To my family
Abstract

This thesis explores the attitude estimation and control problem of a magnetically controlled small satellite in initial acquisition phase. During this phase, large data uncertainties pose estimation challenges, while highly nonlinear dynamics and inherent limitations of the magnetic actuation are primary issues in control. We aim to design algorithms, which can improve performance compared to the state of the art techniques and remain tractable for practical applications.

Static attitude estimation, which is an essential part of a satellite control system, uses vector information and solves a constrained weighted least-square problem. With large data uncertainties, this technique results in large errors rendering divergence or infeasibility in dynamic filtering and control. When static estimation is the primary source of attitude, these errors become critical; for example in low budget small satellites. To address this issue, we formulate a robust static estimation problem with norm-bounded uncertainties, which is a difficult optimization problem due to its unfavorable convexity properties and nonlinear constraints. By deriving an analytical upper bound for the convex maximization, the robust min-max problem is approximated with a minimization problem with quadratic cost and constraints (a QCQP), which is non-convex. Semi-definite relaxation is used to upper bound the non-convex QCQP with a semi-definite program, which can efficiently be solved in a polynomial time. Furthermore, it is shown that the derived upper bound has no gap in solving the robust problem in practice.

Semi-definite relaxations are also applied to solve the robust formulations of a more general class of problems known as the orthogonal Procrustes problem (OPP). It is shown that the solution of the relaxed OPP is exact when no uncertainties are considered; however, for the robust case, only a sub-optimal solution can be obtained.

Finally, a satellite rate damping in initial acquisition phase is addressed by using nonlinear model predictive control (NMPC). Standard NMPC schemes with guaranteed stability show superior performance than existing techniques; however, they are computationally expensive. With large initial rates, the computational burden of NMPC becomes prohibitively excessive. For these cases, an algorithm is presented with an additional constraint on the cost reduction that allows an early termination of the optimizer based on the available computational resources. The presented algorithm significantly reduces the de-tumbling time due to the imposed cost reduction constraint.
Acknowledgements

First of all, I would like to thank my supervisor Dr Eric Kerrigan for his constant support and guidance during my PhD studies. I would like to express my gratitude to him for his encouragement to take decisions during my research, which helped me to establish trust and self-confidence in my abilities. He always motivated me to explore the depths of my research area, while his advices and suggestions helped me to remain focused. His great vision enabled me to see the benefits of his advices in the long run. I am also grateful to Dr Imad Jaimoukha for introducing me to the field of semidefinite relaxations and helping me understanding many practical aspects. He remained a source of encouragement for me during my PhD studies and always showed me a ray of hope whenever I was stuck in my research. It was a great honor and valuable experience to collaborate with such a nice personality. I am also thankful to my examiners Dr James Whidborne and Professor Alessandro Astolfi for their valuable suggestions and comments which helped a lot to improve the thesis.

I would also like to acknowledge all past and present members of our research group for giving feedback on my research and sharing many new ideas. In particular, I want to thank Ammar Hassan, Edo Abraham and Paola Falugi for many fruitful discussions and critical feedback. In addition to the research group, I feel myself lucky to have a group of very nice friends who made my stay a memorable one. I am thankful to all of them for giving me such a good company. I would especially like to thank Abid Rafique, Furqan Tahir and Anas Imtiaz for many useful discussions and help in thesis proof reading. I wish you all a very successful career in your life.

I am also thankful to the Higher Education Commission (HEC) of Pakistan to provide me financial support for my PhD studies.

Finally, my deepest gratitude to my parents who always encouraged me to learn, to my brother and sister for their well wishes and prayers, and most importantly to my wife Sadaf and our beloved daughter Sameen, for their endless love, support and understanding.
Declaration

I hereby declare that this thesis is the result of my own work. Ideas and quotations from the work of other people, published or otherwise, are fully acknowledged.

Shakil Ahmed
Imperial College London
October 2012
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contents</strong></td>
<td>5</td>
</tr>
<tr>
<td><strong>List of Figures</strong></td>
<td>9</td>
</tr>
<tr>
<td><strong>List of Tables</strong></td>
<td>11</td>
</tr>
<tr>
<td><strong>Notation</strong></td>
<td>12</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>15</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>15</td>
</tr>
<tr>
<td>1.2 Review</td>
<td>19</td>
</tr>
<tr>
<td>1.3 Thesis Organization</td>
<td>22</td>
</tr>
<tr>
<td>1.4 Publications</td>
<td>24</td>
</tr>
<tr>
<td><strong>2 Satellite Dynamics</strong></td>
<td>26</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>26</td>
</tr>
<tr>
<td>2.2 Frames and Coordinate Systems</td>
<td>27</td>
</tr>
<tr>
<td>2.3 Nonlinear Equations of Motion</td>
<td>29</td>
</tr>
<tr>
<td>2.3.1 Translational Dynamics</td>
<td>30</td>
</tr>
<tr>
<td>2.3.2 Rotational Dynamics</td>
<td>31</td>
</tr>
<tr>
<td>2.3.3 Kinematics</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Sensor Modeling</td>
<td>33</td>
</tr>
<tr>
<td>2.4.1 Earth Magnetic Field Modeling</td>
<td>34</td>
</tr>
<tr>
<td>2.4.2 Sun Vector Modeling</td>
<td>37</td>
</tr>
<tr>
<td>2.5 Actuator Modeling</td>
<td>39</td>
</tr>
<tr>
<td>2.6 A Closed Loop Simulation Environment</td>
<td>39</td>
</tr>
<tr>
<td>2.6.1 Initial Acquisition Phase Control</td>
<td>39</td>
</tr>
<tr>
<td>2.6.2 Normal Operation Phase Control</td>
<td>42</td>
</tr>
</tbody>
</table>
## CONTENTS

