{\omega}, \dot{\overline{\omega}}) & \rho(J_{rw}, \overline{\omega}, \dot{\Omega}, u_1, u_2) \end{bmatrix} \begin{bmatrix} \theta_J \\ \theta_e \end{bmatrix} + \nu - M_d \,, \tag{6.18}$$

where ν collects all the effects due to measurement noise. In order to simplify the notation, (6.18) will be rewritten as

$$g(h_{rw}, \overline{\omega}) = f(\overline{\omega}, J_{rw}, \Omega, u_1, u_2)\,\theta + \nu - M_d \,, \tag{6.19}$$

where

$$\theta = \begin{bmatrix} \theta_J \\ \theta_e \end{bmatrix} . \tag{6.20}$$

Disturbance Torque

The disturbance torque M_d depends on the following different factors: gravity gradient torque, aerodynamic torque, solar radiation pressure, magnetic torque, as well as some internal oscillation in the satellite structure (e.g. solar panel oscillation). The magnitudes of these components depend on the satellite orbit. The sum of these terms can be well approximated as

$$M_d = M_{d,d} + M_{d,s} \,, \tag{6.21}$$

where $M_{d,d}$ contains the "deterministic" part that can be estimated with sufficient accuracy, while $M_{d,s}$ can be considered as a stochastic term. For a low-Earth orbit satellite, the deterministic part can be approximated as [Sidi, 1997]

$$M_{d,d} = M_0 + M_1 sin(\omega_0 t + \phi_1) + M_2 sin(2\omega_0 t + \phi_2) + M_g \,, \tag{6.22}$$

where ω_0 is the satellite orbital rate and M_g is the gravity gradient torque that can be accurately computed. The stochastic term of the torque is approximated as

$$\dot{M}_{d,s} = -\gamma M_{d,s} + \eta \qquad \text{where} \qquad \eta_x \sim \eta_y \sim \eta_z \sim N(0, \sigma_\eta^2) \,, \tag{6.23}$$

and where $\gamma > 0$ is a scalar value determining the frequency domain of the disturbance. A similar approach has been proposed in [Yoon et al., 2017], where instead a simple continuous random walk process has been utilized ($\dot{M}_{d,s} = \eta$). Since the disturbance torque amplitude is bounded and a common integral random walk grows indefinitely, the process (6.23) has been preferred.

6.3 Filter Design

As in Chapter 5, we perform a linearization of the noise model and from it we design a sub-optimal filter that will ensure low variance estimates from the IV approach. Even before performing any linearization, we can observe that the regressor term $f(\overline{\omega}, J_{rw}, \Omega, u_1, u_2)$ does not have any integrator inside, but only derivatives, therefore for the filter construction we can set $A(s) = 1$, and only $\nu(t)$ and $M_d(t)$ affect the filter design.

From Section 6.2, we have

$$\nu = -\delta\theta_J - \epsilon\theta_J + \tilde{\omega}^{\times}\Big(\sum_{i=1}^{n}(J_{rw,i}\Omega_i u_{1,i} + J_{rw,i}\Omega_i u_{2,i})\Big) - \tilde{\omega}^{\times}\overline{h}_{rw}\,. \tag{6.24}$$

By linearizing the ν term around the operating conditions a linear approximation of ν is obtained (of the type $\frac{D(s)}{C(s)}$, where in this case $C(s) = 1$). Since the noise term bandwidth is not affected by its sign, the linearization point has been chosen as the average of the absolute values of h_{rw}, ω and Ω during the satellite maneuver:

$$h^*_{rw} = \frac{1}{N}\sum_{k=1}^{N}\Big|h_{rw}(t_k)\Big|\,, \qquad \omega^* = \frac{1}{N}\sum_{k=1}^{N}\Big|\overline{\omega}(t_k)\Big| \qquad \Omega^* = \frac{1}{N}\sum_{k=1}^{N}\Big|\Omega(t_k)\Big|\,. \tag{6.25}$$

By linearizing ν, following the same steps shown in (5.33)-(5.37), a noise polynomial matrix of the following form is obtained

$$D_l(s) = D_1\, s + D_2\, s^2\,, \tag{6.26}$$

where a continuous-time notation is kept to have a simpler representation. Since the noise $\tilde{\omega} \simeq \dot{e}$ (as shown in (5.18)), there are no terms without differential operator s. To simplify the filter, the matrices in (6.26) are diagonalized by calculating the diagonal terms as a mean square of the row values.

As done in the previous chapter, we have to combine the noise model $D_l(s)$ with a disturbance torque model in order to obtain the sub-optimal filter. In case of a disturbance free satellite, the inverse of the noise model ($\frac{1}{D_l(s)}$), weighted by the different standard deviations of the three e components, would be the sub-optimal filter.

For the disturbance, as mentioned in Section 6.2, we consider a highly colored noise (that is similar to a bounded random walk process), which represents the residual term after having estimated and removed the constant and sinusoidal components by means of an augmented regressor (see Section 6.4). Also the gravity torque can be removed, since it can be accurately calculated from the satellite attitude with respect to the Earth [Curtis, 2013]. Therefore, the following transfer function is used to describe the stochastic components of the disturbance

$$M_{d,s}(s) = \frac{\gamma}{s+\gamma}\,, \tag{6.27}$$

where γ is an empirical value (e.g. $\gamma = 0.002\,rad/s$) to describe the low-pass behavior of the disturbance torque.

Since these two terms (noise and disturbance) are uncorrelated, the two power spectra add together, with their respective gains, giving an overall spectrum like the one of Figure 6.2. In this figure, the colored solid lines show the two "noise" components, while the black dashed line shows the overall spectrum. While in Chapter 5 we have approximated the overall spectrum with a simple summation of the two single spectra, we want here to find a more elegant solution.

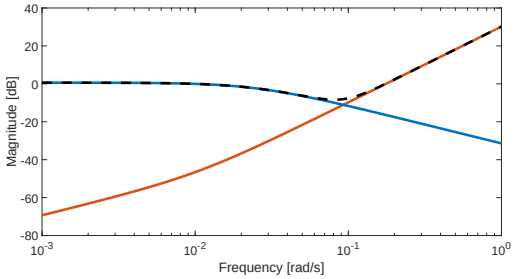


Figure 6.2: Disturbance torque spectrum approximation in solid blue line, measurement noise term in solid red line, and overall noise spectrum in dashed black line.

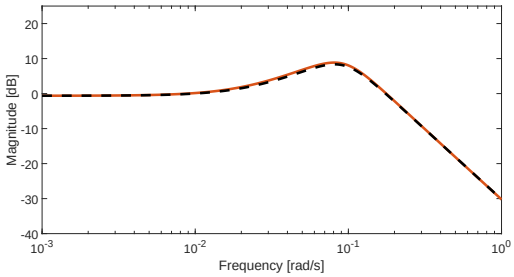


Figure 6.3: Reciprocal of the overall noise spectrum in dashed black line, and its third-order filter approximation in solid red line.

Following the idea of the optimal prefilter for the Box-Jenkins model, we are looking for a filter which has a reciprocal magnitude with respect to the overall noise term. However, the spectrum of Figure 6.2 is not necessarily a rational spectrum and neither its reciprocal. We are therefore looking for a rational transfer function, possibly of a low order, which approximates the reciprocal of the overall noise spectrum. There are different options available in the literature for this purpose. An alternative method, based on the cepstrum properties and described in detail in Appendix A, has been developed and used for this case. The result of this filter synthesis is shown in Figure 6.3, where the dashed black line represents the reciprocal of the overall noise spectrum, while the red line represents an approximation with a third order filter $F(s)$, which will be later used as a prefilter for the parameter estimation algorithm. The structure of $F(s)$ is equivalent to the one described in (5.41) of Section 5.4:

$$F(s) = \frac{c_1 s + c_0}{d_3 s^3 + d_2 s^2 + d_1 s + d_0} \,. \tag{6.28}$$

Calculation of the Prefilter from the Residuals

Instead of following a full grey-box approach for the noise model, and thus the corresponding filter design, a black-box approach can be taken. This consists in estimating the noise ARMA model from the residuals and to use its inverse transfer function as prefilter, as shown in Section 3.3 [Young, 2011]. In this case, the residuals are simply defined as $\xi_r = g(t_k) - f(t_k)\,\hat{\theta}$. However, in this satellite scenario, this refined IV approach presents some additional difficulty with respect to the theoretical approach based on the Box-Jenkins model (Figure 3.4). In fact, since this is an iterative process, the noise model is initialized as a simple white noise and thus the initial prefilter is based only on the system transfer function denominator. For the satellite inverse model, given the presence of no integrator terms, this does result in the use of no prefilter for the first iteration and thus in highly unreliable parameter estimates. In order to avoid this issue, even for the first iteration, a low-pass filter has been applied for the first step of the iteration.

The overall iterative process is then as follows

1. Obtain a first estimate of θ with the use of a second order low-pass filter

$$F(s) = \frac{1}{(\gamma\, s + 1)^2} \,. \tag{6.29}$$

2. With the previously estimated satellite parameters $\hat{\theta}$, compute the residuals

$$\xi_r = g(t_k) - f(t_k)\,\hat{\theta}\,. \tag{6.30}$$

3. Estimate the parameters of the ARMA noise model

$$\xi_r(t_k) = \frac{D(z^{-1})}{C(z^{-1})} e(t_k) \tag{6.31}$$

where $e(t_k)$ is assumed to be white noise source.

4. prefilter the $g(t)$ and the regressor $f(t)$ with the filter

$$F(z^{-1}) = \frac{C(z^{-1})}{D(z^{-1})} \tag{6.32}$$

5. Based on the prefiltered data, compute a new IV estimation on the satellite parameters.

Repeat steps 2-3-4-5 till convergence.

The estimation of the noise model $\xi_r(t_k) = \frac{D(z^{-1})}{C(z^{-1})} e(t_k)$ presents additional difficulties in the satellite scenario due to the system inversion. The majority of the residual are due to the derivation of the noise, so the overall expected noise model is the inverse of a low-pass filter. This is a very uncommon noise model, and not all the commonly used option in the literature can reliably estimate this model. For example, the *ivarma* routine from the CAPTAIN toolbox [Taylor et al., 2007] did perform poorly. The *armax* routine from Matlab, as well at the cepstrum based approach of Appendix A, resulted capable of correctly estimating this ARMA model.

Applying this cepstrum based method, or the *armax* routine the transfer function $\frac{D(z^{-1})}{C(z^{-1})}$ can be estimated from the residuals. The prefilter to be used for the parameter estimation consists then in the inverse of the noise estimated model $\frac{C(z^{-1})}{D(z^{-1})}$. The noise model, and thus the prefilter, can be also converted into a continuous-time model (CARMA) in order to have a form similar to (6.28) and allowing to use the same approach for the filter initial condition estimation. An example of the estimated prefilters $\hat{H}^*(s)$ (one for each axis), in continuous time, is shown in Figure 6.4. For sake of clarity, a unit steady state gain has been imposed for each filter in this plot.

It should be noted that the prefilter design from the residual analysis yields low-pass filters with a cut-off slightly higher than the ones from the tailor-made approach at the beginning of this section (Figure 6.3). From tests based on the experiment results, the lower cut-off of the tailor-made prefilter (6.28) seems to work slightly better than the one obtained from the residuals. Moreover, in order to have a good estimate of the ARMA noise model at low-frequencies, a long experiment is required since for an experiment of duration T_e the frequency resolution is $\frac{1}{2\,T_e}$.

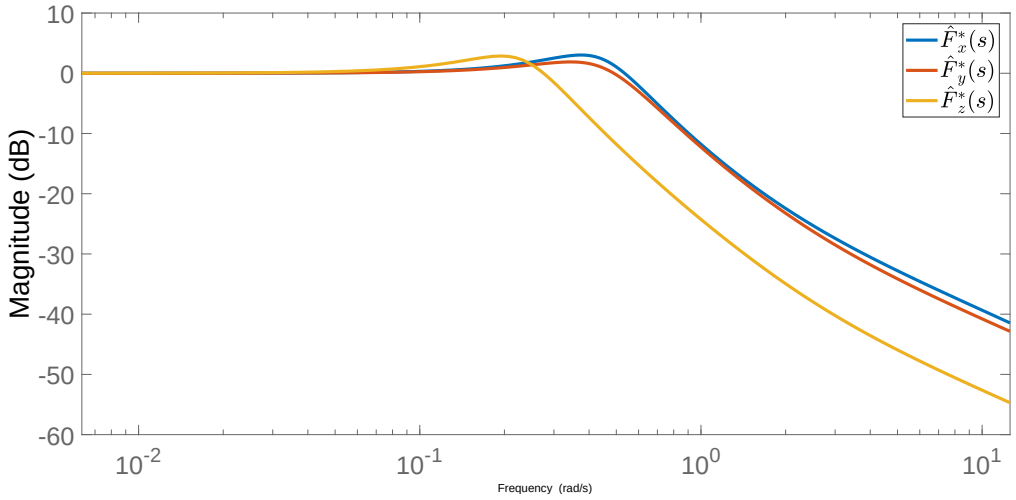


Figure 6.4: Estimated prefilters from the residuals.

6.4 Augmented Model and Delay Estimation

6.4.1 Augmented Regressor and Instrument

In Section 5.2, it has been shown how, depending on the star tracker orientation and noise standard deviation, the noise term ν may have a non-zero expected value. Moreover, the disturbance torque contains some harmonic terms (which depend on the orbital frequency ω_0) that should be estimated and removed. The solution to deal with these terms consists in augmenting the regressor (and therefore the instrument as well):

$$\begin{aligned} f^*_{aug}(t_k) &= \begin{bmatrix} f(t_k) & I_3 & I_3 sin(\omega_0 t_k) & I_3 cos(\omega_0 t_k) & I_3 sin(2\omega_0 t_k) & I_3 cos(2\omega_0 t_k) \end{bmatrix} \\ Z^*_{aug}(t_k) &= \begin{bmatrix} Z(t_k) & I_3 & I_3 sin(\omega_0 t_k) & I_3 cos(\omega_0 t_k) & I_3 sin(2\omega_0 t_k) & I_3 cos(2\omega_0 t_k) \end{bmatrix} \end{aligned} \quad \forall k \in [1 \cdots N], \tag{6.33}$$

where the identity matrix I_3 allows the estimation of the overall constant terms due to the measurement noise and the disturbance $M_d(t)$, while the other terms are added to estimate the sinusoidal components of $M_d(t)$.

In the satellite scenario, the filter (Figure 6.3) has a slow dynamics, therefore the initial conditions of the filter have a significant impact on the inertia estimates. Waiting for the end of the transient behavior may result in discarding a significant portion of the data. A better solution consists in estimating the filter initial condition together with the satellite parameters. To do so, the (filtered) regressor and instrument are further augmented as done in Chapter 5

$$\begin{aligned} f_{aug,f}(t_k) &= \begin{bmatrix} f^*_{aug,f}(t_k) & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k) \end{bmatrix}, \\ Z_{aug,f}(t_k) &= \begin{bmatrix} Z^*_{aug,f}(t_k) & T_{r1}(t_k) & T_{r2}(t_k) & T_{r3}(t_k) \end{bmatrix}, \end{aligned} \tag{6.34}$$

where $T_{ri}(t_k)$ are the transient behaviors corresponding to each one of the filter poles (in this case we considered a third order filter). The fully extended parameter vector θ_{ext} contains 24 parameters in addition to θ.

6.4.2 Modified Regressor for Sensor Delay Compensation

The input and output measurements could be not perfectly synchronized. In our case, the star tracker measures have an unknown constant delay of a fraction of sampling period[21] with respect to the RW speed measures. To overcome this issue, a specific approach is proposed.

Let us consider the following system

$$y(t_k) = \psi(u(t_k))\,\theta\,, \tag{6.35}$$

and suppose the input u is available at instants $t_k - t_\delta$ instead of t_k. Instead of solving the regressor equation[22]

$$\hat{\theta} = \arg\min_{\theta} \sum_{k=1}^{N} \Big|\Big| y(t_k) - \psi(t_k)\,\theta \Big|\Big|^2\,, \tag{6.36}$$

the following augmented equation is solved:

$$\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{bmatrix} = \arg\min_{\theta_1\theta_2} \sum_{k=1}^{N} \Big|\Big| \begin{bmatrix}\psi(u(t_k - t_\delta)) & \psi(u(t_{k+1} - t_\delta))\end{bmatrix} \begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} - y(t_k) \Big|\Big|^2\,. \tag{6.37}$$

The first term in the regression can be considered as a linear interpolation of the function ψ between instants $t_k - t_\delta$ and $t_{k+1} - t_\delta$:

$$\begin{aligned}
&\begin{bmatrix}\psi(t_k - t_\delta) & \psi(t_{k+1} - t_\delta)\end{bmatrix} \begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}\\
&= \Big(\psi(t_k - t_\delta)\frac{\theta_1}{\theta_1+\theta_2} + \psi(t_{k+1} - t_\delta)\frac{\theta_2}{\theta_1+\theta_2}\Big)(\theta_1+\theta_2)\\
&\simeq \psi\Big(t_k + \frac{\theta_2}{\theta_1+\theta_2}\Delta T - t_\delta\Big)(\theta_1+\theta_2)\,,
\end{aligned} \tag{6.38}$$

where the u terms have been omitted for ease of notation. It should be noted that in these equations the operations on θ must be considered element-wise.

Since the best model fit is expected at time t_k, solving for

$$t_k = t_k + \frac{\theta_2}{\theta_1+\theta_2}\Delta T - t_\delta \tag{6.39}$$

we obtain:

$$t_\delta = \frac{\theta_2}{\theta_1+\theta_2}\Delta T \tag{6.40}$$

$$\theta = \theta_1 + \theta_2\,. \tag{6.41}$$

In our specific case, since we want to correct the delay of the star tracker measures with respect to the reaction wheel speed readings, only the first part of the regressor (6.18) related to the inertia parameters θ_J has to be duplicated, leaving the remaining part unchanged. From the two parameter vectors θ_{J1} and θ_{J2} we can calculate the inertia parameter as $\theta_J = \theta_{J1} + \theta_{J2}$.

It should be noted, that in this particular scenario, the gyroscopic term in the regressor (6.7) related to the alignments parameters combines together input and output states with incorrect timing. Therefore, a small error due to the delay is still present, even if negligible as shown in the numerical example hereafter.

[21] The proposed approach can be easily extended to delays larger than one sampling period.

[22] The LS solution is considered here for simplicity, but the extension to the IV solution is straightforward.

6.4.3 Proposed Instrumental Variable Based Algorithm

In this chapter we choose to use a delayed regressor as instrument Z, also to test how this alternative approach performs in this application. The delay should be sufficiently large to guarantee the second property of (3.23), but still to be as small as possible to have a high correlation with the regressor $f(\cdot)$ (first property of (3.23)). For this particular application, since the overall noise comes mainly from a numerical derivative and double derivative of a white noise, a delay of a few samples is sufficient to obtain a negligible correlation with the noise. Moreover, given also the slow dynamics of the satellite, a high correlation between the instrument and the regressor is maintained. The overall estimation algorithm can be synthesized as follows

---

• Apply the Butterworth smoothing filter to the quaternion measurements $\overline{q}(t)$ and compute the time derivatives $\dot{\overline{q}}(t)$, $\ddot{\overline{q}}(t)$;

• Initialize the reaction wheel alignment vectors with their nominal values $\bar{a_i}$;

do

• Compute u_1 and u_2 for each reaction wheel:

for each direction a_i, compute the cross product between a_i and the three main axes x-y-z

$$v_{1_i} = \overline{a}_i^{\times}[1\ 0\ 0]^T\,, \quad v_{2_i} = \overline{a}_i^{\times}[0\ 1\ 0]^T\,, \quad v_{3_i} = \overline{a}_i^{\times}[0\ 0\ 1]^T\,, \tag{6.42}$$

from the vector v_i having the maximum norm, compute

$$u_{1_i} = \frac{v_{j_i}}{||v_{j_i}||}\,, \tag{6.43}$$

compute a second unit vector u_{2_i}, normal to u_{1_i} and to $\overline{a}_i$

$$u_{2_i} = \overline{a}_i^{\times} u_{1_i}\,, \tag{6.44}$$

• Build the linear regression equation (3.8) and filter the regressor $f(\cdot)$, the instrument $Z(\cdot)$ and the input vector $g(\cdot)$ with the prefilter $F(s)$ designed as in Section 6.3,

• Estimate the parameters with the IV method

$$\hat{\theta}_{IV} = \left[\sum_{k=1}^{N} Z_{aug,f}^T(t_k) f_{aug,f}(t_k)\right]^{-1} \left[\sum_{k=1}^{N} Z_{aug,f}^T(t_k) g_f(t_k)\right], \tag{6.45}$$

where $(\cdot)_f$ represents the filtered matrix,

• Update the alignment vectors $\overline{a}_i(k) = \overline{a}_i(k-1) + \alpha_i u_{1_i} + \beta_i u_{2_i}$, and normalize them. Also update the prefilter $F(s)$.

while $(\sum_{i=1}^{n} \alpha_i^2 + \beta_i^2) < threshold$;

6.5 Numerical Simulations and Results

A high-fidelity simulator from CNES, featuring the Microcarb gyroless satellite, has been used to test the proposed estimation method. The true inertia matrix of the spacecraft as well as the reaction wheel alignment matrix are

$$J = \begin{bmatrix} 20.2609 & -1.3756 & -2.1707 \\ -1.3756 & 23.2558 & -0.3744 \\ -2.1707 & -0.3744 & 27.8365 \end{bmatrix} [kg\,m^2]\,, \tag{6.46}$$

$$A_{rw} = \begin{bmatrix} -0.0406 & 0.7302 & -0.0201 & -0.6940 \\ 0.6983 & 0.6832 & 0.7045 & 0.7200 \\ -0.7146 & -0.0067 & 0.7094 & 0.0055 \end{bmatrix}. \tag{6.47}$$

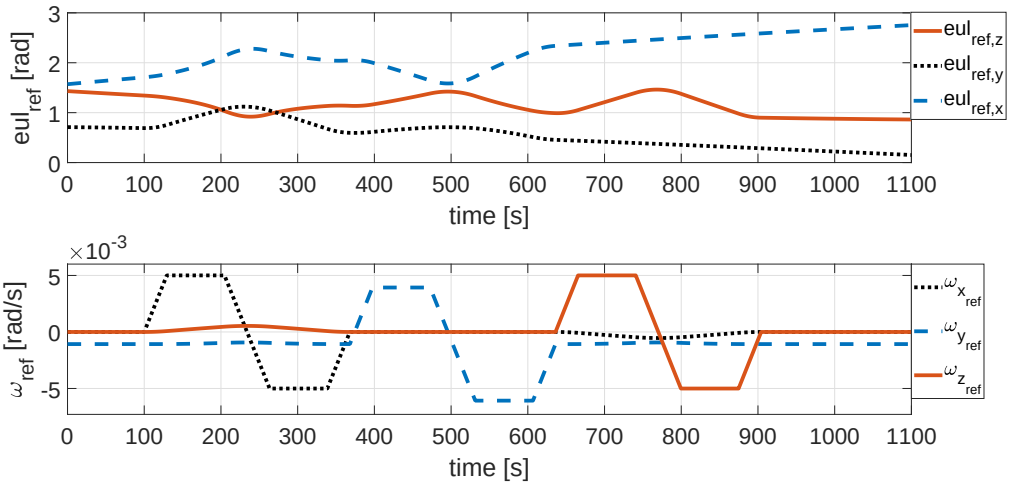


Figure 6.5: Attitude and angular rate reference profile.

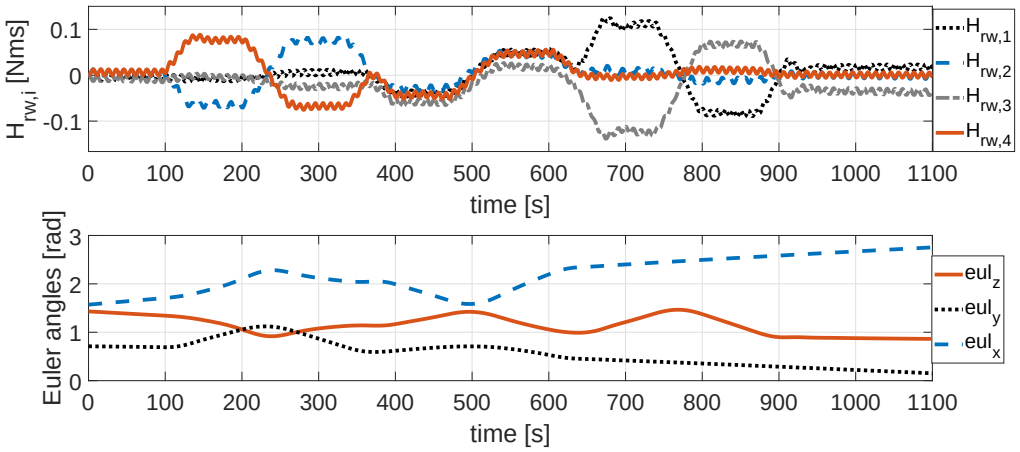


Figure 6.6: Angular momentum for each of the 4 reaction wheels (at the top). Satellite attitude expressed as Euler angles (x-y-z) at the bottom.

Four reaction wheels are used in a pyramidal configuration allowing one level of redundancy in case of reaction wheel failure. This configuration also allows a change of the reaction wheel speeds without changing the overall torque applied to the satellite by moving along the null-space vector. The star tracker sampling frequency is set to 16 Hz, while the reaction wheel speeds were

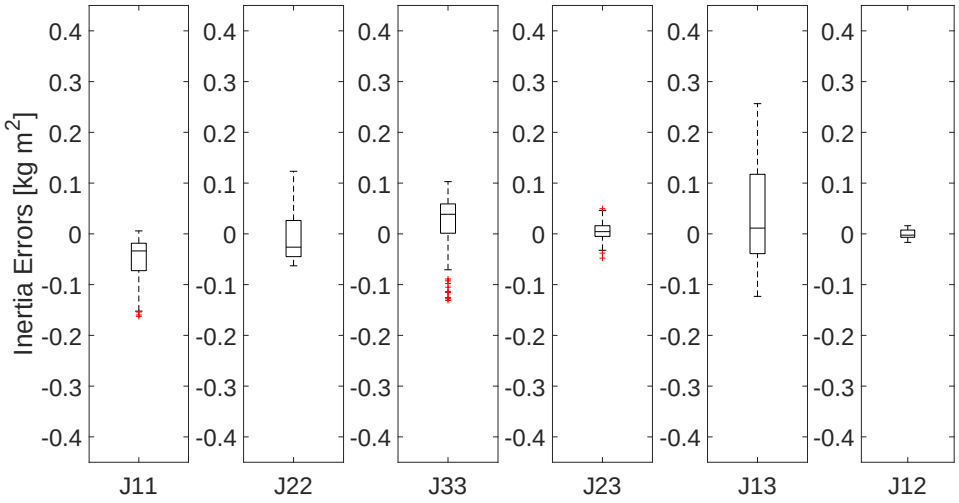


Figure 6.7: Box plot for the inertia estimate errors.

sampled at $1\,Hz$. The star tracker white noise component has the following standard deviation $\sigma_x = \sigma_y = 11.4 \cdot 10^{-6} rad$ and $\sigma_z = 93.0 \cdot 10^{-6} rad$. In the high-fidelity simulator, the star tracker model has also bias and harmonic (at the orbital frequency) noise terms. The bias terms are set to value of around $55 \cdot 10^{-6} rad$. Additionally, a delay of $0.125\,s$ has been set for the attitude sensor readings. All the main types of disturbance torques, typical for a low-Earth orbit satellite, are considered: aerodynamic torque, solar radiation pressure, magnetic torque, gravity gradient torque, as well as internal oscillation in the structure (mainly due to flexible elements like solar panels). During the simulations, the solar panels are kept at the same position in order to avoid any change of the satellite inertia during the maneuver. The initial misalignment for the reaction wheels is set to be around $2°$ for each actuator, while the initial inertia values have been set with an error of about 10% with respect to the true values. The attitude control law is described in [Genin and Viaud, 2018]. Additionally, a mild null space excitation has been added in order to improve the accuracy of the estimated parameters. To do so, the following additional torque has been applied to the 4 reaction wheels

$$T_{rw,0} = C_t N_{rw}\,, \tag{6.48}$$

where N_{rw} is the null vector of A_{rw} ($A_{rw} N_{rw} = 0$), and C_t is the magnitude of the null space torque. In order to avoid reaction wheel saturation, the sign of C_t was changed every $5\,s$. The reference guidance profile is shown in Figure 6.5, and it is kept the same for every simulation, while the orbital position was set each time differently in order to have different disturbance torques for each experiment. Figure 6.6 shows the input-output of the satellite for a single simulation. Since the input and output signal have different sampling frequency, the actuator speed are upsampled at $4\,Hz$, while the attitude measurement are downsampled (after the smoothing operation) at $4\,Hz$. For the instrument Z, a delay of 6 samples, equivalent to $1.5\,s$, was used. This delay is enough to guarantee the second property in (3.23), while, due to the slow maneuvers of the satellite, the correlation between Z and f remains still very high.

The results from a Monte Carlo simulation of 100 runs are shown in Figure 6.7 and 6.8, where the errors of the inertia and alignment parameters have been represented as box plots. The

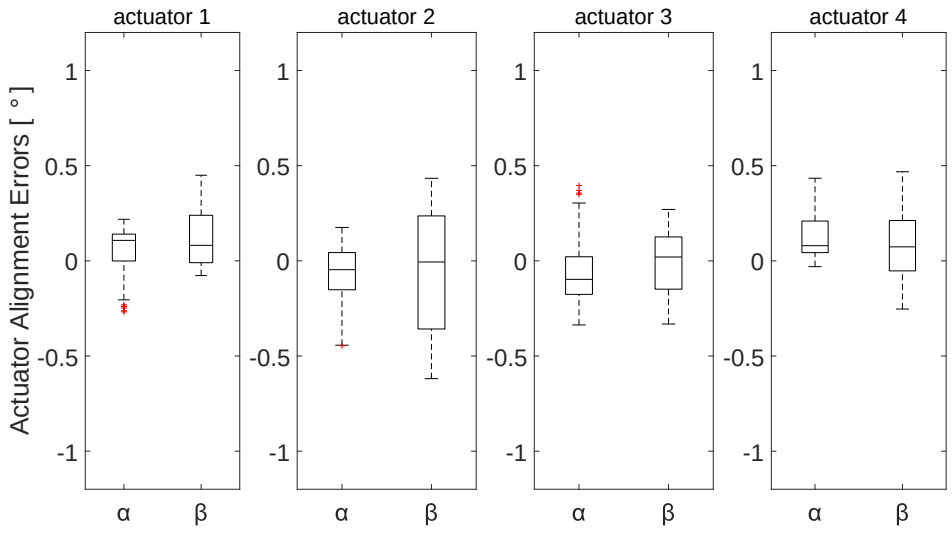


Figure 6.8: Box plot for the alignment estimate error vectors for each of the 4 reaction wheels.

misalignment of the estimates is described as two angles (α, β) that correspond to the alignment correction of (6.5) in order to have a more intuitive representation. There is no evident bias in the inertia estimates as well for the alignments, thanks to the use of the IV method and to the augmentation of the regressor matrix, while the use of the filter allowed to have estimates with a low variance. The delay of the star tracker is correctly estimated, as it is shown in Table 6.1, where $\hat{t}_\delta$ is computed as the average of (6.40) among the three diagonal inertia terms. The overall inertia and actuator alignments estimates are shown in detail in Table 6.2 and 6.3, whereas, differently from the plots, a vector representation is kept for the actuator alignments.

Table 6.1: Star tracker delay results.

true delay (t_{δ_0})	mean ($\hat{t}_{\delta_0}$)	st.d. ($\sigma_{\hat{t}_\delta}$)
$0.125\,s$	$0.120\,s$	$0.002\,s$

Table 6.2: Inertia parameter estimates.

	$J_{11}\,[kgm^2]$		$J_{22}\,[kgm^2]$		$J_{33}\,[kgm^2]$	
	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{J}$	20.209	0.045	23.253	0.055	27.856	0.058
J_0	20.261	-	23.256	-	27.837	-
	$J_{23}\,[kgm^2]$		$J_{13}\,[kgm^2]$		$J_{12}\,[kgm^2]$	
	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{J}$	-0.369	0.018	-2.130	0.096	-1.376	0.009
J_0	-0.374	-	-2.171	-	-1.376	-

In case of satellite affected by a smaller disturbance torque, better results could be obtained since a "stronger" filter (lower cut-off frequency) could be applied to the system equations. Overall, the results are satisfactory, also considering the large number of parameters that are estimated.

Table 6.3: Reaction wheel alignment vector estimates.

	a_1 [-]		a_2 [-]		a_3 [-]		a_4 [-]	
	mean	st.d.	mean	st.d.	mean	st.d.	mean	st.d.
$\hat{a}$	−0.0427	0.0025	0.7293	0.0019	−0.0199	0.0027	−0.6923	0.0016
	0.6975	0.0016	0.6842	0.0020	0.7039	0.0021	0.7216	0.0015
	−0.7153	0.0015	−0.0057	0.0054	0.7100	0.0021	0.0042	0.0030
a_0	−0.0406		0.7302		−0.0201		−0.6940	
	0.6983	-	0.6832	-	0.7045	-	0.7200	-
	−0.7146		−0.0067		0.7094		0.0055	

Table 6.4: Inertia and alignment estimate MSE for the different methods.

	Proposed	Method 1	Method 2	Method 3
$J_{MSE}\,[kg^2m^4]$	0.0312	0.1706	0.2299	0.5084
$a_{MSE}\,[-]$	$1.298 \cdot 10^{-4}$	$6.461 \cdot 10^{-4}$	$8.384 \cdot 10^{-4}$	$36.30 \cdot 10^{-4}$

To further demonstrate the performance of the developed method, and in particular the effectiveness of all the proposed techniques, a comparison is made. We considered three other methods, which follow all the steps presented in Section 6.3-6.4 but one

- Method 1: Instrumental variable with a filter designed only to tackle the noise on the measurements;
- Method 2: Instrumental variable without the estimation of the delay;
- Method 3: Least squares (equivalent to the proposed method, but with no delay for the instrument Z_{aug}, so $Z_{aug} = f_{aug}$).

The results are represented in Figure 6.9 and 6.10, where the proposed method is shown in black, method 1 in blue, method 2 in red, and method 3 in yellow. The proposed approach presents lower bias and variance with respect to the others, showing how all the developed method steps do improve the parameter estimates, especially for the actuator alignments. The comparison with method 2, in particular, shows that a small timing error in the measures can have a non-negligible effect on the estimates. It also shows the effectiveness of the proposed delay compensated method.

A more compact comparison is given in Table 6.4, were the sum of the mean squared errors (MSE) are given for each method. The proposed method has a significantly lower MSE.

In this realistic simulation scenario, the proposed method presents a good performance in both inertia and alignment parameter estimation. Moreover, the several Tailor-made steps and augmentations in the algorithm are well justified as shown by the comparisons.

6.6 Conclusions

In this chapter the estimation problem has been extended to the simultaneous estimation of the inertia and the actuator alignments. To overcome the problem of the unit length constraints of the alignment vectors, an iterative method has been developed, which reduces the number of parameters to be estimated and removes the constraints. The disturbances have been deeply analyzed in order to remove all their deterministic components by means of an augmented regressor. As in Chapter 5, a detailed noise analysis has been performed, allowing for the design

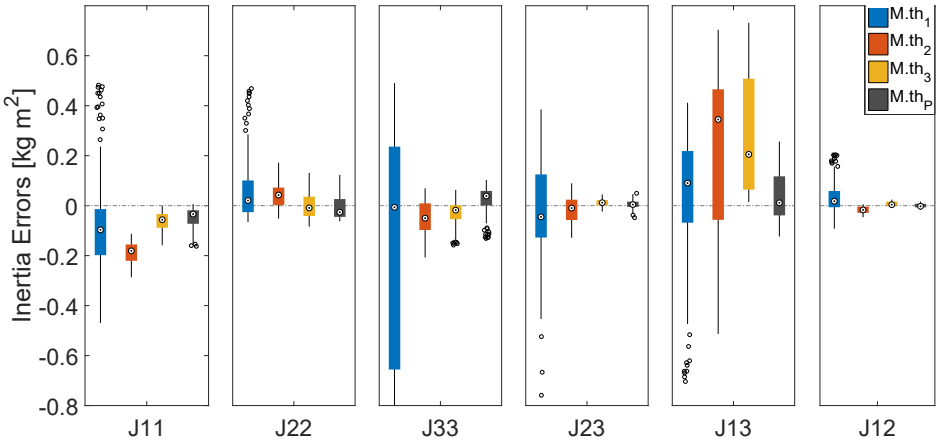


Figure 6.9: Inertia error comparison.