### 2.6.3 Summary

### 3 Estimation Problems with Orthogonality Constraint

#### 3.1 The Attitude Estimation Problem

##### 3.1.1 Definition of Attitude

##### 3.1.2 The Optimization Problem

##### 3.1.3 A Classical Solution

#### 3.2 The Orthogonal Procrustes Problem

##### 3.2.1 An SVD-based Solution

#### 3.3 Semidefinite Relaxations

##### 3.3.1 A Semidefinite Programming Problem

##### 3.3.2 The S-procedure

#### 3.4 Summary

### 4 Robust Attitude Estimation

#### 4.1 Robust Problem Description

#### 4.2 Uncertainty Model

#### 4.3 An Approximation in the Robust Formulation

##### 4.3.1 Comparison of the Analytical Upper Bound

#### 4.4 Addition of a Regularization Term

#### 4.5 Approximate Robust Formulation

#### 4.6 Performance Analysis of the Robust Formulation

##### 4.6.1 Tightness of the Analytical Upper Bound

##### 4.6.2 Robust Performance Comparison for Input Data at One Time Instant

##### 4.6.3 The Effect of the Regularization Term

##### 4.6.4 Robust Performance Comparison for In-Orbit Simulation Data

#### 4.7 Summary

### 5 Solution of the Robust Estimation Problem using Semidefinite Relaxation

#### 5.1 Semidefinite Relaxation for Non-convex QCQP Problems

#### 5.2 Semidefinite Relaxation for the Robust Estimation Problem

#### 5.3 Finding the Robust Quaternion \( (q^*) \)

##### 5.3.1 Relaxation Gap

#### 5.4 Numerical Results
5.4.1 Comparison of the Quaternion Obtained from the Approximate Problem and the SDR ........................................... 74
5.4.2 Comparison of the Solutions for In-Orbit Simulation Data .................................................. 75
5.4.3 Illustration of Sections 5.2 and 5.3: Theoretical Results .................................................. 75
5.5 Summary .................................................. 76

6 Orthogonal Procrustes Problem with Data Uncertainties 78
6.1 Semidefinite Relaxation for the Orthogonal Procrustes Problem .................................................. 78
  6.1.1 Relaxation of the Standard OPP .................................................. 79
  6.1.2 Relaxation of the OPP for Rotations .................................................. 82
6.2 The OPP with Data Uncertainties .................................................. 83
  6.2.1 The Robust Problem .................................................. 84
  6.2.2 Uncertainty Representation in the Data Matrices .................................................. 84
  6.2.3 Semidefinite Relaxation of the Robust OPP .................................................. 84
  6.2.4 Orthogonalization of $X$ .................................................. 86
  6.2.5 Effect on Robust Performance .................................................. 86
  6.2.6 The Robust Problem for Rotations .................................................. 87
6.3 Simulation Results .................................................. 87
  6.3.1 Analysis of the Standard OPP Relaxation .................................................. 87
  6.3.2 Analysis of the OPP Relaxation for Rotations .................................................. 88
  6.3.3 Robust Performance Evaluation .................................................. 88
  6.3.4 Orthogonalization Step .................................................. 90
6.4 Summary .................................................. 90

7 Sub-optimal Predictive Control for Satellite De-tumbling 92
7.1 Introduction .................................................. 92
7.2 Optimal NMPC Formulations .................................................. 93
  7.2.1 Formulation with Terminal Constraint .................................................. 94
  7.2.2 Formulation without Terminal Constraint .................................................. 95
7.3 Sub-optimal NMPC Formulation .................................................. 96
7.4 Satellite Dynamics .................................................. 98
7.5 Simulation Results .................................................. 100
  7.5.1 Performance Comparison between Solving Problems $F_1$ and $F_2$ .................................................. 100
  7.5.2 Comparison with $\beta$-dot Control .................................................. 101
  7.5.3 Suboptimal MPC Performance .................................................. 102
List of Figures

1.1 A comparison of the Earth’s magnetic field strength obtained from the $10^{th}$ order and the $1^st$ order IGRF models against the geocentric latitude. 17
1.2 A comparison of actual attitude and its estimate using a standard algorithm after addition of normally distributed error up to 40% in the measurements. 17

2.1 Satellite frames. 27
2.2 The Earth’s magnetic field against the geocentric latitude. 35
2.3 The Earth’s magnetic field against longitude. 35
2.4 Comparison of the first order Earth’s magnetic field in $F_o$ with the analytical expression of (2.21). 36
2.5 The Sun and the Earth geometry. 38
2.6 Performance of the controller based on derivative of the Earth’s magnetic field. 41