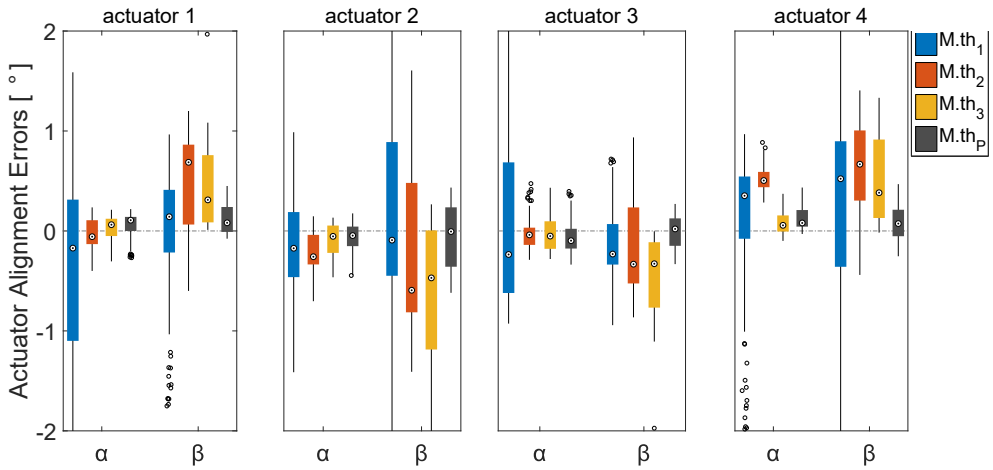


Figure 6.10: Alignment error comparison.

of a prefilter which improves the algorithm estimation performance (low-variance estimates). To further justify this prefilter, also a black-box approach has been tested that estimates the noise ARMA model from the residuals. Both the gray-box and black-box approach for the prefilter design yield similar filters, with a slight difference in the cut-off frequency.

Moreover, the issues due to an incorrect timing of the input signals have been considered and a method has been developed, which allows the estimation the timing error and compensate for its effects on the estimates. All the proposed solutions have been validated through numerical simulations, which use data of a high-fidelity simulator, showing that the developed method has good performances as well that the removal of the timing error improves considerably the estimation accuracy.

Up to this point, the maneuvers for the parameter estimation have been chosen based on some simple rules which in most cases guarantee the persistence of excitation, and the minimization of the variance of the estimates has been obtained with a suitable selection of the prefilter. In the

next chapter we will show that also a suitable choice of the maneuver, by means of experiment design methods, may positively affect the estimation precision.

7

Experiment Design for Satellite Inertia Matrix Estimation

Contents

7.1 Problem Statement 96
7.2 Parametrization Choice for The Maneuver Profile 97
7.3 Experiment Design 98
7.3.1 Fisher Information Matrix Based Experiment Design 98
7.3.2 Noise Model Correction Filter 100
7.3.3 Including Physical Constraints in the Cost Function 101
7.4 Numerical and Experimental Results 102
7.4.1 Numerical Simulations 102
7.4.2 Zero-Gravity Experiment 106
7.5 Conclusions 109

Up to this point the excitation signal used for the experiments has been generated following these simple general rules:

- the signal on the three satellite axes must be excited;
- the excitation of the different components must be linearly independent (e.g. avoid to have the same maneuver on different axes);
- the level of the excitation signal must be as high as possible, compatibly with the system and actuator constraints.

For the satellite inertia estimation it is trivial to generate maneuvers that guarantee persistency of excitation. However, even if the parameters are identifiable, it is well known that the variance of the parameter estimates does also depend on the excitation signal. Therefore, in the whole estimation process, the experiment design, together with the use of an appropriate estimator, is one of the most important steps. The objective of the experiment design is to excite the system in order to obtain data with the highest information content. Several works have studied the optimal excitation problem [Goodwin and Payne, 1977; Armstrong, 1989; Walter and Pronzato, 1990; Franceschini and Macchietto, 2008], however, its application to the satellite parameter estimation is very recent. In [Weiss et al., 2015] the authors propose a receding horizon optimization of the null motion in order to better estimate the reaction wheel alignments.

A different solution for the same problem is described in [Leve and Jah, 2014]. In [Sekhavat et al., 2009] the authors minimize the condition number of the regressor matrix to design the maneuver. Another approach is shown in [Zhai et al., 2017], where a performance index similar to the condition number is used to optimize the maneuver to better estimate the satellite inertia matrix.

Even if the literature on satellite experiment design is quite limited, some related works can be found in the robotic field. Robot dynamics presents indeed many similarities with the rotational dynamics of satellites. Therefore many of these works can be easily adapted to the spacecraft case. In [Gautier and Khalil, 1992], the authors propose an optimization method based on the linear combination of the condition number and the equilibrium of the set of equations that generate the parameters. In [Swevers et al., 1997], a D-optimality criterion on the Fisher information matrix is used. Additionally, instead of optimizing a set of maneuvers points, the maneuvers are represented as a finite Fourier series, thus significantly reducing the number of parameters to be processed. Another similar approach, still based on the finite series Fourier representation, is proposed in [Park, 2006]. In [Calafiore et al., 2001], a genetic algorithm is used to determine excitation trajectories that minimize either the condition number of the regression matrix or the logarithmic determinant of the Fisher information matrix. A different maneuver parametrization is presented in [Rackl et al., 2012], where B-splines are used. The cost function is based on a combination of the condition number and the sum of the joint torques, where the latter term is included in order to improve the signal-to-noise ratio.

In this chapter, we consider the satellite experiment design problem for the inertia parameter estimation. The system under consideration is a small platform that well simulates a satellite behavior. Since this platform is also tested on a parabolic flight [Evain et al., 2019], only very short maneuvers can be performed (for about $5\,s$) adding additional challenges to the maneuver design. The experiment design, and the relative inertia parameter estimation, has been tested both via numerical simulation and through real experiments in a (close to) zero-G environment. This work has been published in [Nainer et al., 2020a].

7.1 Problem Statement

For testing purposes, we consider a satellite-type platform (Figure 7.6). However, this work can be adapted to any satellite since they all have the same system model structure. The platform is equipped with a gyroscope for the angular velocity measurements and it uses 6 control moment gyros (CMGs) to generate the necessary torque for the attitude control. The overall rotational dynamics in the body reference frame is almost identical to the one in (4.2), that is

$$M(t) - \omega(t)^{\times} h_C(t) + M_d(t) = \Big(\Gamma(\dot{\omega}(t)) + \omega(t)^{\times}\Gamma(\omega(t))\Big)\,\theta\,. \tag{7.1}$$

where $M(t)$ is the sum of the torques generated by the CMGs. Like in Chapter 4, by linearly collecting the 6 independent inertia parameters and by rewriting the model as a function of the measured angular rate $\overline{\omega}$ we obtain

$$\underbrace{M(t) - \overline{\omega}(t)^{\times} h_C(t)}_{y(t)} = \underbrace{\Big(\Gamma(\dot{\overline{\omega}}(t)) + \overline{\omega}(t)^{\times}\Gamma(\overline{\omega}(t))\Big)}_{\psi(t)}\,\theta + \nu(t) - M_d\,, \tag{7.2}$$

where the term $\nu(t) \in \mathbb{R}^{3\times 1}$ collects the overall noise effect.

We are now interested in the following problem: *Suppose we have found an unbiased and consistent estimator for θ, find the excitation profile $\omega(t)$ that minimizes the variance of the estimate $\hat{\theta}$.*

Typically, this is a "chicken-and-egg" problem since for the design of optimal maneuvers (to estimate a system) the knowledge of the system itself is required. Anyway, even an approximate knowledge of the system parameters is often sufficient to obtain effective optimized maneuvers.

The experiment design requires three main steps

1. Choose a suitable representation of the reference profile to be optimized,
2. Build a functional to be minimized while including the system constraints[23],
3. Select and apply an optimization solver.

7.2 Parametrization Choice for The Maneuver Profile

Several possible parametrizations exist for the representation of the reference profile $\omega(t)$. The simplest solution consists in optimizing a finite set of equispaced points for the maneuver, and afterwards, interpolate a curve through them. However, the final curve built through interpolation may not necessarily respect all the imposed constraints [Swevers et al., 1997]. This issue can be reduced by increasing the number of points in the optimization, thus increasing the complexity of the cost function. Another approach consists in using a finite Fourier representation as in [Swevers et al., 1997; Park, 2006]. With this parametrization, the number of optimization parameters is significantly reduced and the derivative can be computed analytically, thus without loss of accuracy. However, the Fourier based representation is less flexible with respect to maneuver constraints. For example, in [Park, 2006], in order to guarantee initial and final conditions of the maneuver, a low order polynomial is added to the Fourier representation.

Given the very short maneuver constraint (due to the parabolic flight experiment limitations), a flexible maneuver parametrization should be preferred. For this reason, a cubic B-spline parametrization, as in [Rackl et al., 2012], is used. Splines are piecewise polynomials with pieces that are smoothly connected together. The joining points of the polynomials are called knots. An important property is that, in case of splines with uniform (and unit) knot spacing, we have

$$s(t) = \sum_k c_k \beta^n (t - k)\,. \tag{7.3}$$

This means that a spline $s(t)$ can be expressed as a sum of integer shifts of a B-spline of degree n, denoted as β^n. A B-spline β^n is a symmetrical bell-shaped function obtained from a $(n+1)$-fold convolution of a rectangular pulse β^0. Thanks to the representation defined in (7.3), each spline is defined by its sequence of coefficients c_k, which has the convenient structure of a discrete signal even though the underlying model is continuous. Moreover, B-splines are easy to manipulate. For instance, the derivative can be calculated as

$$\frac{d\beta^n(t)}{dt} = \beta^{n-1}(t + \frac{1}{2}) - \beta^{n-1}(t - \frac{1}{2})\,. \tag{7.4}$$

One more interesting property of the splines is the so-called convex hull property [Schumaker, 2007]. This means that the curve lies in the convex hull defined by its control points, as shown in

[23] Some optimization solvers can also handle constraints. However, due to the complexity of the constraints, we preferred to include them in the cost function as soft constraints.

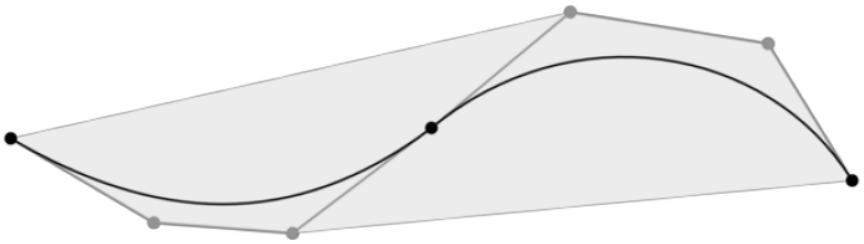


Figure 7.1: Convex hull property of a cubic spline.

Figure 7.1. This property will be exploited in order to impose the respect of the constraints also at intermediate points without interpolating the function values (if the control points respect the constraints, all the curve will do).

7.3 Experiment Design

7.3.1 Fisher Information Matrix Based Experiment Design

The problem of finding effective excitation trajectories is a common problem in the field of parameter identification and different approaches have been presented in the literature. In the design of experiments for estimating a statistical model, optimal design, in combination with an unbiased estimator, allows parameters to be estimated with minimum variance[24]. An optimal excitation design requires a lower number of experimental runs (or shorter runs) to achieve a sufficient estimation accuracy. Intuitively, data with higher information content with respect to the parameters to be estimated should yield more accurate estimates. This concept is represented by the Fisher information matrix [Goodwin and Payne, 1977], which is a measure of the amount of information that a measured variable x provides on the unknown parameter θ. It is formally defined as

$$I(\theta) = E[\nabla_\theta \, log(p(x|\theta)) \nabla_\theta \, log(p(x|\theta))^T] \,, \tag{7.5}$$

or equivalently by

$$I(\theta) = -E[\nabla^2_\theta \, log(p(x|\theta))] \,, \tag{7.6}$$

where $p(x|\theta)$ is the probability of the observed variable x for a given value of θ. In case of a linear regression equation

$$y = \psi \, \theta + e \,, \tag{7.7}$$

where e is assumed to be a vector of independent identically distributed random variables with zero mean and covariance matrix $\sigma^2 I$, it can be easily demonstrated that the Fisher information is given by ([Ljung, 1999], Ch.7 and Ch.13)

$$I(\theta) = \frac{\psi^T \psi}{\sigma^2} \,. \tag{7.8}$$

[24] As in the majority of this work, we considered $\mathcal{S} \in \mathcal{M}(\theta)$ (3.12).

In fact, for a SISO system, considering (7.7), the probability density function of $y - \psi\,\theta$ is

$$p(y - \psi\theta|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\psi\theta)^2}{2\sigma^2}} . \tag{7.9}$$

Taking the logarithm of (7.9) and differentiating with respect to θ we obtain:

$$\nabla_\theta(log(p(y - \psi\theta|\theta)) = \frac{y\psi^T}{\sigma^2} - \frac{\psi\theta\psi^T}{\sigma^2} . \tag{7.10}$$

Taking another derivative, the expectation of the Hessian matrix, which is the negative of the Fisher information, is

$$E(\nabla^2_\theta(log(p(y - \psi\theta|\theta))) = -\frac{\psi^T\psi}{\sigma^2} . \tag{7.11}$$

In case of N measures, the values of $y(t_i)$ and $\psi(t_i)$ can be stacked respectively into a vector Y and a matrix Ψ as

$$Y = \begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_N) \end{bmatrix} , \qquad \Psi = \begin{bmatrix} \psi(t_1) \\ \psi(t_2) \\ \vdots \\ \psi(t_N) \end{bmatrix} . \tag{7.12}$$

Since the Fisher information matrix associated to a set of independent samples is equal to the sum of the individual Fisher information matrices, we have

$$I(\theta) = \frac{\Psi^T\Psi}{\sigma^2} . \tag{7.13}$$

The Fisher information matrix is related to the Cramer-Rao lower bound [Kay, 1998; Young, 2011]. The Cramer-Rao inequality states that, for any unbiased linear estimator for θ, we have

$$Var(\hat{\theta}) \geq I(\theta)^{-1} . \tag{7.14}$$

The maximization of the Fisher information corresponds therefore to the minimization of the lower bound of the variance of the estimates.

The least squares (LS) solution

$$\hat{\theta}_{LS} = (\Psi^T\Psi)^{-1}\Psi^T Y \tag{7.15}$$

has variance

$$Var(\hat{\theta}) = \sigma^2(\Psi^T\Psi)^{-1} , \tag{7.16}$$

which is exactly the inverse of the Fisher information.

In the satellite application, the regressor $\psi(\overline{\omega})$ is correlated with the noise ν, and therefore a least squares method would yield biased estimates [Jun et al., 2010]. To obtain consistent parameter estimates, as shown in Chapter 4-5-6, an instrumental variable (IV) method can be used [Söderström and Stoica, 1983; Young, 2011]. For the basic IV, the analyical solution (repeated here for sake of clarity), similar to the least squares one, is given by

$$\hat{\theta} = (Z^T\Psi)^{-1}Z^T Y , \tag{7.17}$$

where Z is an instrument, that should be correlated with the regressor Ψ and uncorrelated with the noise ν. Given (7.7), the IV parameter variance is

$$Var(\hat{\theta}_{IV}) = \sigma^2(Z^T\Psi)^{-1}Z^T Z(\Psi^T Z)^{-1} , \tag{7.18}$$

which is similar to (7.16). Considering that the optimal instrument Z is the noise-free version of Ψ, optimizing a functional based on (7.15) and (7.16) is effective also if the IV method is used. In fact, (7.18) can be well approximated as

$$Var(\hat{\theta}_{IV}) \simeq \sigma^2 (Z^T Z)^{-1} . \tag{7.19}$$

and, since the noise-free signal is used for the optimization, using (7.19) or (7.16) is equivalent. For a detailed description of the IV approach the reader may refer to the Section 3.3.

As a result, in the simple regression model (7.7), an optimal experiment design should maximize (7.13) or almost equivalently minimize (7.16).

Most of the design strategies that can be found in the literature have as objective the minimization of some functional $\rho(Var(\theta))$. Among them, the most common criteria are [Franceschini and Macchietto, 2008; Walter and Pronzato, 1990]:

- **D-optimality**

  This optimality criterion requires to maximize the determinant of $\Psi^T\Psi$. Since the determinant is the product of the eigenvalues, which is inversely proportional to the product of the axes of the confidence ellipsoid, maximizing the determinant is equivalent to minimize the volume of the confidence ellipsoid of $\hat{\theta}$.

- **E-optimality**

  It requires the maximization of the smallest eigenvalue of the Fisher information matrix, equivalent to maximizing the minimum eigenvalue of $\Psi^T\Psi$. Geometrically, the E-optimal design minimizes the maximum diameter of the asymptotic confidence ellipsoids of $\hat{\theta}$.

- **A-optimality**

  This criterion seeks to minimize the trace of $(\Psi^T\Psi)^{-1}$. It results in minimizing the average variance of the estimate $\hat{\theta}$.

- **Condition Number Minimization**

  This optimality criterion minimizes the condition number for inversion of the matrix $\Psi^T\Psi$, thus reducing the sensitivity of the estimate to perturbations in the input data and to round-off errors made during the solution process. It should be noted that the condition number is not influenced by any scaling factor on the regressor Ψ. Therefore, the cost function should also include a penalty term in order to improve the signal to noise ratio.

7.3.2 Noise Model Correction Filter

Equation (7.16) assumes that $e(t)$ in (7.7) is a white noise. However, in the satellite model (7.2), the noise term $\nu(t)$ does not respect this assumption. In fact, even if nonlinear, ν can be well approximated as

$$\nu \simeq (\gamma s + 1)\, e\,, \tag{7.20}$$

where e is a white noise source, γ is a scalar real positive value and s is the differential operator. This approximation, also used in Chapter 4, has been shown in detail in Chapter 5, where a similar noise model has been obtained (5.37). The additional differential operator s in (5.37) with respect to the new model (7.20) is simply due to the use of attitude measurements instead of angular velocity measurements. We remind the reader, that this unusual improper transfer function comes from the use of an inverse model.

Many of the papers cited in the introduction, that deal with satellite and robot experiment design, did not consider that the white noise assumption was not fulfilled. This comes, again, from the use of the inverse model to obtain equations in linear regression form. In this case, even if the sensor is affected by an ideal white noise, the overall noise term ν presents a high-frequency amplification behavior.

Therefore, by not respecting the white noise assumption, optimizing a functional based on (7.16) does not generate an optimal maneuver. A solution consists in prefiltering (7.2) with the inverse of the noise transfer function in order to respect the hypotheses in (7.7). A low-pass filter

$$F(s) = \frac{1}{\gamma\, s + 1} \tag{7.21}$$

is therefore applied to both sides of (7.2) to come up with

$$y_f(t) = \psi_f(t)\theta + \nu_f(t)\,, \tag{7.22}$$

where $(\cdot)_f$ means that the signal is filtered by $F(s)$. This approach is very similar to what is done in the optimal IV method [Young, 2011] in order to have white residuals. For this filtered equation, the assumption of white noise ($\nu_f(t) \sim N(0, \sigma^2)$) is "approximately" fulfilled. Therefore, the scheme described in the previous subsection can now be applied to (7.22) and the variance to be minimized becomes

$$Var(\hat{\theta}) = \sigma^2(\Psi_f^T \Psi_f)^{-1}\,. \tag{7.23}$$

7.3.3 Including Physical Constraints in the Cost Function

Considering Section 7.1.1 and 7.1.2, the optimization aims at minimizing a functional $\rho(\cdot)$ based on one of the possible criteria. For example, using the D-optimality criterion, the following cost function can be formulated

$$\rho_D = -log\Big(det\big(\Psi_f(\omega, \dot{\omega})^T \Psi_f(\omega, \dot{\omega})\big)\Big)\,. \tag{7.24}$$

However, in addition to the minimization of the cost function, an optimal trajectory should also respect some physical constraints (e.g. saturation of the platform actuators). The easiest solution consists in treating the system constraints as soft constraints, and therefore adding some penalty terms to the functional to be minimized that reach high values when close to the saturation/physical limits. Three types of constraints are considered: torque saturation, angular momentum saturation, and initial and final conditions of the maneuver. However, for more complex scenarios, other constraints could be added the same way, such as constraints on the satellite attitude in order to avoid the blinding of the star tracker.