4.1 Comparison of the analytical bound given in (4.10) and the bound obtained by solving the semidefinite program (4.16). The bound difference is given as a percentage of relative error between the SDR and the analytical bounds. 62
4.2 Histogram showing the distribution of the cost for the robust and non-robust solutions. The plot shows data for 2000 runs. In each run, uniformly distributed random error is added in the test vectors within the given uncertainty bounds. 63
4.3 The effect of the change of the tuning parameter $\eta$ on robust performance. 64
4.4 A comparison of the attitude angles obtained using the robust and non-robust algorithms. The dotted line shows the original data without errors while the other two cases include errors within the chosen uncertainty bound. 65
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Difference between the attitude obtained by solving (4.18) and its semidefinite relaxation. In the right-hand side plots, $e_\phi, e_\theta$ and $e_\psi$ are errors in the Euler angles.</td>
</tr>
<tr>
<td>5.2</td>
<td>Eigenvalues of the matrix $L_{1,1}$ matrix.</td>
</tr>
<tr>
<td>5.3</td>
<td>Gap between the maximum cost of (4.18) and its upper bound $\bar{\gamma}^*$ obtained using the solution of (5.3).</td>
</tr>
<tr>
<td>5.4</td>
<td>Difference between the quaternions obtained by solving (4.18) and its semidefinite relaxation.</td>
</tr>
<tr>
<td>6.1</td>
<td>Simulation results for the semidefinite relaxation of the standard OPP (6.4).</td>
</tr>
<tr>
<td>6.2</td>
<td>Simulation results for the semidefinite relaxation of the OPP for rotations (6.19).</td>
</tr>
<tr>
<td>6.3</td>
<td>Evaluation of the robust performance.</td>
</tr>
<tr>
<td>6.4</td>
<td>The first subplot shows number of orthogonal solutions (%) versus the uncertainty bounds, and second subplot compares cost using the SDR solution before and after orthogonalization.</td>
</tr>
<tr>
<td>7.1</td>
<td>Performance comparison between solving $F_1$ and $F_2$.</td>
</tr>
<tr>
<td>7.2</td>
<td>Rate damping performance between solving $F_1$ and using $\beta$-dot control.</td>
</tr>
<tr>
<td>7.3</td>
<td>Performance of the IPOPT solver while solving NMPC problems $F_1$ and $F_2$ with initial rates $P = 6.9$ deg/s, $Q = -9.7$ deg/s and $R = 6.3$ deg/s. Comparison is given for two cases, i.e., one with maximum iterations set to 300 and second with 50.</td>
</tr>
<tr>
<td>7.4</td>
<td>Performance comparison between solving $F_2$ and $F_3$ for maximum iterations set to 50.</td>
</tr>
<tr>
<td>7.5</td>
<td>Performance of sub-optimal MPC with different maximum iterations (left) and $\gamma$ (right).</td>
</tr>
<tr>
<td>7.6</td>
<td>Rotational energy reduction of a well-tuned $\beta$-dot controller compared to $F_2$ and Algorithm 1.</td>
</tr>
<tr>
<td>A-1</td>
<td>Geometry of the Earth and the satellite orbits</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Notation used to define state vector. ......................... 30
2.2 Data used in the nonlinear simulation. ....................... 40

4.1 Vector data without errors. Vectors in the body frames are represented by
$\mathbf{b}_i$ and in the reference frames by $\mathbf{r}_i$. The index $i$ is 1 for Earth magnetic
field measurement and 2 is for the sun sensor data. ................. 61
4.2 Vector data with added error. Vectors in the body frames are represented
by $\mathbf{b}_i$ and in the reference frames by $\mathbf{r}_i$. ...................... 63

5.1 Comparison of the elements of quaternion $q_i, i = 1, 2, \ldots, 4$ obtained from (4.18)
and (5.3) for the vector set given in Table 4.2 ....................... 77

7.1 Simulation parameters for MPC and $\beta$-dot performance comparison. ... 101
Notation

- **small letters vectors and scalars**
- **capital letters matrices**
- $\mathbb{R}$ set of real numbers
- $A := B$ $A$ is defined by $B$
- $| \cdot |$ absolute value of a scalar
- $\| \cdot \|$ vector norm
- $\| x \|_1$ 1-norm of the vector $x : \| x \|_1 := \sum_{i=1}^{n} |x_i|$ 
- $\| x \|_2$ 2-norm of the vector $x : \| x \|_2 := \sqrt{x^T x}$ 
- $\| x \|_p$ weighted 2-norm of the vector $x : \| x \|_p := \sqrt{x^T Px}$ 
- $\| x \|_\infty$ infinity-norm of the vector $x : \| x \|_\infty := \max_i (|x_i|)$ 
- $\| A \|_F$ Frobenious-norm of the matrix $A : \| A \|_F := \sqrt{\text{tr}(A^T A)}$ 
- $I_n$ identity matrix in $\mathbb{R}^{n \times n}$ 
- $0_{n \times m}$ a matrix of $n$ rows and $m$ columns with zero entries 
- $1_{n \times m}$ a matrix of $n$ rows and $m$ columns with unit entries 
- $A^T$ transpose of the matrix $A$ 
- $\det(A)$ determinant of the matrix $A$ 
- $\text{tr}(A)$ trace of the matrix $A$ 
- $\mathcal{N}(A)$ null space of the matrix $A$ 
- $\dim(A)$ dimension of the matrix $A$ 
- $\text{diag}(a_1, \ldots, a_n)$ a matrix having only non-zero diagonal elements $a_1, \ldots, a_n$. 
- $\lambda(A)$ eigenvalues of the matrix $A$ 
- $\sigma(A)$ singular values of the matrix $A$ 
- $A \succeq 0$ matrix $A$ is positive semidefinite 

**Dynamics (Chapter 2)**

$F_b$ body frame 
$F_i$ inertial frame 
$F_e$ earth frame 
$F_o$ orbit frame 
$p$ position vector 
$v$ velocity vector 
$\omega$ rotational velocity vector 
$q$ quaternion
x \quad \text{state vector in } \mathbb{R}^{13} \text{ (translational and rotational dynamics)}

m \quad \text{control vector in } \mathbb{R}^{3}

P,Q,R \quad \text{body rates with respect to } F_o

\beta \quad \text{earth magnetic field vector in } \mathbb{R}^{3}

\tau_c \quad \text{control torque in } \mathbb{R}^{3}

\tau_{gg} \quad \text{gravity gradient torque in } \mathbb{R}^{3}

\psi_e \quad \text{geocentric latitude}

l_e \quad \text{geocentric longitude}

J \quad \text{moment of inertia matrix in } \mathbb{R}^{3 \times 3}

g \quad \text{gravity in } \mathbb{R}^{3}

\omega_o \quad \text{orbit rotation rate}

i \quad \text{orbit inclination}

\text{Estimation (Chapters 3-6)}

b \quad \text{information vector in } F_b

r \quad \text{information vector in } F_r

B(b) \quad \text{uncertainty set for measurement vector } b

R(r) \quad \text{uncertainty set for reference vector } r

A, B, B, R \quad \text{data matrices}

\bar{A}, \bar{B}, \bar{B}, \bar{R} \quad \text{uncertain data matrices}

C \quad \text{transformation matrix in } \mathbb{R}^{3 \times 3}

\gamma \quad \text{uncertainty bound}

\bar{\gamma} \quad \text{upper bound}

\delta \quad \text{uncertainty vector}

\Delta \quad \text{uncertainty matrix}

J \quad \text{cost function}

\text{Control (Chapter 7)}

x \quad \text{state vector in } \mathbb{R}^{7} \text{ (only rotational dynamics)}

T \quad \text{prediction horizon}

\delta \quad \text{sample time}

V(\cdot) \quad \text{cost function}

\ell(\cdot) \quad \text{stage cost}

F(\cdot) \quad \text{terminal cost}

P, Q \quad \text{weighing matrices}

x(\cdot), u(\cdot) \quad \text{state and input trajectories}

x_{\text{shifted}}(\cdot), u_{\text{shifted}}(\cdot) \quad \text{shifted state and input trajectories for warm start}

u(\cdot, x(t), t) \quad \text{input trajectory with initial state } x(t) \text{ at time } t

x^*(\cdot, x(t), t) \quad \text{state trajectory for the control trajectory } u(\cdot) \text{ with initial state } x(t) \text{ at time } t

u^*(\cdot) \quad \text{optimal or sub-optimal control trajectory}

V^*(x(\cdot), t) \quad \text{optimal or sub-optimal value function for a known control trajectory } u^*(\cdot)

\mathbb{U} \quad \text{set of inputs}

X_f \quad \text{terminal state set}
<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECI</td>
<td>Earth Centered Inertial</td>
</tr>
<tr>
<td>ECEF</td>
<td>Earth Centered Earth Fixed</td>
</tr>
<tr>
<td>IGRF</td>
<td>International Geomagnetic Reference Field</td>
</tr>
<tr>
<td>LEO</td>
<td>Low Earth Orbit</td>
</tr>
<tr>
<td>LMI</td>
<td>Linear Matrix Inequality</td>
</tr>
<tr>
<td>OPP</td>
<td>Orthogonal Procrustes Problem</td>
</tr>
<tr>
<td>QCQP</td>
<td>Quadratically Constrained Quadratic Program</td>
</tr>
<tr>
<td>NMPC</td>
<td>Nonlinear Model Predictive Control</td>
</tr>
<tr>
<td>SDR</td>
<td>Semidefinite Relaxation</td>
</tr>
<tr>
<td>SDP</td>
<td>Semidefinite Program</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>WMM</td>
<td>World Magnetic Model</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Motivation
This research is motivated by the need to develop reliable attitude estimation and control algorithms for initial acquisition phase of a Low-Earth-Orbit (LEO) small satellite. These small satellites have strict size, weight and budget constraints. Due to these constraints, the satellites depend heavily on the Earth’s magnetic field for estimation and control. The main reasons of this dependence are its free availability giving benefits like economy in cost and weight, easy on-board management especially for magnetic actuation, and making operational life independent of fuel. However, these advantages are obtained at the price of a more complicated estimation and control task. These satellite applications are gaining increased popularity, both in the academic community [38; 49], as well as in many future space applications [6]. We are mainly interested in the static attitude estimation, which is an integral part of the satellite control system. We will also address the satellite de-tumbling control problem, where the main objective is to reduce the time to damp the high body rates induced due to launch disturbances to zero with less use of available power resources.

Static attitude estimation requires information of some vector quantities, such as the Earth’s magnetic field, sun and star directions, satellite position, each obtained from two different sources, e.g. a sensor and a mathematical model. In the initial acquisition phase, static estimation has to face increased level of uncertainties, where in addition to sensing and modeling errors, coupling between magnetic actuators and the Earth magnetic field (EMF) measurement results in a highly uncertain environment. In these conditions, the existing techniques, which are mainly based on solving a constrained weighted least-squares problem, can give potentially wrong attitude information. The
situation becomes worse when static estimation is the only or primary source of attitude information for control, which is mostly the case in these small satellites [85]. The large attitude error under worst uncertainties can significantly affect the performance of any dynamic filtering and control algorithms [25], e.g., Kalman filtering or model predictive control. These algorithms can face divergence or infeasibility issues, especially due to large initial estimation errors [42].

In the existing literature on static attitude estimation, statistical methods are generally used to analyze the sensitivity of the estimated attitude against measurement noise [76]. These studies mostly neglect modeling errors, which could be significant. Moreover, this analysis does not directly modify the obtained solution for uncertainties; although, some studies suggested optimal weighing schemes to achieve minimum variance for the weighted least-squares problem [55; 76]. Under worst-case uncertainties, the performance of these algorithms may significantly degrade. Consider for example, the Earth’s magnetic field, which is the most commonly used vector information in such applications. EMF can be measured by magnetometers installed on the body of the satellite, while different high fidelity models, such as the International Geomagnetic Reference Field (IGRF) [72], World Magnetic Model (WMM) [60] are used to predict the EMF. However, there is always some error between the Earth’s magnetic field given by these models and the actual field values. For example, the error between these higher order models and the actual field can be around 20% [72]. Furthermore, use of simple models such as the low order IGRF model [72], which are normally preferred due to lower computational cost, may result up to an additional 10% error. A comparison of such a high and low order models is shown in Figure 1.1. Moreover, in the early launch phase, interaction between the Earth’s magnetic field and magnetic actuators results in an increased noise level in magnetic field measurements. All such uncertainties lead to errors in the attitude estimate. Considering an uncertainty of 30% of the norm of the vector measurements, a nominal static attitude estimation algorithm gives large errors, as shown in Figure 1.2 for an in-orbit satellite simulation data. However, it is important to note that this performance degradation is in the worst-case scenario.