First, the actuator constraints should be converted into state constraints

$$|M_C(t)| \leq M_{max} \quad \Rightarrow \quad |\dot{\omega}(t)| \leq \dot{\omega}_{max}\,, \tag{7.25}$$

$$|h_C(t)| \leq h_{max} \quad \Rightarrow \quad |\omega(t)| \leq \omega_{max}\,. \tag{7.26}$$

In this work a limit has been set for each of the components of the angular rate ω and of the angular acceleration $\dot{\omega}$. Moreover, some small safety margins are considered in the conversions $M_{max} \rightarrow \dot{\omega}_{max}$ and $h_{max} \rightarrow \omega_{max}$.

Once the constraints are defined on ω and $\dot{\omega}$, the full cost function to be minimized, making use of (7.23), (7.25), (7.26), has the following form

$$\omega_{ref} = \arg\min_{\omega} \Big(\rho\Big(\big(\Psi_f(\omega, \dot{\omega})^T \Psi_f(\omega, \dot{\omega}) \big) \Big) + \text{penalty} \Big) \tag{7.27}$$

where $\rho(\cdot)$ is a functional that depends on the chosen optimization criterion and the penalty, containing the physical limitation and the initial and final conditions of the maneuver, is defined as

$$\begin{aligned} \text{penalty} = {} & K_1(\omega(t_1) - \omega_{initial})^6 + K_1(\omega(t_N) - \omega_{final})^6 + \\ & + K_2 \sum \left[Q(\omega(t_k), \omega_{max}) + Q(-\omega(t_k), -\omega_{max}) \right] + \\ & + K_3 \sum \left[Q(\dot{\omega}(t_k), \dot{\omega}_{max}) + Q(-\dot{\omega}(t_k), -\dot{\omega}_{max}) \right], \end{aligned} \tag{7.28}$$

with

$$Q(x, threshold) = \begin{cases} 0 & x \leq threshold \\ (x - threshold)^6 & x > threshold \end{cases} . \tag{7.29}$$

and where K_1, K_2 and K_3 are some constants to weigh the different penalty terms. The summation in (7.28) should be extended to a thick grid of points in order to ensure that the constraints are respected for the whole signal.

A simpler, but effective approach is to exploit the spline convex hull property and extend the summation for the penalty only to the spline control points $\omega_{cp,k}$. This leads to a slightly more conservative constraints.

This nonconvex optimization problem can be solved using one of the available optimization tools. For this work we used the optimization toolbox from Matlab, and in particular the *fminunc* routine.

7.4 Numerical and Experimental Results

To illustrate the effectiveness of the maneuver design, the method is firstly tested via a Monte Carlo simulation. Then, a real experiment is performed on a parabolic flight, where the platform is commanded with the optimized maneuver in a zero-G environment typical of a satellite.

7.4.1 Numerical Simulations

For the numerical simulations, the platform-satellite inertia J was set to

$$\begin{bmatrix} 0.022 & 0 & 0 \\ 0 & 0.022 & 0 \\ 0 & 0 & 0.027 \end{bmatrix} [kg\, m^2], \tag{7.30}$$

and the CMG cluster saturation limits have been set to

$$\begin{cases} M_{C,x} \leq 0.040\, Nm \\ M_{C,y} \leq 0.040\, Nm \\ M_{C,z} \leq 0.047\, Nm \end{cases} \qquad \begin{cases} h_{C,x} \leq 0.0210\, Nms \\ h_{C,y} \leq 0.0210\, Nms \\ h_{C,z} \leq 0.0254\, Nms \end{cases} . \tag{7.31}$$

To take some safety margins to avoid saturation, the maximum angular velocity and acceleration for the optimized maneuver have been set to $|\omega_{max}| \simeq 0.3\, rad/s$ and $|\dot{\omega}_{max}| \simeq 0.9 rad/s^2$. The

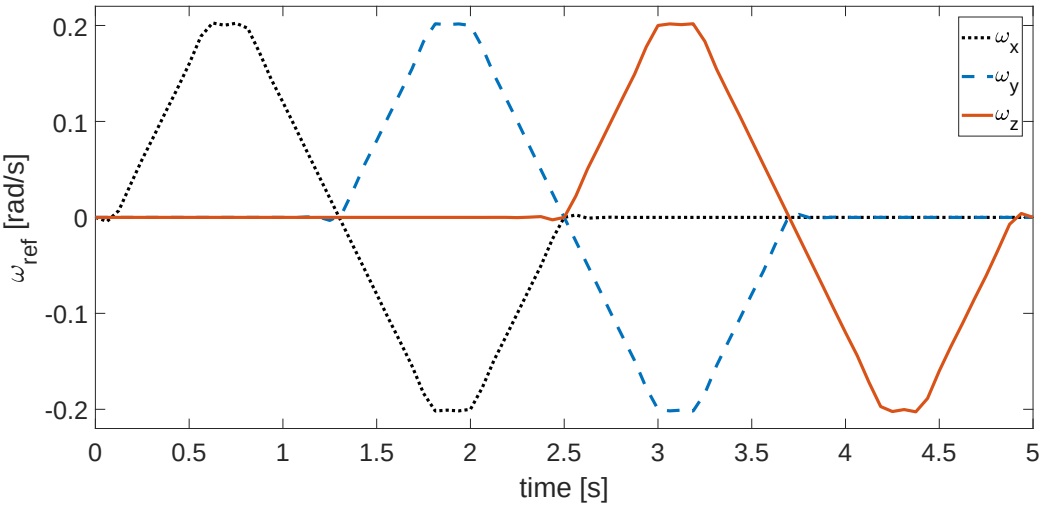


Figure 7.2: Initial reference profile used for the optimization.

	minimum eigenvalue ↑	trace ↓	determinant ↑	condition number ↓
basic	0.099	21.27	0.0036	3.271
A-opt	1.663	**2.431**	330.6	1.546
D-opt	1.683	2.434	**331.3**	1.568
E-opt	**1.746**	2.528	253.8	1.465
Cond. N° - opt	1.532	2.883	115.3	**1.426**

Table 7.1: Main parameters for different optimality criteria. The arrows indicate whether a large ↑ or small ↓ value is better.

initial and final angular velocity for the maneuver were set equal to zero, while the experiment duration was set to $5\,s$. The maneuvers are represented by cubic splines with a grid step of $0.1\,s$. Figure 7.2 shows the basic reference profile used as initial maneuver for the optimization algorithm. The choice of this initial profile was based on the two main conditions in order to avoid close to singular regressor: all three axes should be excited and the maneuvers should not be the same on different axes. The prefilter was set as

$$F(s) = \frac{1}{\gamma\, s + 1}\,, \tag{7.32}$$

where $\gamma = 0.8$. With a posteriori check we have verified that, with this value of γ, the error ν_f becomes reasonably white. Firstly, an optimized maneuver has been computed for each one of the four criteria of Section 7.3.1. The main criterion parameters of the differently generated maneuvers are shown in Table 7.1, where it is clearly visible how the optimized maneuvers have better "criterion" values with respect to the basic profile. On the other hand, for this particular application, the different optimization criteria do not yield significantly different maneuver "properties". For example, it may be observed that optimizing the determinant we obtain also a good value of minimum eigenvalue and of the trace.

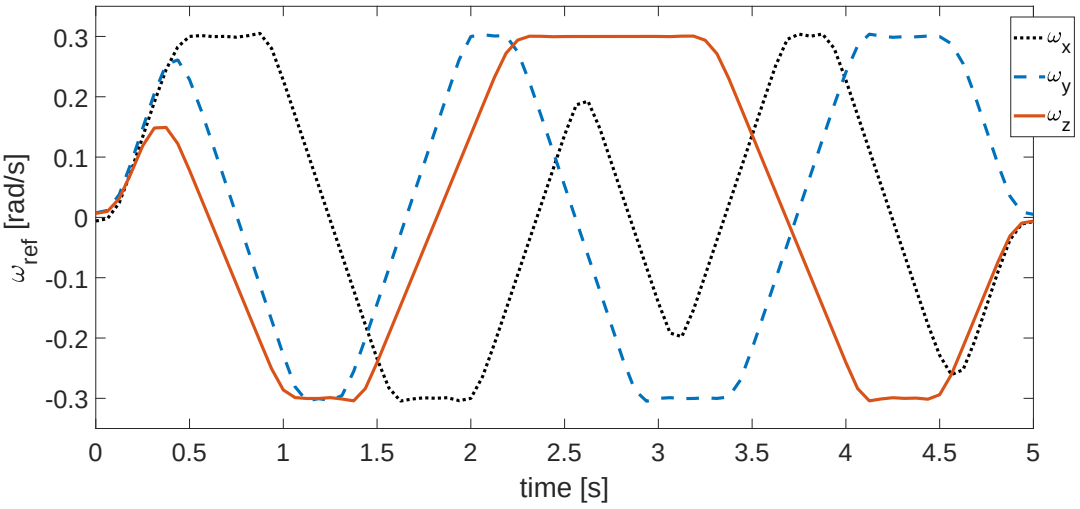


Figure 7.3: D-opt maneuver.

The practical advantage of the experiment design has been studied via a Monte Carlo simulations, where the satellite inertia has been estimated both from the basic and the optimized reference profiles. Additionally, also a scaled version of the basic profile is used in order to have the same maximum angular velocity ω amplitude compared to the optimized maneuver ($|\omega_{max}| \simeq 0.3\, rad/s$). This "scaled" maneuver (ω^*_{ref}) is simply the basic maneuver of Figure 7.2 multiplied by 1.5 ($\omega^*_{ref} = 1.5 \cdot \omega_{ref}$).

Since there is very little difference among the different optimization criteria, only the D-opt maneuver has been used (Figure 7.3) in this comparison. The noise standard deviation on the angular velocity measurement was set as $\sigma = 0.025\, rad/s$ for each of the three axes. As mentioned in Chapter 4-5-6, since the regressor $\psi(\overline{\omega})$ is correlated with the noise ν [Jun et al., 2010], an IV method has been used to obtain consistent estimates of the inertia parameters. Given the high similarity with the estimation problem of Chapter 4, the same IV implementation has been used with the main but fundamental difference that also the filter initial conditions have been estimated (similarly to Chapter 5). In fact, given the very short experiment duration, if wrong initial conditions are set, the filter transient has an important impact on the accuracy of the estimates [Garnier et al., 2003], making the overall algorithm unreliable.

The results from a Monte Carlo simulation of 100 runs are shown in Table 7.2 and in Figure 7.5. The increase of accuracy with the optimized reference profile is significant as it can be observed from the much smaller standard deviation in all the parameter estimates. The use of optimized maneuvers is therefore fully justified. Moreover, as shown in Figure 7.4, for the optimized maneuver, the actuator speeds are consistently within their operation range ($|\Omega_{max}| = 5\, rad/s$)for the majority of the simulation experiment.

		J_{11}	J_{22}	J_{33}	J_{23}	J_{13}	J_{12}
basic	avg.	22.24	22.33	27.30	0.14	0.12	0.15
	st.d.	1.12	1.67	1.30	1.27	1.01	1.20
scaled	avg.	22.14	22.15	27.15	0.05	0.05	0.05
	st.d.	0.67	0.96	0.76	0.74	0.60	0.69
D-opt	avg.	22.12	22.11	27.02	0.04	0.10	0.01
	st.d.	0.38	0.33	0.35	0.23	0.24	0.28
true		22.00	22.00	27.00	0.00	0.00	0.00

Table 7.2: Inertia estimates from the different reference profiles $[10^{-3} kg\, m^2]$

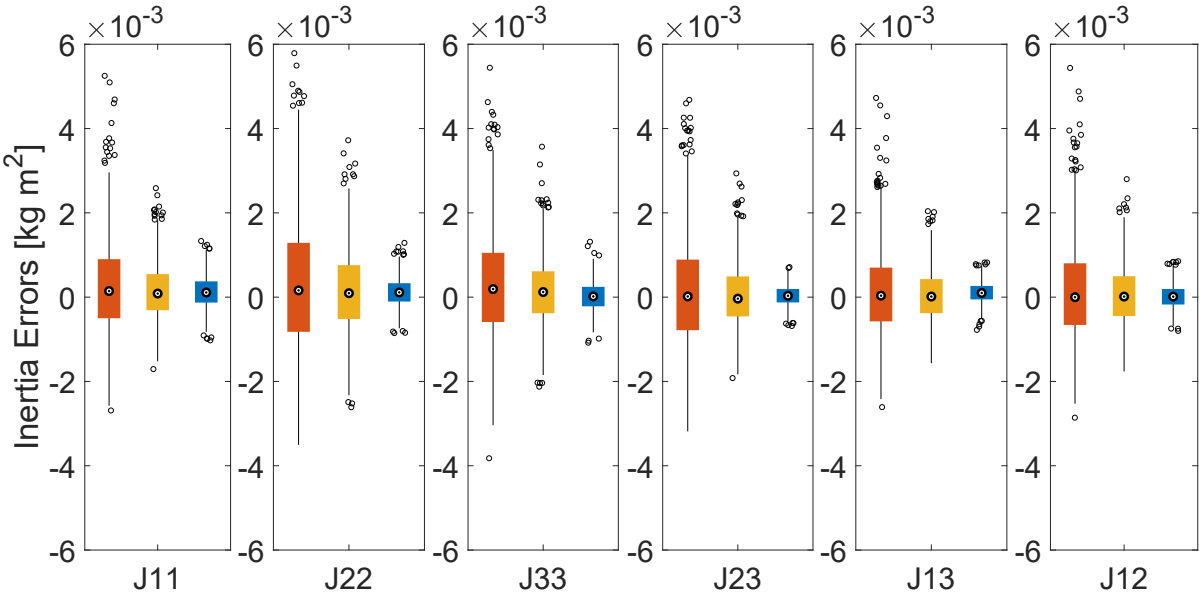


Figure 7.4: Box plot for the inertia estimate errors. In red the results from the basic profile, in yellow the ones from the scaled basic profile, and in blue the ones of the D-opt profile.

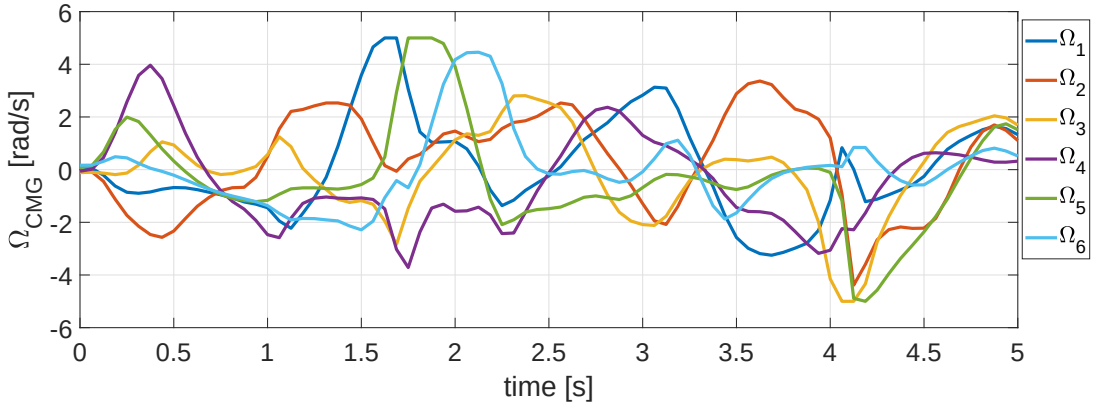


Figure 7.5: CMG speeds obtained from the D-opt maneuver.

7.4.2 Zero-Gravity Experiment

Experiments have been performed on a real platform developed by ISAE-SUPAERO, ONERA and CNES (Figure 7.6). The platform, named SCRAT-0g, has been tested on a parabolic flight that simulated a close to zero gravity environment. The purpose of this experiment was to test a satellite attitude controller on a platform which simulates a real satellite. The test has been repeated a few times, collecting the values of angular rate (from the gyroscope) and actuator angular momentum (generated by the 6 CMGs) at a sample rate of $16\,Hz$. The gyroscope white noise had a standard deviation of $\sigma = 0.095°/s$ on each axis. The experiment was in free floating, therefore, after the platform release, the unavoidable relative motion with respect to the airplane caused premature collisions with the inner walls. Since the final part of the maneuver was unusable, it was necessary to truncate the data. For the inertia estimation, 5 experiment data have been used, with a duration that varied from $2\,s$ up to around $4\,s$. The input-output data of the longest achieved experiment is shown in Figure 7.7. Since the platform tracking was not very accurate, the angular velocity ω does not perfectly match the commanded reference optimized profile. Similarly to the simulation case, the inertia has been estimated by an IV method with a prefilter to deal with the derivatives in the system equation and their effect on the noise.



Figure 7.6: A picture of the SCRAT-0g platform tested on a parabolic flight (courtesy of ISAE SUPAERO and ONERA).

In order to improve the estimation accuracy, the final inertia has been computed as a weighted average of the 5 different estimates, where the weights were given by the covariance of the

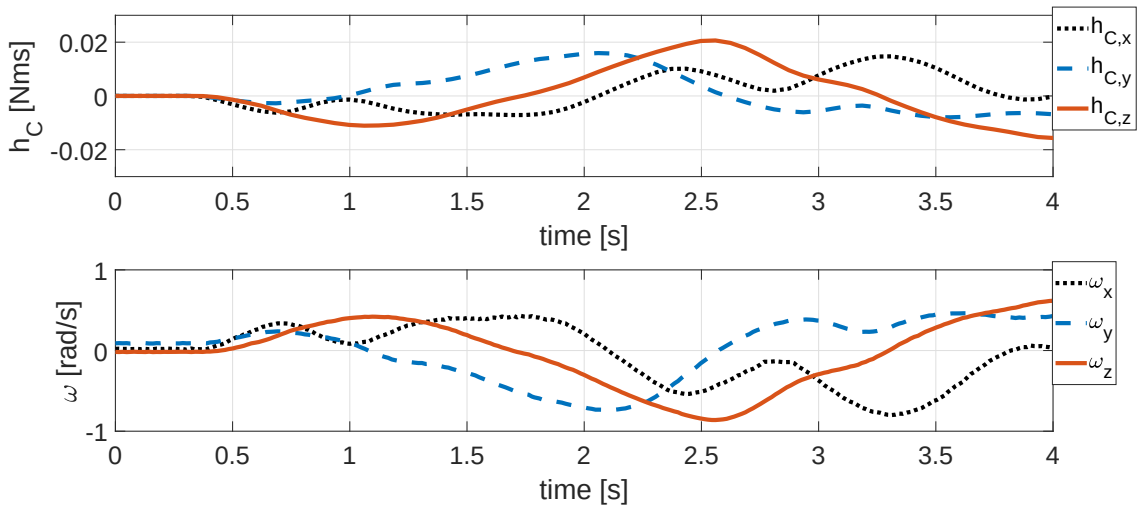


Figure 7.7: Input-Ouput data from the zero-G experiment.

residuals ([Ljung, 1999], Ch.14):

$$\hat{\theta} = P \sum_{i}^{n} (P_i)^{-1} \hat{\theta}_i \,, \tag{7.33}$$

where $\hat{\theta}_i$ is the i-th inertia estimate and P_i is its corresponding covariance, and where

$$P = \left(\sum_{i}^{n} (P_i)^{-1} \right)^{-1} . \tag{7.34}$$

The final result is shown in Table 7.3, where it is compared to the inertia from the approximated CAD (computer-aided design) model.

	J_{11}	J_{22}	J_{33}	J_{23}	J_{13}	J_{12}
$\hat{\theta}$ $[10^{-3} kg\, m^2]$	20.04	21.26	25.75	0.09	0.04	1.08
θ_{CAD} $[10^{-3} kg\, m^2]$	20.80	20.80	26.00	0	0	0

Table 7.3: Inertia Estimate from zero-G test

In this case we do not have a true value for comparison as for the simulations, but the inertia estimate is reasonably close to the nominal CAD values[25]. Introducing the estimated parameters in (7.2), it is possible to do a cross validation by comparing the total torque ($M - \omega^{\times} h_C$) with its effect on the system ($\psi\,\hat{\theta}$). This comparison is depicted in Figure 7.8, which shows a good match among the signals. To have a better way to evaluate these experiment results, Figure 7.9 shows the torque cross-validation[26] for a realization of the numerical simulation of the previous subsection (still using the D-opt reference profile). The fit in the torque cross-validation for the real experiment is not very different from the numerical simulation case, further illustrating the

[25] Ideally, the inertia values for the simulation and for the CAD model should be the same. However, last minute changes on the platform resulted in a slight difference between the two.

[26] In this case, the noise-free ω is available and it can be directly used to build ψ.

accurate inertia estimation. The slightly worse match is justified by the presence of disturbance torques and other nonidealities in the zero-G experiment.

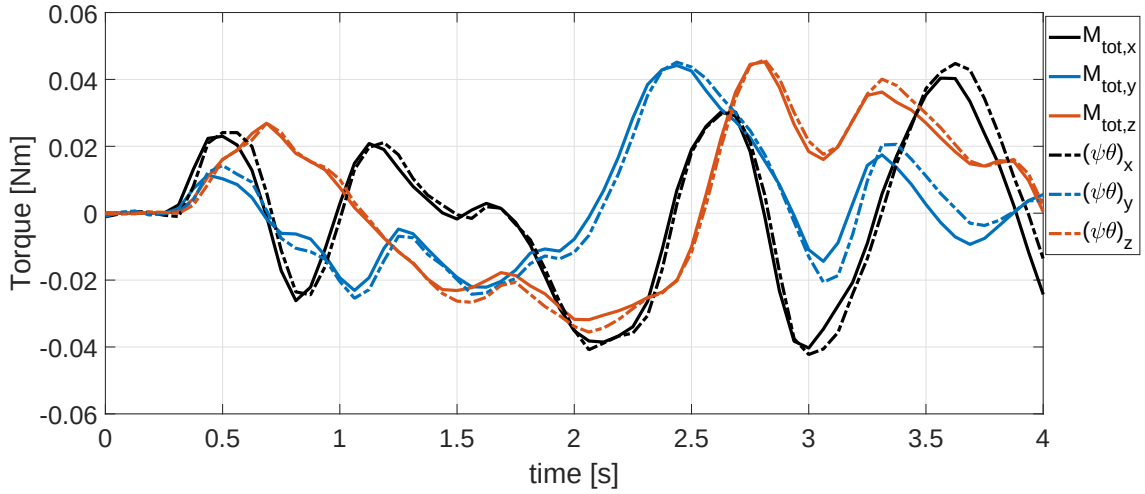


Figure 7.8: Torque cross-validation for the zero-G experiment: measured (solid lines), one-step ahead prediction (dash-dot line).

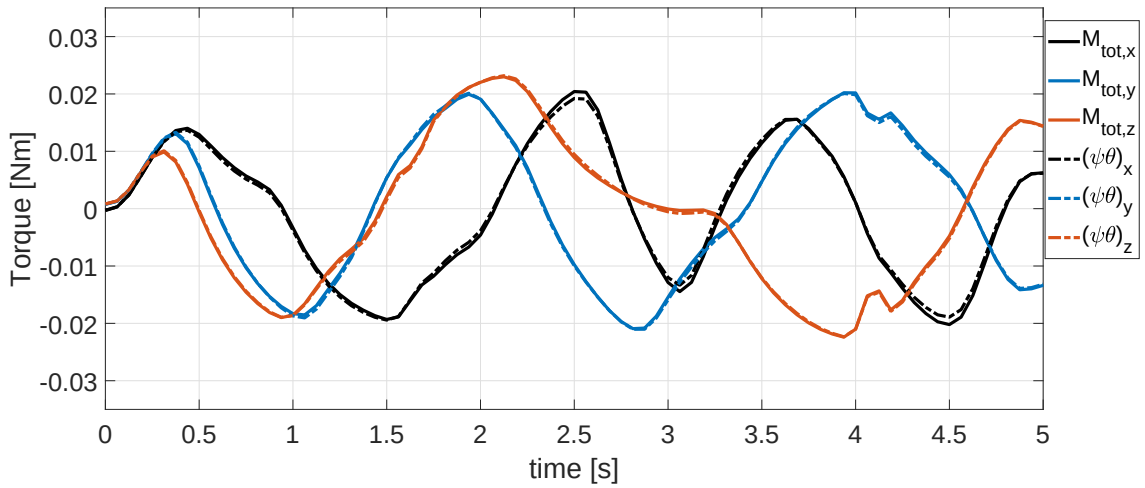


Figure 7.9: Torque cross-validation for the numerical simulation: measured (solid lines), one-step ahead prediction (dash-dot line).

Both the experiment design and the inertia estimate can be considered satisfactory, not just in the numerical results, but also in the real experiment on the parabolic flight. The results are rather impressive, despite the very short duration of the experiments and the disturbance torques affecting the platform.

7.5 Conclusions

In this work, we addressed the problem of finding optimal excitation maneuvers for the estimation of satellite inertia parameters. A cubic B-spline signal representation was introduced and a constrained nonlinear optimization was formulated, based on a Fisher information cost function minimization. The performance and practicability of the maneuvers obtained by the proposed approach have been tested both through numerical simulations and in a real experiment on a parabolic flight. The estimation results have illustrated the effectiveness of the proposed optimized maneuver, with positive effects in the real zero-G experiment, where good estimates have been obtained, especially considering the very small dataset.

Even if, given the prototype constraints, only very short maneuvers have been considered, the extension to the typical satellite scenarios which allows for longer maneuvers is quite straightforward. The maneuver length can be directly changed in the optimization function as well as the actuator constraints, while, given the typical slower dynamics of satellite with respect to SCRAT-0g, the sampling time can be increased thus maintaining the overall computational complexity almost unchanged. If really long maneuvers must be performed, a different and more compact signal representation must be chosen to limit the optimization computational complexity, based for example on a finite Fourier series.

8

Conclusions and Future Works

Contents

8.1 Contribution Overview 111
8.1.1 Analysis, Development and Design of a Tailor-made IV Method for Satellite Parameter Estimation 111
8.1.2 Satellite Inertia Estimation for Gyroless Satellites 112
8.1.3 Extension to the Actuator Alignment Estimation 112
8.1.4 Design of Maneuvers for Satellite Inertia Estimation 113
8.2 User Parameter Guide 113
8.3 Future Works 114
8.3.1 Recursive Parameter Estimation 114
8.3.2 Estimation with Attitude and Angular Velocity Sensors 114
8.3.3 Null Space Experiment Design 114

In this final chapter, we give an overview of the contributions related to the satellite parameter estimation problem. Additionally, possible future works, especially regarding the extension of the proposed methods, are suggested.

Although this thesis is focused on the satellite application, many of the analyses here presented can be extended to other fields of application. In particular, the noise analysis and the IV tailored solution can be extended to other estimation problems that use an inverse model of the real system (e.g. robot dynamics estimation).

8.1 Contribution Overview

In the following subsections, the main contributions of this work are briefly described. They are presented in the same order as they appear in the thesis, starting from Chapter 3 and 4.

8.1.1 Analysis, Development and Design of a Tailor-made IV Method for Satellite Parameter Estimation

In Chapter 3, it is shown how by using the inverse system model of the satellite a useful model in linear regression form can be obtained. However, it is also highlighted how this system inversion transforms the original model into an EIV model, thus highly limiting the choice of estimation methods that are capable to yield unbiased parameter estimates.

In Chapter 4, the IV method is proposed for the inertia parameter estimation and it is shown, both theoretically and by numerical simulations, how the use of the IV can overcome the least square limitation and obtain unbiased results. The closed-loop nature of the system, typically an issue in identification, is exploited to build a closed-loop auxiliary model that is used to generate the noise-free estimates to build the instrument. Additionally, a first step towards a grey-box noise model for the inverse model is made. From a practical perspective, a SVF filter is applied in order to compute the state time derivatives as well as to work as a prefilter for the IV method. Finally, the inertia estimation problem from biased gyroscope measurements has been considered, and the gyroscope bias effect on the estimates has been analyzed. A modified IV estimation algorithm has been proposed, which is able to estimate and remove the effect of measurement bias.

8.1.2 Satellite Inertia Estimation for Gyroless Satellites

In Chapter 5, the gyroless case is studied. Firstly, since the SVF filter is no more applicable, alternative methods are considered to deal, in part, with the noise amplification due to the time derivatives. Secondly, a detailed noise analysis is performed that studies the effect of the noise on the attitude measurements on the full inverse model. The result of this analysis, together with an approximated model of the disturbance torques, are used to the design of a sub-optimal filter for the IV method. Additionally, the model and the respective estimation algorithm are augmented in order to estimate and compensate the filter initial condition and to remove some of the deterministic components of the disturbance torque. The estimation of the initial condition of the prefilter allows for a better use of the telemetry data since the initial filter transient is estimated and removed, avoiding the need of discarding the initial data. This has a significant importance for short experiments, as the one described in in Chapter 7, but it is also important for standard experiments, where the greater availability of data still has some benefits on the estimation accuracy.

The good performance the algorithm is shown through the numerical simulation as well by using real telemetry data from the Picard CNES satellite.

8.1.3 Extension to the Actuator Alignment Estimation

In Chapter 6, the identification problem is extended to the actuator alignment estimation. An iterative algorithm has been implemented in order to avoid dealing with the quadratic constraints implied by the alignment vectors. Additionally, a non-perfect synchronization among the sensors is considered. As far as the author knows, this issue has never been considered in the literature. However, as shown by the simulations, even a small syncronization error may have non negligible negative effects on the estimation accuracy. A method, which uses a double regressor, has been developed to estimate the time shift of the data and remove its effect. In this IV implementation, a delay regressor has been used as instrument, thus no information about the closed-loop system and in particular the satellite controller are needed. In this part of the work an alternative method has been developed and tested for the design of the prefilter. The method exploits the cepstrum properties to obtain an approximation of the filter impulse response and it seems well suited for the type of CARMA noise model of this application. Again, the algorithm is tested on data generated by a high-fidelity simulator from CNES in order to validate the performance in this new case, as well as to justify all the "extensions" of the method.

8.1.4 Design of Maneuvers for Satellite Inertia Estimation

The final part of this work (Chapter 7) shifts the focus from the estimation algorithm to the experiment design. The objective was to design maneuvers that result in telemetry data with higher information content. In this particular case, the maneuver design was focused on a small real prototype that has been also tested on a parabolic flight. In order to design optimized trajectories, a B-spline parameterization is chosen for the reference profile, and then the maneuver is computed by minimizing a functional based on the Fisher information matrix. The optimization algorithm considers the very peculiar noise model resulted by the inversion of the satellite model. Different optimality criteria are tested, and the optimized maneuvers are compared with a typical testing maneuver, showing how the proposed experiment design results in drastically more precise inertia estimates. Finally, the maneuver generated by the experiment design has been tested on the real prototype in close to zero-G conditions. Also in the real experiment, accurate inertia estimates have been obtained, especially considering the extremely short data sets.

8.2 User Parameter Guide

In this thesis a tailored instrumental variable method has been used for the inertia parameter estimation. However, depending on the specific application, different approaches have been implemented as well as different way to filter the model equations. Hereafter is a brief guide for a correct use of the algorithms:

- **Instrument "construction":** Two main approaches have been used in this work. The first consists in building and running a closed-loop auxiliary model to obtain noise-free signal estimates, while the second approach uses a delayed regressor as instrument. For the latter case, the choice of the delay plays an important role. As already mentioned, a large delay reduces the correlation between the instrument and the regressor and it results in a higher standard deviation, while a too small delay may not be sufficient to guarantee the uncorrelation property with the noise term, determining unconsistent estimates. However, in this satellite case, the tuning of the delay is quite straightforward since, considering the presence of first and second order derivatives in the noise term, a delay of a few samples is sufficient to obtain a negligible correlation between the instrument and the regressor. In this scenario, a delay between 4 and 6 sample is a good trade-off to obtain consistent estimates without increasing significantly the variance. The short delay, of 4 samples is suggested in case of fast maneuvers with respect to the sampling frequency (like in Chapter 7), while a delay of 6 sample can be used for less agile maneuvers (like in Chapter 6). Nonetheless, if the computational complexity is not an issue and the controller is known, the use of the closed-loop auxiliary model is suggested.

- **User-parameter tuning:** Despite the overall IV implementation may seem complex (like in Chapter 5 and 6), there are only two main parameters that need to be chosen: the sensor noise standard deviation and the disturbance torque standard deviation. This presents many similarities with the EKF parameter estimation approach that makes use of the noise covariance matrix (often named R) and the process covariance matrix (often named Q) as user parameters. The sensor noise standard deviation can be directly taken from the datasheet, while, as for the Q matrix in the EKF, the disturbance standard deviation tuning is more difficult. Since the estimation is performed offline, an analysis of the residuals can

be performed to guide the choice of the filter, as done in Chapter 6. In case of long experiments, the whole filter design can be based on the ARMA model obtained from the residuals. In this case, we can still exploit the grey-box modeling of the noise by imposing a $-40\,dB/decade$ at the high frequencies.

8.3 Future Works

8.3.1 Recursive Parameter Estimation

A recursive implementation of the IV method would be an interesting additional development, allowing for an online parameter estimation. Since the IV method has a cost function closely related to the LS, there are already tools that can be applied to transform the IV method into a recursive algorithm (i.e. the QR decomposition may be a valid technique to solve this problem). One of the main points that needs to be tackled is the generation of the instrument. The use of an auxiliary model to generate the instrument is well suitable for offline estimation, and its implementation to an online algorithm can be difficult due to its computational complexity. However, there are other solutions that can be applied without increasing the overall computational complexity. The simplest solution consists in using a delayed regressor as instrument, as done in Chapter 6. If instead the satellite tracking performance is very precise ($\omega_{ref} \simeq \omega$), the instrument may use directly as noise-free signals the reference signals given to the controller [Brunot and Janot, 2018].

8.3.2 Estimation with Attitude and Angular Velocity Sensors

In this thesis we considered the estimation from gyroscope measurements or from star tracker measurements. If the accuracy of the two sensors is similar, they can be both used in order to obtain better parameter estimates. Given the very different characteristics of the two sensors, a complementary filter can be used to merge the measurements [Higgins, 1975; Mahony et al., 2008] as an alternative to the Kalman filter. The low-pass filter should be applied to the angular velocity computed from the attitude measurements, since at low-frequencies the differentiation amplifies only slightly the noise, while the measurements from the gyroscope would be filtered be a high-pass filter. An additional advantage of this approach is that the effect of the bias from the gyroscope sensor, being at low frequency, is directly removed.

8.3.3 Null Space Experiment Design

In Chapter 7, satellite maneuvers have been optimized in order to improve the inertia estimation. However, the same can be done for the actuator alignment estimation. In this case, also a null-space excitation does improve the Fisher information matrix. In [Leve and Jah, 2014] and [Weiss et al., 2015], the authors tried to exploit the null space, however they only considered the experiment design for the actuator alignment. In this scenario, for a satellite with 4 reaction wheels, the null space consists in a 4-dimensional vector with fixed orientation, thus making the experiment design almost trivial. However, the problem could be made more complex by performing an experiment design for both inertia and actuator alignment estimation. In this case, the amount of freedom due to actuator saturations changes significantly during the experiment. Therefore, an optimal experiment design should exploit the null-space by considering that the amount of torque available within the saturation limits changes significantly during the maneuver.

Appendices

A

Cepstrum Based Filter Synthesis

In this appendix a novel filter synthesis method is proposed, as an alternative to the *armax* routine of Matlab or to the *ivarma* from the CAPTAIN toolbox [Taylor et al., 2007]. The purpose of this method is to determine the coefficients of a rational transfer function $\frac{D(z^{-1})}{C(z^{-1})}$ which best approximates the behavior of the measured residuals. Unlike methods which are based on a high order AR approximation of the noise filter (e.g. *ivarma*), in this case a high order FIR (finite impulse response) is used. The use of a FIR representation has been proved more suitable for the noise models used in this work, which contains only derivatives.

The input of the algorithm is the time sequence of the estimation residuals, from which the power spectrum density is calculated. However, if the power spectrum is already known, as for the prefilter calculation of Chapter 6, this step can be skipped and the calculation can begin at step 3 of the algorithm.

A detailed description of the method is given beneath.

Method

In general, when a wide sense stationary signal $e(k)$ passes through a linear system $H(z^{-1})$, the autocorrelation of the output $\xi(k)$ is related to to the autocorrelation of the input $e(k)$ by the relation:

$$R_\xi = R_e * R_h, \tag{A.1}$$

where $R_h(k) = \sum_i h(i)h(k+i)$ is the autocorrelation of the filter impulse response.