In general, to address the uncertainties in worst-case scenarios, a robust min-max problem need to be solved which will give the best uncertainty immunized attitude information. One can find much work on the unconstrained robust least-squares problem [19; 26; 27; 32; 69; 86]. However, the static attitude estimation problem, which is a weighted least-squares problem with nonlinear constraints, is less addressed for deter-
1.1 Motivation

Figure 1.1: A comparison of the Earth’s magnetic field strength obtained from the 10\textsuperscript{th} order and the 1\textsuperscript{st} order IGRF models against the geocentric latitude.

Figure 1.2: A comparison of actual attitude and its estimate using a standard algorithm after addition of normally distributed error upto 40% in the measurements.
ministic uncertainties. Consideration of robustness in min-max settings complicates the nonlinear constrained optimization problem, making it computationally intractable or very difficult to solve. In this thesis, we study how best we can solve this robust estimation problem, keeping in view the tractability for practical applications. We rely on approximations to obtain tractable formulations of the robust problem, which is otherwise hard to solve. We are mainly interested in developing algorithms that can give better attitude estimates in the presence of large uncertainties, reducing the likelihood of divergence in filtering or infeasibility in control algorithms due to estimation errors.

The type of constrained estimation problem we consider in attitude estimation is also closely related to a well-known mathematical problem known as the orthogonal Procrustes problem (OPP) [34; 39], which has diverse applications. Another motivation of this research is that robustness consideration in the OPP may be of interest to a wider class of audience. The orthogonal Procrustes problem deals with finding a geometrical transformation with an orthogonality constraint. In fact, the attitude determination problem is a subclass of OPP. Other formulations of the OPP can address rotations, reflections and translations and have applications in image processing, machine learning, computer vision and statistics. In image processing and machine learning, the OPP is used in pose estimation, which involves the estimation of an object’s position and orientation, either relative to the model reference frame, or at a previous time using a camera or a range sensor [37]. This application involves both orientation and translation. In statistics, the OPP is used in principal component analysis, a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. However, in this thesis, we are mainly interested in the formulations, which address rotations and reflections. Like the attitude estimation problem, in all these applications, the input data for the constrained least square problem (a camera image for example), is prone to uncertainties. When uncertainties are large, the robustness consideration becomes a primary requirement. In this thesis, following similar lines adopted for robust attitude estimation, we use semidefinite relaxations to solve the orthogonal Procrustes problem both with and without data uncertainties.

Lastly, we look at the de-tumbling control problem in the initial acquisition phase of a small satellite with magnetic actuation. Launch disturbances induce high body rates, which result in highly nonlinear dynamics. Due to nonlinearities and use of magnetic actuation, rate damping normally takes a long time with existing techniques. However,
1.2 Review

This phase is required to be short, since it may restrict many crucial deployments such as solar panels, communication antennas. Some of these deployments are critical from the operational point of view. For example, these small satellites depend on solar panels for electrical power in normal operation, while during the initial acquisition phase, as solar panels are not yet deployed, they have to rely on on-board batteries. This adds to the requirement for the de-tumbling phase to be short. At the same time it is also a requirement that the control system makes less use of batteries. In this research, we study nonlinear model predictive control (NMPC) to optimally solve this problem and compare its performance with exiting methods. However, one main concern with NMPC is high computational requirements. We are mainly interested to study how NMPC performance is affected if optimization is terminated early and how we can develop algorithms, which can give acceptable performance despite sub-optimality.

1.2 Review

As discussed in Section 1.1, one can find much work on the robust linear least-squares problems, such as [19; 20; 26; 27; 32; 69; 86]. However, there is less discussion on robustness against data uncertainties in the least-squares estimation problems with orthogonality and determinant constraints, such as the attitude estimation problem and other applications of the orthogonal Procrustes problem. In this section, we give a review of existing approaches to address uncertainties. We also discuss robustness in the least-squares problems. Lastly, we briefly mention some existing results in semidefinite relaxations, which are helpful to solve the estimation problems addressed in this thesis.

In the attitude estimation problem, especially in the context of satellite applications, it has been a point of concern since early days to analyze the sensitivity of the obtained solution against measurement noise. Since noise generally has a probabilistic description, performance of these algorithms is generally analyzed in a stochastic sense. Expressions for covariance propagation of the attitude transformation matrix are presented and analyzed in the literature. [76] presents covariance analysis for the TRIAD algorithm, one of the earliest algorithm used for static estimation [13], and compares it with the covariance expression of his new algorithm QUEST, which proved to be very successful in practical applications. [55] discusses both covariance analysis and sensitivity analysis for his singular values decomposition based algorithm. Most of the covariance studies, such as [23; 56], only consider errors in the measurement vectors, while completely neglecting errors in the modeled vectors. These studies only provide a way to analyze the varia-
tion in the estimated attitude against measurement errors but have no direct affect on the solution of the attitude problem. However, some studies suggested optimal weighing schemes to achieve minimum variance for the weighted least-squares problem [23]. [64] discusses measurement uncertainties in the problem formulation and presented a multiplicative uncertainty model, which could maintain a unit norm of the input information vectors. However, they considered uncertainties with a known probability distribution. They have formulated a modified measurement vectors and have presented algorithms to solve the standard Wahba problem [88]. Some algorithms, such as [24], also consider uncertainties in the input measurements in a stochastic framework, which is based on minimum variance recursive estimation. However, most of these studies mainly concentrate on measurement noise only, while neglecting all other possible error sources, such as sensor installation errors, offsets, biases, and mainly modeling errors, which could be significant.