This statement can be demonstrated by the following equations:

$$\begin{aligned} R_\xi(k) &= E[\xi(n)\xi(n+k)] = \\ &= E\Big[\sum_i \sum_j h(i)h(j)e(n-i)e(n+k-j)\Big] = \\ &= \sum_i \sum_j h(i)h(j)E[e(n-i)e(n+k-j)] = \\ &= \sum_i \sum_j h(i)h(j)R_e(k+i-j) \end{aligned}$$

assuming that $m = j - i$

$$R_\xi(k) = \sum_l R_e(k-m) \sum_i h(i)h(m+i) = \\ = \sum_l R_e(k-m)R_h(m) = R_e(k) * R_h(k)\,.$$

Since in our case the input is supposed to be a white noise, its autocorrelation is simply a Kronecker pulse multiplied by a constant σ_e^2 and therefore

$$R_\xi = \sigma_e^2 R_h\,. \tag{A.2}$$

The application of the Fourier transform to (A.2) yields:

$$\mathcal{F}(R_\xi) = \sigma_e^2 |H(e^{j\omega})|^2\,. \tag{A.3}$$

Taking the square root of (A.3) and neglecting the constant term σ_e (since we are interested only in the shape of the filter) we obtain the magnitude of the filter transfer function $|H(e^{j\omega})|$. By exploitation of the cepstrum properties, it is possible to find a minimum phase sequence $\hat{h}(n)$ which corresponds to a filter $\hat{H}(e^{i\omega})$ with the same magnitude. It is now possible, from the sequence $\hat{h}(n)$, to estimate the coefficients of low order polynomials $C(z^{-1})$ and $D(z^{-1})$, which yield an impulse response $h_2(n)$ that approximates $\hat{h}(n)$ in a least square sense. The inverse of this transfer function $\frac{C(z^{-1})}{D(z^{-1})}$ is taken as prefilter.

Based on the previous considerations, the proposed filter synthesis method is composed by the following steps:

Algorithm 1: Main steps of the filter estimation algorithm

1. calculate the autocorrelation sequence of the estimation residuals;
2. apply a Fast Fourier Transform to the autocorrelation sequence, obtaining a power spectrum density;
3. calculate the magnitude of the residual spectrum taking the square root of the power spectrum;
4. calculate the prefilter magnitude, which is the reciprocal of the spectrum magnitude;
5. find a minimum phase impulse response of the prefilter from its magnitude, exploiting the cepstrum properties;
6. use a linear regression to find the coefficients of the filter $\frac{D(z^{-1})}{C(z^{-1})}$ that best approximate the calculated impulse response;
7. take the inverse $\frac{C(z^{-1})}{D(z^{-1})}$ as prefilter.

Step 5 requires an additional explanation. The noise transfer function $H(e^{i\omega})$ is supposed to be stable. This means that its impulse response becomes negligible after a few samples. It can be therefore approximated with a FIR (moving average) filter with sufficient length. The calculation of the inverse transform of $H(e^{i\omega})$ can be done numerically by means of a FFT. The result would be a finite length non-causal zero phase sequence. In order to obtain a simpler time response, a minimum phase solution is sought. A minimum-phase FIR transfer function is characterized by the fact that all the zeros are located on or inside the unit circle in the Z plane. Therefore, in order to obtain a minimum phase solution without affecting the magnitude of the transfer function, all the zeros outside the unit circle must be moved inside, replacing

them with their reciprocals. The *polystab* function of Matlab cannot be used for this purpose, since for large size vectors (the length of the estimated FIR transfer function is of the order of 200 samples in our experiments) it becomes unstable producing unreliable results. A solution to this problem has been found exploiting the properties of the cepstrum sequences.

The cepstral sequence associated with a function $H(z^{-1})$ is defined as the inverse Z-transform of $log(H(z^{-1}))$:

$$c(n) = Z^{-1}[log(H(z^{-1}))]\,. \tag{A.4}$$

The FIR approximation $\hat{H}(e^{i\omega})$ of the transfer function $H(e^{i\omega})$ can be considered as a polynomial with coefficients $h(n)$, that can be decomposed as:

$$\hat{H}(z^{-1}) = Az^{N/2}\prod_{k=1}^{N}(1 - z_k z^{-1})\,. \tag{A.5}$$

It can be easily shown that the zeros inside the unit circle and the ones on it contribute to the causal component of $c(n) = Z^{-1}[log(H(z^{-1}))]$, while the ones outside the unit circle contribute to the anti-causal component of $c(n)$. In fact, for $|z_k| < 1$

$$Z^{-1}[log(1 - z_k z)] = \begin{cases} \frac{z_k^{-n}}{n} & n < 0 \\ 0 & n \geq 0 \end{cases} \tag{A.6}$$

and

$$Z^{-1}[log(1 - z_k z^{-1})] = \begin{cases} \frac{-z_k^{n}}{n} & n > 0 \\ 0 & n \leq 0 \end{cases}\,. \tag{A.7}$$

It follows that a time reversal of the values of $c(n)$ for $n < 0$ corresponds to move the zeros z_k^{-1} external to the unit circle to their reciprocal position z_k and gives the cepstral sequence $c_2(n)$ associated to a minimum phase filter:

$$c_2(n) = \begin{cases} 0 & n < 0 \\ c(0) & n = 0 \\ c(n) + c(-n) & n > 0 \end{cases}\,. \tag{A.8}$$

The sequence $c_2(n)$ corresponds therefore to a FIR transfer function $H_2(z^{-1})$ with the same magnitude of $H(z^{-1})$ but with a minimum phase. At this point, one could go back to the impulse response $h_2(n)$ of $H_2(z^{-1})$ by an inversion of the cepstral transformation. However, there is a shorter way using the recursion formula (Oppenheim and Schafer, 1975, Ch.10):

$$h_2(n) = \begin{cases} 0 & n < 0 \\ e^{c(0)} & n = 0 \\ c_2(n)h_2(0) + \sum_{k=1}^{n-1}\frac{k}{n}c_2(k)h_2(n-k) & n > 0 \end{cases}\,. \tag{A.9}$$

Example

As an example, the following ARMA model is considered

$$H(z^{-1}) = \frac{3.851z^{-3} - 3.752z^{-2} - 0.1481z^{-1} + 3.95}{1.926z^{-1} + 1.975}\,, \tag{A.10}$$

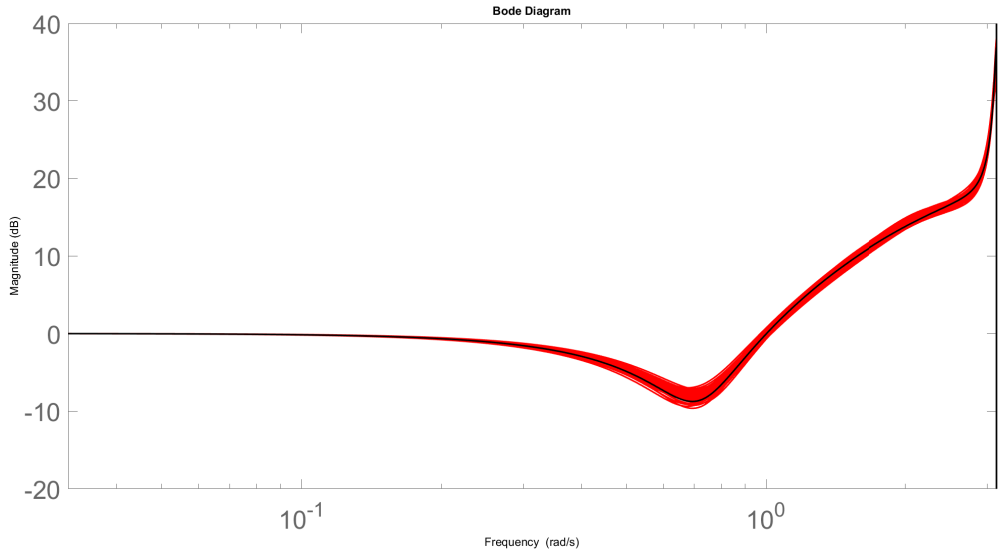


Figure A.1: Bode plot of the true (black line) and estimated model with the proposed method (red lines) from a Monte Carlo of 100 runs.

where its transfer function has a similar “shape" to the one that we expect from the satellite inverse model overall noise[27] (a high-pass amplifier due to the derivatives term, and a low-pass due to the disturbance torque).

The proposed cepstrum-based ARMA model estimation is applied to the ARMA output ξ

$$\xi(k) = H(z^{-1})\, e(k)\,, \tag{A.11}$$

where $e(t_k)$ is a zero-mean white noise signal. For this numerical simulation, 5000 output samples have been considered and used by the proposed algorithm.

The results are shown in Figure A.1 where the black line represents the overall true spectrum of the residuals (a high-pass amplifier due to the derivatives term, and a low-pass due to the disturbance torque) while in red the third order estimated model with the cepstrum based approach from a Monte Carlo of 100 runs. In case of the IV method, the prefilter is then simply the reciprocal of the estimated ARMA transfer function.

[27] Chapter 5 and 6 for further details.

B

Application of Delayed State Variables Filter

As shown in Chapter 5, there is the need to filter the quaternion signal before calculating its derivative. Since the system equation is nonlinear, no signal distortion is allowed. The solution exploited in Chapter 5 has been a zero lag smoother, which uses a forward and a backward filtering step to ensure a zero lag. However, this method can only be applied for off-line estimation. An alternative can be found using a delayed state-variable filter (DSVF). The DSVF is a filter that has a frequency response presenting constant unit gain and an almost linear phase with respect to the frequency in the passband of the filter. This means that all the signal components in the passband are subject to a constant delay [Tsang and Billings, 1994]. These filters are commutative with a nonlinear function. This means that, given a signal $x(t)$ passing through a nonlinear function $F(\cdot)$, applying the delay filter at the resulted signal $y(t)$ will give the same output $y_f(t)$ as applying the filter directly to $x(t)$ before going through function $F(\cdot)$ (See Figure B.1).

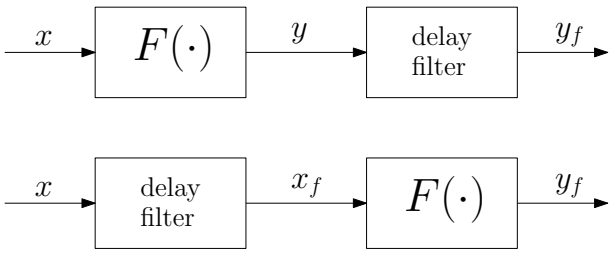


Figure B.1: The order of the nonlinear function $F(\cdot)$ and of the delay filter does not change the output.

Delay filter

A delay filter has a frequency response that presents constant unit gain and an almost linear phase with respect to the frequency in the passband of the filter. No finite order filters can produce such response exactly, therefore an approximation of such response will be sought. The Butterworth filter, given a cut-off frequency Ω_C presents a maximally flat amplitude characteristic and an almost linear phase characteristic for frequency below the cut-off frequency Ω_C. For these reasons, a Butterworth filter is considered for the approximation of the delay filter.

The transfer function for a third-order Butterworth filter is as follows:

$$H_b(s) = \frac{\Omega_C^3}{s^3 + 2s^2\Omega_C + 2s\Omega_C^2 + \Omega_C^3} \tag{B.1}$$

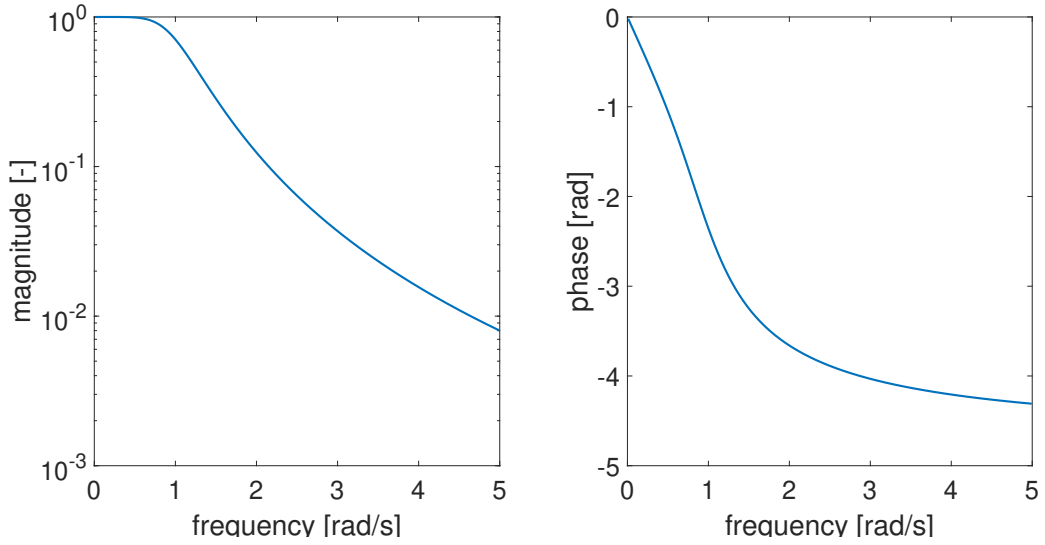


Figure B.2: Bode plot for a third-order Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$.

However, as it can be seen from Figure (B.2), the phase is not exactly linear in the the pass band (below the cut-off frequency Ω_C), or equivalently, the group delay is not constant (Figure B.3). The Butterworth filter can be improved by using equalization filters. Since the group delay of the Butterworth filter needs to be constant, while its magnitude need to be kept the same, a constant gain all-pass equalizer should be considered:

$$H_e(s) = \frac{(s-\alpha)^2 + \beta^2}{(s+\alpha)^2 + \beta^2}\,. \tag{B.2}$$

The equalizer $H_e(s)$ group delay is

$$T_e(\omega) = \frac{4\alpha(\omega^2 + \alpha^2 + \beta^2)}{(\alpha^2 + \beta^2 - \omega^2)^2 + 4\alpha^2\omega^2}\,, \tag{B.3}$$

where in this case ω represents the frequency (and no more the satellite angular rate).

The values of α and β are selected in order to have an overall group delay that best approximates a constant delay. For such objective, the following cost function is minimized

$$J = \int_0^{\Omega_{Tmax}} \big(T_0 - T_b(\omega) - T_e(\omega)\big)^2 d\omega\,, \tag{B.4}$$

where T_0 is the nominal group delay, $T_b(\omega)$ is the group delay of the Butterworth filter, and Ω_{Tmax} is the frequency where the Butterworth filter group delay is a maximum.

The group delay of an equalized butterworth filter is shown in Figure B.4, and it is significantly more flat in the frequency of interest (below the cut-off frequency).

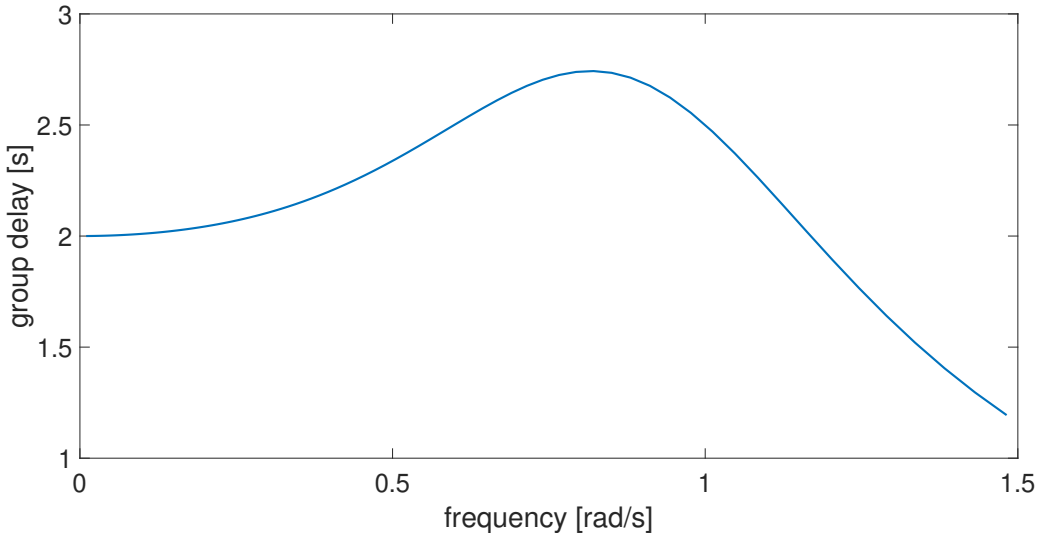


Figure B.3: Group delay for the third-order Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$.

Delayed state variable filter implementation

The derivative estimation through a DSVF filter is very similar to the SVF one shown in Chapter 4. Let us consider the following forth-order Butterworth filter

$$F_{B4}(s) = \frac{\Omega_C^4}{s^4 + (c_1 + c_2)\Omega_C s^3 + (c_1 c_2 + 2)\Omega_C^2 s^2 + (c_1 + c_2)\Omega_C^3 s + \Omega_C^4} \tag{B.5}$$

with

$$c_1 = -2\,cos\left(\frac{5}{8}\pi\right) \qquad \text{and} \qquad c_2 = -2\,cos\left(\frac{7}{8}\pi\right). \tag{B.6}$$

as well as a second-order equilizer (all-pass filter) (B.2).

By denoting a signal $y(t)$ and $y_f(t)$ being the corresponding filtered version by the Butterworth filter, its successive filtered derivatives are as follows

$$\dot{y}_f = \Omega_C^4 F_{B3}(s) y - \Omega_C^4 F_{B3}(s) y_f\,, \tag{B.7}$$

$$\ddot{y}_f = \Omega_C^4 F_{B2}(s) y - ((c_1 + c_2)\Omega_C^3) F_{B2}(s)\dot{y}_f - \Omega_C^4 F_{B2}(s) y_f\,, \tag{B.8}$$

$$\dddot{y}_f = \Omega_C^4 F_{B1}(s) y - (c_1 c_2 + 2)\Omega_C^2 F_{B1}(s)\ddot{y}_f - (c_1 + c_2)\Omega_C^3 F_{B1}(s)\dot{y}_f - \Omega_C^4 F_{B1}(s)\,, \tag{B.9}$$

where

$$F_{B3}(s) = \frac{1}{s^3 + (c_1 + c_2)\Omega_C s^2 + (c_1 c_2 + 2)\Omega_C^2 s + (c_1 + c_2)\Omega_C^3}\,, \tag{B.10}$$

$$F_{B2}(s) = \frac{1}{s^2 + (c_1 + c_2)\Omega_C s + (c_1 c_2 + 2)\Omega_C^2}\,, \tag{B.11}$$

$$F_{B1}(s) = \frac{1}{s + (c_1 + c_2)\Omega_C}\,. \tag{B.12}$$

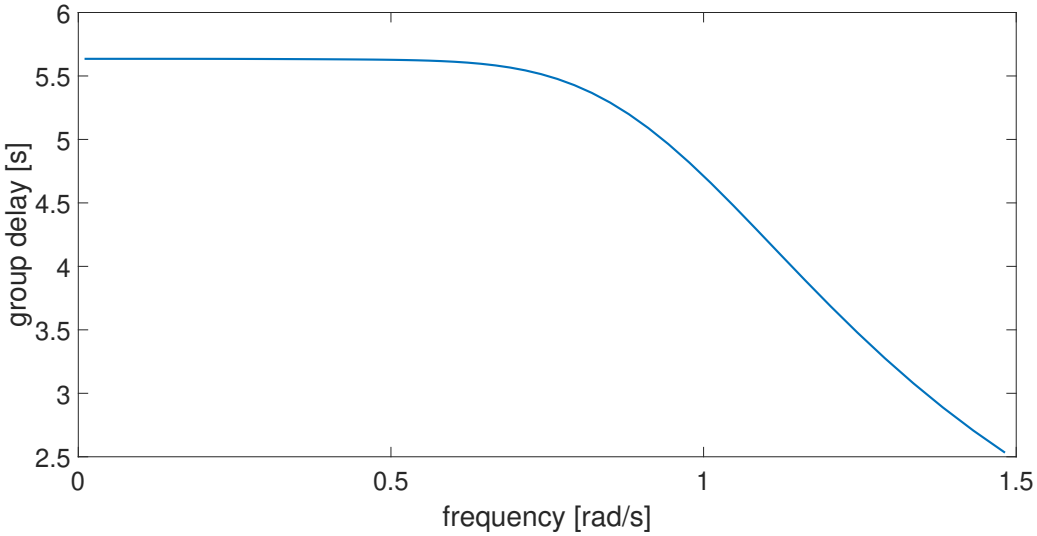


Figure B.4: Group delay for the equalized Butterworth filter with a cut-off frequency $\Omega_C = 1\,rad/s$.

Résumé en français

Cette thèse traite de l'estimation des paramètres du modèle satellite qui régissent la dynamique de rotation et qui affectent son contrôle d'attitude. L'attitude est l'orientation tridimensionnelle d'un vaisseau spatial par rapport à un référentiel spécifié. La capacité de contrôler son attitude est essentielle pour presque tous les satellites. Très souvent, une précision de pointage élevée est requise, en particulier pour les satellites construits pour des applications qui impliquent l'observation ou la mesure au sol ou dans l'espace. En général, la précision de pointage requise de la structure de l'engin spatial est déterminée par sa mission. Les applications les plus exigeantes sont celles qui impliquent un télescope, où l'erreur de pointage admissible est de l'ordre de 10^{-6} degrés.

Les méthodes de contrôle d'attitude peuvent être divisées en deux catégories : les méthodes actives, qui contrôlent le satellite en utilisant des actionneurs de couple, et les méthodes passives, qui utilisent la réaction du satellite à un certain effet physique pour assurer sa stabilité (par exemple stabilisation du gradient de gravité) ou qui utilisent propriétés du moment angulaire (stabilisation de spin). Les méthodes passives ne peuvent fournir qu'un contrôle d'attitude grossier, adapté aux satellites simples, qui ne nécessitent qu'un alignement de base. Dans ce travail, nous considérons le contrôle d'attitude actif, qui est utilisé par la plupart des satellites. Le système de contrôle d'attitude (parfois appelé système de détermination et de contrôle d'attitude, ou ADCS) comprend des capteurs, des actionneurs, des ordinateurs, des logiciels et des équipements de support au sol. Le schéma de commande typique d'un ADCS est illustré à la Figure 1.

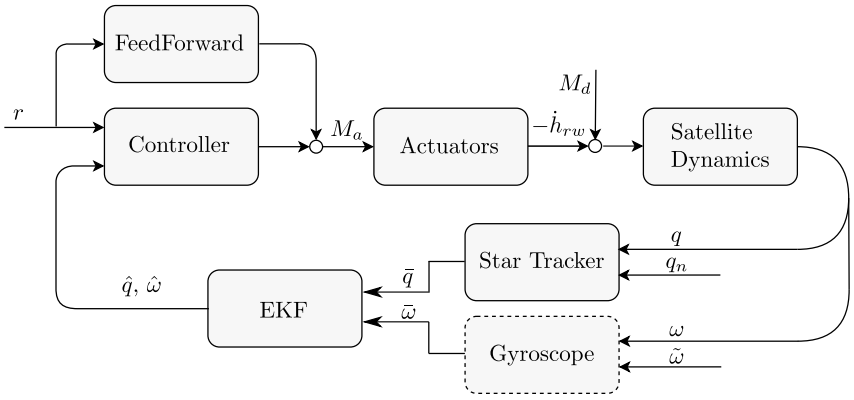


Figure 1: Schéma de contrôle d'attitude.

La connaissance précise du modèle de dynamique de rotation des satellites est fondamentale pour plusieurs raisons. Tout d'abord, pour un réglage optimal du système de détermination d'attitude et de contrôle, en particulier si un contrôleur prédictif est utilisé. Ensuite, des modèles mathématiques précis d'un satellite sont nécessaires à de nombreuses autres fins, comme le

diagnostic, la simulation ou encore l'estimation d'état.

Par ailleurs, une estimation précise des paramètres de dynamique de rotation des satellites n'est possible qu'en orbite. En fait, plusieurs facteurs modifient la masse globale et sa distribution après le lancement du satellite, tels que : la consommation de carburant, le déploiement des panneaux solaires et la déformation de la structure du vaisseau spatial. De plus, l'alignement des actionneurs peut légèrement changer après le lancement en raison de la distorsion du cadre. Cela conduit au sujet de cette thèse : estimation des paramètres des satellites à partir des données de télémétrie. Tous les algorithmes d'estimation proposés dans ce travail sont basés sur des données de télémétrie collectées lors d'une manœuvre d'attitude appropriée.

Le séquencement des chapitres de cette thèse suit en partie la chronologie des activités du travail doctoral. Tout d'abord, un modèle mathématique de la dynamique d'attitude du satellite a été défini, basé sur deux ensembles de paramètres : la matrice d'inertie et la matrice d'alignement de l'actionneur. Ensuite, une méthode fiable et cohérente a été recherchée en commençant par des conditions relativement simples et proches de l'idéal. Plusieurs sources d'erreur et d'autres objectifs ont été progressivement ajoutés, jusqu'à ce que le modèle mathématique se rapproche le plus possible de l'environnement réel. Aussi des simulations proches des conditions réelles de l'environnement ont été recherchées: un simulateur haute fidélité du CNES a été utilisé et des tests ont également été effectués sur les données collectées à partir d'un vrai satellite. Enfin, après avoir obtenu des algorithmes d'estimation satisfaisants, le plan d'expérience a été pris en considération, dans le but d'améliorer encore la précision des estimations.

L'organisation détaillée de la thèse est présentée ci-dessous.

Le premier chapitre donne un aperçu du contexte du problème et définit les motivations ainsi que la portée de ce travail. Une brève description de la structure du satellite est donnée et le problème général de l'identification du système est présenté. Un ordre de grandeur typique est également donné concernant les poids et les orbites des satellites.

Le Chapitre 2 commence par quelques définitions qui sont utilisées dans la suite de ce travail : les principaux systèmes de référence d'une part et les méthodes de représentation d'attitude d'autre part, avec un accent sur les propriétés du quaternion, puisque les quaternions sont largement utilisés dans la suite du document. Ensuite, les équations qui régissent la dynamique de rotation des satellites sont présentées. Ces équations sont utilisées pour construire un modèle de type boîte grise, où les paramètres inconnus sont représentés par l'inertie du satellite (ou par l'inertie et les vecteurs d'alignement de l'actionneur dans le cas plus général). Le modèle est ensuite complété par le couple de perturbation. Le chapitre se termine par une description des principaux capteurs satellites (capteur d'étoiles, gyroscope, encodeur de roue de réaction) et des actionneurs de contrôle d'attitude (roues de réaction, actionneur gyroscopique et magnéto-coupleur).

Le problème de l'estimation des paramètres est abordé au Chapitre 3. L'état de l'art présente les différentes solutions proposées dans la littérature pour l'estimation des paramètres satellitaires. De nombreux auteurs proposent des méthodes basées sur les moindres carrés ou sur le filtre de Kalman. Dans la littérature, aucun auteur n'a jamais utilisé la méthode de la variable instrumentale (IV), malgré ses propriétés intéressantes. Quelques considérations sont faites sur les problèmes possibles pour cette estimation des paramètres. Tout d'abord, il est brièvement montré comment dans cette application particulière, la structure du système en boucle fermée n'affecte pas de manière critique l'identification grâce à la présence d'un filtre dans la boucle de rétroaction. Deuxièmement, les avantages et les inconvénients liés à l'utilisation du modèle satellite inverse sont présentés, avec un accent particulier sur les similitudes par rapport au modèle « errors-in-variables » (EIV). Après une description du modèle satellite pour l'estimation des paramètres et des méthodes d'estimation possibles, le chapitre se termine par le choix de

méthode de l'auteur: la variable instrumentale (IV). Afin de justifier ce choix, les principaux avantages de l'approche IV sont présentés, l'accent étant mis sur les conditions EIV qui rendent plusieurs autres méthodes biaisées. Cependant, comme mentionné dans ce chapitre, compte tenu également de la nouveauté de l'approche IV pour cette application, plusieurs étapes supplémentaires et des ajustements sur mesure doivent être apportés à la conception de la méthode IV afin de fonctionner efficacement.

Au Chapitre 4, le problème d'estimation de l'inertie pour un satellite équipé de gyroscopes est examiné. Il s'agit du scénario le plus courant considéré dans la littérature. L'objectif principal de ce chapitre consiste à montrer comment la méthode de variable instrumentale peut également être adaptée à ce problème d'estimation. Des mesures de fréquence angulaire ont été envisagées, qui passent par le modèle inverse non linéaire, et une analyse détaillée du bruit est effectuée, qui est ensuite utilisée pour définir un filtre optimal. Ce filtre a pour fonction de réduire la variance des estimations, mais il est également utilisé pour éviter la différenciation numérique nécessaire au calcul de l'accélération angulaire. L'algorithme proposé utilise un modèle auxiliaire pour construire l'instrument, qui est mis à jour de manière itérative avec les paramètres calculés. La méthode a été validée par des simulations de Monte Carlo, où également une estimation par moindres carrés a été exécutée pour comparaison. Il est démontré que la méthode IV peut surmonter les limites des moindres carrés et obtenir des résultats non biaisés, malgré le bruit dans le régresseur. La nature en boucle fermée du système, qui pose généralement un problème lors de l'identification, est ici exploitée pour garantir que le modèle auxiliaire en boucle fermée, qui est utilisé pour générer l'instrument, n'a aucune dérive par rapport au système réel. Dans la dernière section de ce chapitre, l'effet de la dérive du gyroscope sur les estimations a été analysé et une solution a été proposée pour estimer et supprimer une partie de l'effet de dérive. Cette solution a également été validée par des simulations, montrant que la méthode est robuste même en cas de valeurs de biais très élevées.

Au Chapitre 5, la méthode de variable instrumental est étendue aux satellites sans gyroscope. Cela permet d'estimer les paramètres d'inertie également pour les petits satellites, qui utilisent souvent des gyroscopes MEMS bon marché et inexacts (ou parfois pas de gyroscope du tout). Premièrement, le filtre du Chapitre 4 n'étant plus applicable, d'autres méthodes sont envisagées pour traiter, en partie, de l'amplification du bruit due aux dérivées temporelles. Deuxièmement, une analyse détaillée du bruit est effectuée qui étudie l'effet du bruit sur les mesures d'attitude sur le modèle inverse complet. étant donné que le nouveau signal d'entrée est basé sur les quaternions d'attitude, qui sont liés à la vitesse angulaire et à l'accélération angulaire par des équations non linéaires, certaines linéarisations ont été nécessaires pour concevoir le préfiltre (mais aucune linéarisation n'est effectuée dans le modèle pour les paramètres estimés). De plus, en raison de la relation non linéaire entre les quaternions et la vitesse angulaire, il n'a pas été possible d'utiliser un filtre à variable d'état pour éviter la différenciation numérique. Un filtre Butterworth a été utilisé à la place, suivi d'un préfiltre. Les résultats de cette analyse, ainsi qu'un modèle approximatif des couples de perturbation, sont utilisés pour la conception d'un filtre sous-optimal pour la méthode IV. De plus, le modèle et l'algorithme d'estimation respectif sont augmentés afin d'estimer et de compenser la condition initiale du filtre et de supprimer certaines des composantes déterministes (constantes et sinusoïdales) du couple de perturbation. L'estimation de l'état initial du préfiltre permet une meilleure utilisation des données de télémétrie puisque le transitoire de filtre initial est estimé et supprimé, évitant ainsi d'avoir à jeter les données initiales. Cela a une importance significative pour les expériences courtes, comme celles décrites au Chapitre 7, mais il est également important pour les expériences standards, où la plus grande disponibilité des données présente encore certains avantages sur la précision de l'estimation. Pour les simulations numériques, toutes les données utilisées dans ce chapitre ont été générées par un simulateur haute

fidélité du CNES, où toutes les sources possibles de perturbation sont présentes. Les conditions de simulation sont donc aussi proches que possible de l'environnement réel (Figure 2). Enfin, la méthode a également été testée avec des données réelles enregistrées lors d'une expérience sur le satellite Picard avant sa fin de vie.

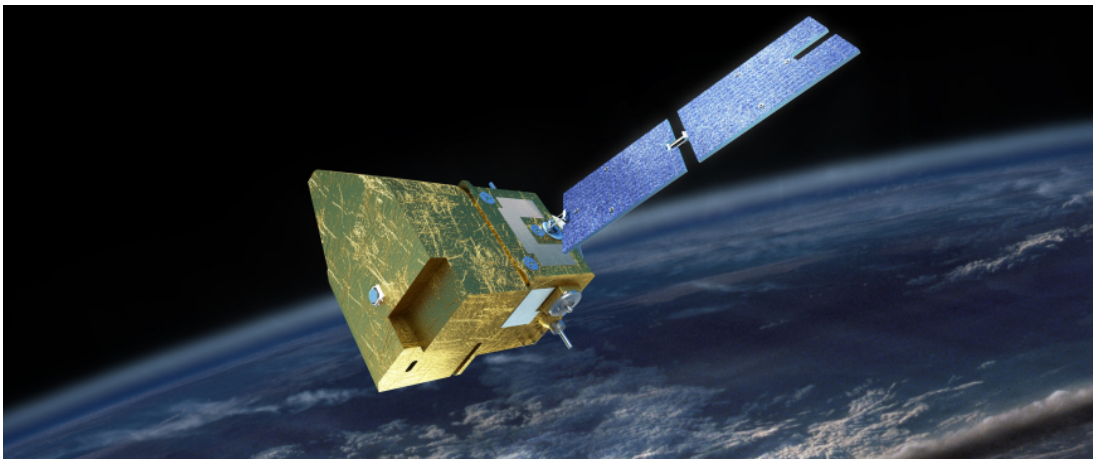


Figure 2: Satellite MicroCarb du simulateur haute fidélité (source: CNES).

Au Chapitre 6, le problème d'identification est étendu à l'estimation de l'alignement de l'actionneur, tout en considérant un satellite sans gyroscope. Pour l'estimation de l'alignement, un algorithme itératif a été développé afin d'éviter de traiter les contraintes quadratiques impliquées par les vecteurs unitaires. De plus, une synchronisation non parfaite entre les capteurs est prise en compte (les mesures du suiveur stellaire ne sont pas parfaitement synchronisées avec les lectures de vitesse de l'actionneur). À la connaissance de l'auteur, cette question n'a jamais été abordée dans la littérature. Cependant, comme le montrent les simulations, même une petite erreur de synchronisation peut avoir des effets négatifs non négligeables sur la précision des estimées. Une méthode, qui utilise un double régresseur, a été développée pour estimer le décalage temporel des données et supprimer son effet. Dans cette implémentation IV, un régresseur retardé a été utilisé comme instrument, donc aucune information sur le système en boucle fermée et en particulier le contrôleur satellite n'est nécessaire. Dans cette partie du travail, une méthode alternative a été développée et testée pour la conception du préfiltre. La méthode exploite les propriétés du cepstre pour obtenir une approximation de la réponse impulsionnelle du filtre et elle semble bien adaptée au type de modèle de bruit de cette application. Là encore, l'algorithme est testé sur des données générées par le simulateur haute fidélité du CNES afin de valider les performances dans ce nouveau cas, ainsi que de justifier toutes les «extensions» de la méthode.

La partie finale et la dernière contribution de cette thèse (Chapitre 7) se concentrent sur la conception de l'expérience pour l'estimation de l'inertie du satellite. Jusqu'à ce point, les expériences ont été réalisées en considérant l'exigence de persistance de l'excitation. Ceci a été obtenu en suivant quelques règles simples pour les manœuvres: les trois axes du satellite doivent être excités, l'excitation sur différents axes doit être linéairement indépendante (par exemple éviter d'avoir la même manœuvre sur différents axes) et le niveau du signal d'excitation doit être le plus élevé possible, compatible avec le système et l'actionneur contraintes. La persistance de l'excitation est une condition nécessaire, mais on peut faire plus mieux. Dans cette dernière partie du travail, des techniques de conception d'expériences ont été exploitées dans le but d'améliorer encore la précision de l'estimation des paramètres. Afin de concevoir des manœuvres optimales, la méthode proposée utilise une représentation cubique B-spline de la trajectoire. La manœuvre

optimisée est obtenue grâce à la minimisation d'une fonction basée sur la matrice d'informations de Fisher. Le processus d'optimisation prend également en compte les contraintes physiques dues à la saturation des actionneurs. Différents critères d'optimalité sont testés et les manœuvres optimisées sont comparées à une manœuvre de test typique, montrant comment la conception de l'expérience proposée aboutit à des estimées des paramètres de la matrice d'inertie nettement plus précises. Enfin, la manœuvre générée par la conception de l'expérience a été testée sur un vrai prototype, un modèle réduit simulant un satellite, dans des conditions proches de zéro-G (vol zéro-G). Toujours dans l'expérience réelle, des estimations précises de l'inertie ont été obtenues, en particulier compte tenu des ensembles de données extrêmement courts.

Le dernier chapitre résume les principales contributions de ce travail et propose quelques pistes de réflexion pour de futurs travaux. La première proposition est l'utilisation de filtres complémentaires, comme alternative au filtre de Kalman, pour combiner le signal du suiveur stellaire et du gyroscope. La deuxième proposition, pour l'estimation de l'alignement des actionneurs, est d'exploiter l'espace nul noyau des actionneurs, en les utilisant près de leurs limites de saturation.

Bibliography

[Ahm+98] Ahmed, J., Coppola, V. T., and Bernstein, D. S. "Adaptive asymptotic tracking of spacecraft attitude motion with inertia matrix identification". In: *Journal of Guidance, Control, and Dynamics* 21.5 (1998), pp. 684–691.

[Ang03] Angelo, J. A. *Space technology*. Greenwood Publishing Group, 2003.

[Arm89] Armstrong, B. "On finding exciting trajectories for identification experiments involving systems with nonlinear dynamics". In: *The International journal of robotics research* 8.6 (1989), pp. 28–48.

[BOG94] Bar-On, J. and Grasse, K. "Global optimization of a quadratic functional with quadratic equality constraints". In: *Journal of Optimization Theory and Applications* 82.2 (1994), pp. 379–386.