The issue of robustness against measurement and modelling errors has also been explored in the orthogonal Procrustes problem. We refer to [37; 47], who have discussed robustness in pose estimation, which is used in many machine vision problems. Both these references also rely on statistical methods to deal with outliers. Reference [37] improves the least-squares estimation to make it robust against outliers by converting the estimation procedure to an iterative reweighed least-squares, where the weights for each observation depend on the residual error. They also used a nonlinear regression technique known as M-estimator and proposed some modifications to improve robustness. Techniques given in [47] also used M-estimation and LMS (least median of squares) of the residual error to design robust algorithms. Both references showed through experimental results that the robust techniques can effectively suppress outliers or mismatched points. An optimal framework for robust pose estimation is considered in [28]. They derived necessary conditions for $\ell_\infty$ optimality and show how to use them in a branch and bound setting to find the optimum and to detect outliers. They treated translations and rotations separately and derived efficient robust algorithms for both cases. Specifically for the orthogonal Procrustes problem, [79] has presented a perturbation bound for the rotation problems, which are tighter than existing bounds derived from the polar factors. The presented result relates sensitivity of the solution with the condition number of the information matrix. They also presented conditions for which their derived bounds become valid for the general OPP. All the mentioned references mostly address only measurement noise in a stochastic framework based on covariance of a given sensor.
1.2 Review

Robustness in the attitude estimation problem or the orthogonal Procrustes problem in a min-max setting is not much addressed in literature. However, as mentioned at the start of this review, robustness in un-constraint least-squares problems has been extensively addressed using min-max optimization. We mention here the seminal work by [20; 32], which also motivated us to address the robust attitude estimation problem on similar lines. [32] addressed the deterministic robustness of the least-squares problems in which the perturbations are deterministic and unknown but bounded. For unstructured perturbations, they minimized the worst-case residual error using convex second-order cone programming (SOCP), while for a known perturbation structure, they solved a semidefinite program (SDP). Both SOCP and SDP give an exact worst case value when perturbation enters linearly in the data matrices, and can be solved with good efficiency in practice. They also compared the robust least-squares (RLS) solution with the solution of total least-squares (TLS) [33] and gave conditions when the least-squares, the TLS and the RLS coincide. They also proposed to use the TLS in conjunction with RLS to get perturbation bounds. They link their RLS results with the regularized LS, such as Tikhonov regularization and proposed a method to choose regularization parameter, which is optimal for robustness by solving an SOCP. Similar results were derived by [20] for an unstructured RLS problem; however, they provided a nice geometric interpretation of the obtained solution.

We mention some work on solving the non-convex quadratic problems. One very effective method for such difficult problems is the S-procedure, which is used to convert the non-convex quadratic optimization problems or the robust min-max problems to tractable semidefinite programs. This approximation technique was first mathematically developed by [31; 95], and later on many researchers worked on its different extensions and applications. Reference [16] discusses many applications of the S-procedure. [32] also used this technique to solve the structured RLS problems. It is also widely used in recent work on robust optimization [10; 41].

Lastly, a brief discussion on control of the satellites using magnetic actuators is given. Use of the Earth magnetic field for satellite attitude control, although started in late 70’s [82], has gained much interest in the last two decades. Attitude control based on the Earth’s magnetic field poses many challenges, such as inherent under-actuation of the magnetic coils [12; 71], time varying control authority due to position dependent Earth magnetic field, highly nonlinear dynamics and constraints on control, computations and available power. For attitude control using magnetic actuators, the most commonly used
technique is known as the $\beta$-dot control [29; 82], which uses derivative of the Earth’s magnetic field and is based on the principle of kinetic energy reduction. Nonlinear control technique proposed by [50; 51; 78] use state feedback and mainly concentrate on ensuring stability. In [90], a sliding mode control technique has been used for the same problem. Some linear optimal control and model predictive control techniques have also been proposed [70; 71; 81; 91; 93], which are useful for the normal operational phase of the satellite. However, these linear schemes are not suitable during the tumbling phase due to high nonlinearities. Recently some work has been done using NMPC for de-tumbling control [14; 15]. They demonstrated an improvement in the rate damping performance over a classical proportional control scheme.

1.3 Thesis Organization

This section presents the structure of the thesis. We give a brief description of each chapter along with its main contributions.

Chapter 2: Satellite Dynamics

Chapter 2 provides necessary background, which is required to formulate the attitude estimation and control problem for a satellite, which are the main topics addressed in this thesis. We mainly discuss coordinate frames, satellite dynamics, and mathematical modelling of the Earth’s magnetic field and the sun vector. We also present a closed-loop simulation environment using a classical control technique to bring the satellite to an equilibrium from high body rates induced by launch disturbances. This framework is used both to motivate and test the developed algorithms.

Chapter 3: Estimation Problems with Orthogonality Constraint

Chapter 3 introduces two problems, which are considered with uncertainties in Chapters 4-6. Firstly, a satellite attitude determination problem, along with one solution strategy which provides basis for various algorithms used in many real-life satellite applications. The Orthogonal Procrustes Problem is introduced next, along with a solution based on singular value decomposition. Lastly, a brief discussion on semidefinite relaxations for non-convex optimization problems is presented, which is helpful to understand the solutions presented, mainly in Chapter 5 and 6.
1.3 Thesis Organization

Chapter 4: Robust Attitude Estimation

Chapter 4 develops an approximate robust attitude estimation problem. Firstly, a robust attitude estimation problem is presented which is based on a weighted least-squares approach with nonlinear constraints. An uncertainty model is introduced considering modeling errors, measurement noise, sensor biases and offsets as infinity-norm bounded uncertainties. Using this uncertainty model and a quaternion transformation, the robust min-max problem is simplified by replacing the matrix optimization variable to a quaternion vector. This transformation also removes one nonlinear constraint. The maximization part in the min-max problem is concave and so hard to solve. Hence, an analytical upper bound is proposed, which transforms the min-max problem into an approximate non-convex minimization problem. A new regularization scheme is also proposed to improve the robust performance. The usefulness of the proposed algorithm in the presence of uncertainties is evaluated with the help of numerical examples. This chapter is mainly based on [3; 5].