[BOG97] Bar-On, J. and Grasse, K. "Global optimization of a quadratic functional with quadratic equality constraints, part 2". In: *Journal of Optimization Theory and Applications* 93.3 (1997), pp. 547–556.

[Ber+87] Bergmann, E. V., Walker, B. K., and Levy, D. R. "Mass property estimation for control of asymmetrical satellites". In: *Journal of Guidance, Control, and Dynamics* 10.5 (1987), pp. 483–491.

[BH+92] Brown, R. G., Hwang, P. Y., et al. *Introduction to random signals and applied Kalman filtering*. Vol. 3. Wiley New York, 1992.

[BJ18] Brunot, M and Janot, A. "A New Recursive Instrumental Variables Approach for Robot Identification". In: *18th IFAC Symposium on System Identification* 51.15 (2018), pp. 132–137.

[Bru+18] Brunot, M, Janot, A, Young, P., and Carrillo, F. "An improved instrumental variable method for industrial robot model identification". In: *Control Engineering Practice* 74 (2018), pp. 107–117.

[Cal+01] Calafiore, G, Indri, M., and Bona, B. "Robot dynamic calibration: Optimal excitation trajectories and experimental parameter estimation". In: *Journal of robotic systems* 18.2 (2001), pp. 55–68.

[Che+16] Chen, F., Garnier, H., Gilson, M., Agüero, J. C., and Liu, T. "Refined instrumental variable parameter estimation of continuous-time Box–Jenkins models from irregularly sampled data". In: *IET Control Theory & Applications* 11.2 (2016), pp. 291–300.

[CB97] Coca, D and Billings, S. "Continuous-time system identification for linear and nonlinear systems using wavelet decompositions". In: *International Journal of Bifurcation and Chaos* 7.01 (1997), pp. 87–96.

[CJ11] Crassidis, J. L. and Junkins, J. L. *Optimal estimation of dynamic systems*. Chapman and Hall/CRC, 2011.

[Cra+07] Crassidis, J. L., Markley, F. L., and Cheng, Y. "Survey of nonlinear attitude estimation methods". In: *Journal of guidance, control, and dynamics* 30.1 (2007), pp. 12–28.

[Cur13] Curtis, H. D. *Orbital mechanics for engineering students*. Butterworth-Heinemann, 2013.

[Dam01] Damilano, P. "Pleiades high resolution satellite: a solution for military and civilian needs in metric-class optical observation". In: *15th Annual/USU Conference on Small Satellites* (2001).

[Eva+19] Evain, H., Alazard, D., Rognant, M., Solatges, T., Brunet, A., Mignot, J., Rodriguez, N., and Dias-Ribeiro, A. "Satellite Attitude Control with a six-Control Moment Gyro Cluster tested under Microgravity Conditions". In: *International Symposium on Space Flight Dynamics* (2019).

[For99] Forssell, U. "Closed-loop identification: Methods, theory, and applications". PhD thesis. Linköping University Electronic Press, 1999.

[For+11] Fortescue, P., Swinerd, G., and Stark, J. *Spacecraft systems engineering*. John Wiley & Sons, 2011.

[FN09] Fosbury, A. and Nebelecky, C. "Spacecraft actuator alignment estimation". In: *AIAA Guidance, Navigation, and Control Conference*. 2009, p. 6316.

[FM08] Franceschini, G. and Macchietto, S. "Model-based design of experiments for parameter precision: State of the art". In: *Chemical Engineering Science* 63.19 (2008), pp. 4846–4872.

[Gar15] Garnier, H. "Direct continuous-time approaches to system identification. Overview and benefits for practical applications". In: *European Journal of Control* 24 (2015), pp. 50–62.

[Gar+03] Garnier, H., Mensler, M., and Richard, A. "Continuous-time model identification from sampled data: implementation issues and performance evaluation". In: *International journal of Control* 76.13 (2003), pp. 1337–1357.

[GK92] Gautier, M. and Khalil, W. "Exciting trajectories for the identification of base inertial parameters of robots". In: *The International journal of robotics research* 11.4 (1992), pp. 362–375.

[GV18] Genin, F. and Viaud, F. "An innovative control law for Microcarb microsatellite". In: *32nd annual AAS Guidance and Control Conference* (2018).

[GH05] Gilson, M. and Hof, P. Van den. "Instrumental variable methods for closed-loop system identification". In: *Automatica* 41.2 (2005), pp. 241–249.

[Gil+11] Gilson, M., Garnier, H., Young, P. C., and Hof, P. M. Van den. "Optimal instrumental variable method for closed-loop identification". In: *IET control theory & applications* 5.10 (2011), pp. 1147–1154.

[Gle+12] Gleyzes, M. A., Perret, L., and Kubik, P. "Pleiades system architecture and main performances". In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 39.1 (2012), pp. 537–542.

[GVL80] Golub, G. H. and Van Loan, C. F. "An analysis of the total least squares problem". In: *SIAM journal on numerical analysis* 17.6 (1980), pp. 883–893.

[GVM91] Golub, G. H. and Von Matt, U. "Quadratically constrained least squares and quadratic problems". In: *Numerische Mathematik* 59.1 (1991), pp. 561–580.

[GP77] Goodwin, G. C. and Payne, R. L. *Dynamic system identification: experiment design and data analysis*. Academic press, 1977.

[Hay+01] Haykin, S. S. et al. *Kalman filtering and neural networks*. Wiley Online Library, 2001.

[Hig75] Higgins, W. T. "A comparison of complementary and Kalman filtering". In: *IEEE Transactions on Aerospace and Electronic Systems* 3 (1975), pp. 321–325.

[HE18] Ho, D. and Enqvist, M. "On the equivalence of forward and inverse IV estimators with application to quadcopter modeling". In: *18th IFAC Symposium on System Identification* 51.15 (2018), pp. 951–956.

[Hof98] Hof, P. Van den. "Closed-loop issues in system identification". In: *Annual reviews in control* 22 (1998), pp. 173–186.

[IM10] Isermann, R. and Münchhof, M. *Identification of dynamic systems: an introduction with applications*. Springer Science & Business Media, 2010.

[Jan+13] Janot, A., Vandanjon, P.-O., and Gautier, M. "A generic instrumental variable approach for industrial robot identification". In: *IEEE Transactions on Control Systems Technology* 22.1 (2013), pp. 132–145.

[JB06] Jun, B.-E. and Bernstein, D. S. "Least-correlation estimates for errors-in-variables models". In: *International Journal of Adaptive Control and Signal Processing* 20.7 (2006), pp. 337–351.

[Jun+10] Jun, B.-E., Bernstein, D. S., and McClamroch, N. H. "Identification of the inertia matrix of a rotating body based on errors-in-variables models". In: *International Journal of Adaptive Control and Signal Processing* 24.3 (2010), pp. 203–210.

[Kay98] Kay, S. M. *Estimation theory*. Prentice Hall PTR, 1998.

[Kei+06] Keim, J. A., Acikmese, A. B., and Shields, J. F. "Spacecraft inertia estimation via constrained least squares". In: *2006 IEEE Aerospace Conference*. IEEE. 2006, 6–pp.

[KM06] Klein, V. and Morelli, E. A. *Aircraft system identification: theory and practice*. American Institute of Aeronautics and Astronautics Reston, VA, 2006.

[LW92] Larson, W. J. and Wertz, J. R. *Space mission analysis and design*. Tech. rep. Torrance, CA (United States); Microcosm, Inc., 1992.

[LW02] Lee, A. Y. and Wertz, J. A. "In-flight estimation of the Cassini spacecraft's inertia tensor". In: *Journal of spacecraft and rockets* 39.1 (2002), pp. 153–155.

[LJ14] Leve, F. and Jah, M. "Spacecraft actuator alignment determination through null-motion excitation". In: *IEEE Transactions on Aerospace and Electronic Systems* 50.3 (2014), pp. 2336–2342.

[Lju79] Ljung, L. "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems". In: *IEEE Transactions on Automatic Control* 24.1 (1979), pp. 36–50.

[Lju99] Ljung, L. *System identification: Theory for the user*. 2nd. Prentice-Hall, 1999.

[MG06] Mahata, K. and Garnier, H. "Identification of continuous-time errors-in-variables models". In: *Automatica* 42.9 (2006), pp. 1477–1490.

[Mah+08] Mahony, R., Hamel, T., and Pflimlin, J.-M. "Nonlinear complementary filters on the special orthogonal group". In: *IEEE Transactions on automatic control* 53.5 (2008), pp. 1203–1218.

[MP17] Manchester, Z. R. and Peck, M. A. "Recursive inertia estimation with semidefinite programming". In: *AIAA Guidance, Navigation, and Control Conference*. 2017, p. 1902.

[MVH07] Markovsky, I. and Van Huffel, S. "Overview of total least-squares methods". In: *Signal processing* 87.10 (2007), pp. 2283–2302.

[Nai+18] Nainer, C., Garnier, H., Gilson, M., and Pittet, C. "In-Orbit Data Driven Identification of Satellite Inertia Matrix". In: *18th IFAC Symposium on System Identification* (2018).

[Nai+19] Nainer, C., Garnier, H., Gilson, M., Evain, H., and Pittet, C. "In-Flight inertia matrix estimation of a gyroless satellite". In: *5th CEAS Conference on Guidance Navigation and Control* (2019).

[Nai+20a] Nainer, C., Gilson, M., Garnier, H., Pittet, C., and Evain, H. "Design of Satellite Maneuvers for Inertia Parameter Estimation". In: *21st IFAC World Congress*. 2020.

[Nai+20b] Nainer, C., Garnier, H., Gilson, M., Evain, H., and Pittet, C. "Parameter Estimation of a Gyroless Micro-Satellite from Telemetry Data". In: *[submitted] Journal of Guidance, Control, and Dynamics*. 2020.

[Nor+11] Norman, M. C., Peck, M. A., and O'shaughnessy, D. J. "In-orbit estimation of inertia and momentum-actuator alignment parameters". In: *Journal of Guidance, Control and Dynamics* 34.6 (2011), pp. 1798–1814.

[OS75] Oppenheim, A. V. and Schafer, R. W. "Digital signal processing(Book)". In: *Prentice-Hall* (1975).

[Par06] Park, K.-J. "Fourier-based optimal excitation trajectories for the dynamic identification of robots". In: *Robotica* 24.5 (2006), pp. 625–633.

[Pec01] Peck, M. A. "Estimation of wheel and CMG alignments from on-orbit telemetry". In: *2001 Flight Mechanics Symposium*. Vol. 1. 2001.

[Pit+15] Pittet, C, Luzi, A., Peaucelle, D., Biannic, J.-M., and Mignot, J. "In-flight results of adaptive attitude control law for a microsatellite". In: *CEAS Space Journal* 7.2 (2015), pp. 291–302.

[Psi05] Psiaki, M. L. "Estimation of a spacecraft's attitude dynamics parameters by using flight data". In: *Journal of Guidance Control and Dynamics* 28.4 (2005), p. 594.

[Rac+12] Rackl, W., Lampariello, R., and Hirzinger, G. "Robot excitation trajectories for dynamic parameter estimation using optimized B-splines". In: *2012 IEEE international conference on robotics and automation*. IEEE. 2012, pp. 2042–2047.

[SG64] Savitzky, A. and Golay, M. J. "Smoothing and differentiation of data by simplified least squares procedures." In: *Analytical chemistry* 36.8 (1964), pp. 1627–1639.

[Sch11] Schafer, R. W. "On the frequency-domain properties of Savitzky-Golay filters". In: *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*. IEEE. 2011, pp. 54–59.

[Sch07] Schumaker, L. *Spline functions: basic theory*. Cambridge University Press, 2007.

[Sek+09] Sekhavat, P., Karpenko, M., and Ross, I. "UKF-based spacecraft parameter estimation using optimal excitation". In: *AIAA Guidance, Navigation, and Control Conference*. 2009, p. 5786.

[Sid97] Sidi, M. J. *Spacecraft dynamics and control: a practical engineering approach*. Vol. 7. Cambridge university press, 1997.

[SS83] Söderström, T and Stoica, P. *Instrumental variable methods for system identification, Lectures Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1983.

[Söd07] Söderström, T. "Errors-in-variables methods in system identification". In: *Automatica* 43.6 (2007), pp. 939–958.

[SM02] Söderström, T. and Mahata, K. "On instrumental variable and total least squares approaches for identification of noisy systems". In: *International Journal of Control* 75.6 (2002), pp. 381–389.

[Swe+97] Swevers, J., Ganseman, C., Tukel, D. B., De Schutter, J., and Van Brussel, H. "Optimal robot excitation and identification". In: *IEEE transactions on robotics and automation* 13.5 (1997), pp. 730–740.

[TW97] Tanygin, S. and Williams, T. "Mass property estimation using coasting maneuvers". In: *Journal of Guidance, Control, and Dynamics* 20.4 (1997), pp. 625–632.

[Tay+07] Taylor, C. J., Pedregal, D. J., Young, P. C., and Tych, W. "Environmental time series analysis and forecasting with the Captain toolbox". In: *Environmental Modelling & Software* 22.6 (2007), pp. 797–814.

[Thi+08] Thienel, J., Luquette, R., and Sanner, R. "Estimation of spacecraft inertia parameters". In: *AIAA Guidance, Navigation and Control Conference and Exhibit*. 2008, p. 6454.

[TB94] Tsang, K. and Billings, S. "Identification of continuous time nonlinear systems using delayed state variable filters". In: *International Journal of Control* 60.2 (1994), pp. 159–180.

[Ttv] *TTVS*. CNES, 2019.

[VHV89] Van Huffel, S. and Vandewalle, J. "Comparison of total least squares and instrumental variable methods for parameter estimation of transfer function models". In: *International journal of control* 50.4 (1989), pp. 1039–1056.

[VHV91] Van Huffel, S. and Vandewalle, J. *The total least squares problem: computational aspects and analysis*. Vol. 9. Siam, 1991.

[Van+04] VanDyke, M. C., Schwartz, J. L., and Hall, C. D. "Unscented Kalman filtering for spacecraft attitude state and parameter estimation". In: *Advances in the Astronautical Sciences* 118.1 (2004), pp. 217–228.

[Vee+13] Veen, G. Van der, Wingerden, J.-W. van, Bergamasco, M., Lovera, M., and Verhaegen, M. "Closed-loop subspace identification methods: an overview". In: *IET Control Theory & Applications* 7.10 (2013), pp. 1339–1358.

[WP90] Walter, É. and Pronzato, L. "Qualitative and quantitative experiment design for phenomenological models - a survey". In: *Automatica* 26.2 (1990), pp. 195–213.

[Wei+12] Weiss, A., Leve, F., Kolmanovsky, I. V., and Jah, M. "Reaction wheel parameter identification and control through receding horizon-based null motion excitation". In: *Itzhack Y. Bar-Itzhack Memorial Symposium on Estimation, Navigation, and Spacecraft Control*. Springer. 2012, pp. 477–494.

[Wei+15] Weiss, A., Leve, F., Kolmanovsky, I. V., and Jah, M. "Reaction wheel parameter identification and control through receding horizon-based null motion excitation". In: *Advances in Estimation, Navigation, and Spacecraft Control*. Springer, 2015, pp. 477–494.

[Wer12] Wertz, J. R. *Spacecraft attitude determination and control*. Vol. 73. Springer Science & Business Media, 2012.

[Wil+02] Wilson, E., Lages, C., and Mah, R. "On-line gyro-based, mass-property identification for thruster-controlled spacecraft using recursive least squares". In: *The 2002 45th Midwest Symposium on Circuits and Systems, 2002. MWSCAS-2002.* Vol. 2. IEEE. 2002, pp. II–II.

[Yoo+17] Yoon, H., Riesing, K. M., and Cahoy, K. "Kalman filtering for attitude and parameter estimation of nanosatellites without gyroscopes". In: *Journal of Guidance, Control, and Dynamics* 40.9 (2017), pp. 2272–2288.

[You81] Young, P. "Parameter estimation for continuous-time models—a survey". In: *Automatica* 17.1 (1981), pp. 23–39.

[You08] Young, P. C. "The refined instrumental variable method". In: *Journal Européen des Systemes Automatisés* 42.2-3 (2008), pp. 149–179.

[You11] Young, P. C. *Recursive estimation and time-series analysis: An introduction for the student and practitioner*. 2nd. Springer, 2011.

[You+08] Young, P. C., Garnier, H., and Gilson, M. "Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models". In: *Identification of continuous-time models from sampled data*. Springer, 2008, pp. 91–131.

[Zha+17] Zhai, K., Wang, T., and Meng, D. "Optimal excitation design for identifying inertia parameters of spacecraft". In: *Acta Astronautica* 140 (2017), pp. 370–379.

Résumé

Cette thèse porte sur la détermination des paramètres du modèle dynamique d'un satellite pour son contrôle à l'aide de la théorie de l'identification des systèmes. Plusieurs algorithmes d'estimation paramétrique sont développés et adaptés à différentes configurations de mesures de données (gyroscope ou star tracker). Ces algorithmes permettent d'estimer la matrice d'inertie du satellite ainsi que l'alignement des actionneurs à partir des données de télémétrie. Pour réaliser cette estimation, la méthode de variable instrumentale est privilégiée. Des filtres sont élaborés afin d'améliorer significativement la précision de l'estimation, et ce même en présence de bruits de capteur et de perturbations au niveau des actionneurs. Les performances des algorithmes proposés sont analysées et validées à l'aide de simulations Monte Carlo à partir de données issues d'un simulateur haute-fidélité du CNES. La deuxième contribution concerne l'optimisation de la richesse des manœuvres réalisées par le satellite tout en respectant les contraintes physiques du système. L'efficacité des nouvelles trajectoires proposées est démontrée d'une part via des simulations de Monte Carlo et d'autre part à l'aide de tests effectués lors d'un vol zéro gravité en avion.

Mots-clés: Satellites, Estimation des Paramètres, Plan d'expérience, Variable Instrumentale, Modèle boîte grise, Traitement du Signal.

Abstract

The system identification for satellite attitude control is investigated in this thesis. Several parameter estimation algorithms are developed and adapted to the different types of sensor (gyroscope or star tracker). These algorithms allow to estimate, from the telemetry data, the satellite inertia matrix as well as the actuator alignments. For these estimation algorithms, an instrumental variable approach is considered. Filters are designed in order to significantly improve the accuracy and precision of the estimates, even in presence of sensor noise and disturbance torques. The performances of the proposed algorithms are analyzed and validated via Monte Carlo simulations using data from a high-fidelity simulator from CNES. The second main contribution concerns the optimization of maneuvers to improve the information content in the data, while respecting the physical constraints of the satellite. The effectiveness of the generated trajectory is evaluated both via Monte Carlo simulations and through real experiments in a zero-gravity environment.

Keywords: Satellites, Parameter Estimation, Experiment Design, Instrumental Variable, Grey-box Model, Signal Processing.