Chapter 5: Solution of Robust Estimation Problem using Semidefinite Relaxation

Chapter 5 presents a tractable method for solving the proposed robust problem, which is a non-convex quadratically constrained quadratic program, using semidefinite relaxation (SDR). The relaxed formulation is convex with a linear objective and linear matrix inequality constraints. It is also shown how to extract the robust attitude information from the SDR solution. Moreover, the optimality properties of the SDR solution are studied and it is theoretically shown that there is no gap between the approximate problem and its semidefinite relaxation under a condition, which is often satisfied in practice. Lastly, numerical simulations are presented to support the theoretical results. This chapter is mainly based on [4; 5].

Chapter 6: Orthogonal Procrustes Problem with Data Uncertainties

Chapter 6 discusses solution of the orthogonal Procrustes problem using semidefinite relaxations. We only consider the reflection and rotation problems. It is shown that the relaxation approach for the standard problem results in zero gap, giving optimal solution. A robust formulation for the OPP is also presented. Semidefinite relaxation is used to solve this problem. However, it is observed that the formulation with data uncertainties does not always give zero gap, resulting in a non-orthogonal solution. In
these cases, a procedure to find the nearest orthogonal matrix is presented. It is shown through simulation results that the frequency of occurrence of the non-orthogonal solution is related with the size of the uncertainty. The discussions presented in Chapter 6 are based on [2].

Chapter 7: Sub-optimal Predictive Control for Satellite De-tumbling

This chapter addresses the de-tumbling problem for a satellite in the initial acquisition phase. Two standard NMPC schemes are studies to analyze performance in comparison with a classical control scheme for rate damping. Both NMPC schemes give superior performance when initial rates are small. However, with larger rates, one scheme faces infeasibility, while the other scheme needs a large number of iterations to reach an optimal point, which is not not acceptable for the considered application. For such cases, a sub-optimal algorithm is proposed with an additional constraint on cost reduction, which allows an early termination of the optimizer. The performance of the sub-optimal algorithm is analyzed using numerical simulations.

Chapter 8: Conclusions and Future Directions

The last chapter summarizes the main contributions of this thesis and highlights possible future research directions.

1.4 Publications

Most of the work in this thesis is based on the following publications:

Journal Publications


Conference Publications

1.4 Publications


Chapter 2

Satellite Dynamics

This chapter presents dynamic equations for a small Low-Earth-Orbit (LEO) satellite with magnetic actuators. Firstly, a satellite geometry is introduced followed by dynamics and mathematical modeling for sensors and actuators. In the last part of the chapter, a classical control scheme to obtain a closed loop simulation setup is discussed. A nonlinear simulation, based on the presented dynamic model, is used to motivate, define and test the estimation and control problems addressed in this thesis.

2.1 Introduction

To develop dynamic equations we consider a low cost CubeSat [38; 85]: a pico-satellite moving in a circular orbit at an average altitude of 650 km above the Earth surface. For attitude determination only two measurements are available, namely the Earth’s magnetic field and the sun direction vector. The Earth’s magnetic field is measured with the magnetometers installed on the satellite body. We consider two sets of magnetometers, one is installed inside the satellite, which is mainly used in the post-launch phase when the satellite is recovering from launch disturbances, while the other is installed on an extended boom, which is deployed once the satellite has achieved an equilibrium. The sun vector is sensed by a pair of sun sensors also installed on the satellite body. For control, three magnetic actuators, called magnetorquers, are used, which generate electromagnetic field by passing current through the coils, which interact with the Earth’s magnetic field to generate a control torque. For on-board power, satellite mainly relies on sun energy in the normal operation phase, while in the initial acquisition phase, when solar panels are not deployed, total dependence is on on-board batteries, which is a limited resource.
2.2 Frames and Coordinate Systems

To develop the dynamic model, some coordinate frames need to be defined. For this, consider a basic geometry of a satellite moving in a circular orbit around the Earth’s center, as shown in Figure 2.1. Here $F_b$ is the body frame, $F_o$ is the orbit frame (reference frame in this case), $F_i$ is the inertial frame, also known as Earth-Centered Inertial (ECI) frame and $F_e$ is the Earth frame, also known as Earth-Centered, Earth-Fixed (ECEF) frame. Both $F_b$ and $F_o$ are required to define the estimation and control problem, $F_i$ is used for developing the equation of motion of the satellite and is also required by the Earth’s magnetic field model, while $F_e$ is used by the Earth’s gravity model. The coordinate systems attached to these frames are Cartesian, right-handed and orthogonal and are defined with the origin and three basis vectors. The formal definition of these frames for the satellite is being presented. Some related terminology, which is used in this description, is defined in Appendix A.

![Figure 2.1: Satellite frames.](image)

**Body Frame ($F_b$)**

The body frame is fixed within the body of the satellite with its origin at the satellite center of mass. The $z_b$-axis is aligned with an extended boom and directed towards the tip (see Figure 2.1). The other two axes $x_b$ and $y_b$ can be fixed with respect to
some installed instruments. Further, we assume that the body axes are aligned with the principal axes, which are defined as the axes about which the moment of inertia matrix is diagonal.

**Orbit Frame** \((F_o)\)

The orbit frame has its origin at the satellite’s center of mass, the \(z_o\)-axis aligns with the line joining the satellite center with the Earth’s center, the \(x_o\)-axis is tangent to the orbit plane in the direction of the orbital angular velocity vector for a circular orbit, which is the case being considered here, and the \(y_o\)-axis completes the right-handed axes system. This frame rotates at a constant rate \(\omega_o\) around the Earth’s center for circular orbits.

**Inertial Frame** \((F_i)\)

The ECI frame has its origin at the Earth’s center of mass. The coordinate system is defined with \(x_i\)-axis towards the Vernal Equinox (see Appendix A), the \(z_i\)-axis towards the celestial north pole and \(y_i\)-axis completes the right-handed axes system.

**Earth Frame** \((F_e)\)

The Earth frame is rotating and translating with the Earth’s center, also known as the Earth-Centered-Earth-Fixed frame. The coordinate system is defined with origin at the Earth’s center, the \(x_e\)-axis points towards the Prime Meridian (the zero longitude line at Greenwich), the \(z_e\)-axis points towards celestial North pole (aligned with \(z_i\)) and the \(y_e\)-axis completes the right-handed axes system. The ECEF frame rotates with the rotation rate of the Earth around its spin axis.

In the Earth pointing equilibrium state, both the body and the orbit frames are aligned. During the initial acquisition phase, the control system is required to bring the satellite to an equilibrium state as soon as possible. In normal operation, the control system maintains this equilibrium. The attitude of the satellite is used by the control system to determine the error between the body frame and the required equilibrium.

**Frame Transformation**

Transformations between these frames are frequently used during the development of the satellite dynamics, as well as for defining the estimation and control problem. We represent a transformation matrix from the coordinate system \(F_a\) to the coordinate system \(F_b\) with \(C_{b/a}\), especially if it is not clear from the context. As all the transformations are
2.3 Nonlinear Equations of Motion

orthogonal we can write $C_{b/a} = C_{a/b}^T$. We also use a quaternion to represent a transformation matrix. A quaternion is defined as $q \in \mathbb{R}^4 := [q_1 \ q_2 \ q_3 \ q_4]^T$ satisfying $q^T q = 1$, where $q \in \mathbb{R}^3 := [q_1 \ q_2 \ q_3]^T$ depends on the Euler axis of rotation, while $q_4 \in \mathbb{R}$ depends on the angle of rotation. Use of a quaternion for frame transformation not only avoids singularity issues, which are a major concern in the kinematic update equations based on Euler angles [80, eq. 1.3-22a], but also makes the transformation independent of the Euler angle sequence. They have some added advantages for solving attitude estimation problem, which will be discussed later in the thesis. A transformation matrix $C$ can be written in terms of a quaternion as [76, eq. 41, 42]

$$C := (q_4^2 - q^T q)I + 2qq^T + 2q_4Q,$$

where

$$Q := \begin{bmatrix} 0 & q_3 & -q_2 \\ -q_3 & 0 & q_1 \\ q_2 & -q_1 & 0 \end{bmatrix}.$$  

A quaternion $q$ corresponds to a specific transformation matrix. To avoid any confusion, when it is not clear from the context, we also represents a quaternion as $q_{b/a}$, which gives a transformation from frame $a$ to $b$. Different properties of the quaternion can be found in [77; 80].

2.3 Nonlinear Equations of Motion

This section presents both the rotational and the translational dynamics of the satellite. The translational dynamics is mainly required to find the position of the satellite in the orbit. This information is used both by the vector-based static attitude estimation as well as some control schemes. Both of these techniques use the Earth’s magnetic field, which is obtained from mathematical models only if the position of the satellite in the orbit is known.

The State and Control Vectors

The equation of motion is given as a set of first order ordinary differential equations given as

$$\dot{x}(t) = f (x(t), m(t), t),$$  

(2.3)
### Table 2.1: Notation used to define state vector.

<table>
<thead>
<tr>
<th>State Vector</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbf{p}_{CM/O}$</td>
<td>Position of the satellite center of mass with respect to the center of $F_i$ expressed in $F_i$</td>
</tr>
<tr>
<td>$\mathbf{v}_{CM/e}^b$</td>
<td>Velocity of the satellite center of mass with respect to $F_e$ expressed in $F_b$</td>
</tr>
<tr>
<td>$\omega_{b/o}^b$</td>
<td>Angular velocity of $F_b$ with respect to $F_o$ expressed in $F_b$</td>
</tr>
<tr>
<td>$\mathbf{q}_{b/o}$</td>
<td>Quaternion vector for transformation from $F_o$ to $F_b$</td>
</tr>
</tbody>
</table>

where the state vector $\mathbf{x}(t) \in \mathbb{R}^{13}$ and control $\mathbf{m}(t) \in \mathbb{R}^3$ are given as

$$
\mathbf{x}(t) := \begin{bmatrix} 
(p_{CM/O}^i)^T & (v_{CM/e}^b)^T & (\omega_{b/o}^b)^T & (q_{b/o})^T 
\end{bmatrix}^T,
$$

$$
\mathbf{m}(t) := \begin{bmatrix} m_x & m_y & m_z 
\end{bmatrix}^T.
$$

In the state vector, $\mathbf{p}_{CM/O}^i \in \mathbb{R}^3$, $\mathbf{v}_{CM/e}^b \in \mathbb{R}^3$, $\omega_{b/o}^b \in \mathbb{R}^3$ and $q_{b/o}$ represents quaternion for rotation from the orbit frame to the body frame. The subscript $CM$ represents satellite center of mass and $O$ is the origin of the inertial frame. The definition of these vectors is given in Table 2.1. In the control vector $\mathbf{m}(t)$, each component $m_x$, $m_y$ and $m_z$ represents dipole moment of the magnetic actuator installed about the corresponding axis of the body frame.

### 2.3.1 Translational Dynamics

The translational dynamics consists of state equations for both the position and the velocity of the satellite. Using Newton’s second law for the translational motion of the satellite, we can write

$$
\frac{1}{m} \sum f_{dis} + \tilde{g} = t \mathbf{v}_{CM/i},
$$

where $\sum f_{dis}$ represents the sum of the disturbance forces, $m$ is the mass of the satellite and $\tilde{g}$ represents the gravitation term force. Using the equation of Coriolis [80, eq. 1.2-7] we relate $\mathbf{v}_{CM/i}$ and $\mathbf{v}_{CM/e}$ and write the state equation for $t \mathbf{p}_{CM/O}^i$ in $F_i$ as

$$
t \mathbf{p}_{CM/O}^i = \mathbf{v}_{CM/i}^i = \mathbf{v}_{CM/e}^i + \omega_{e/i}^i \times \mathbf{p}_{CM/O}^i.
$$
2.3 Nonlinear Equations of Motion

where the left superscript represents the frame with respect to which the derivative is taken. Using the derivative of (2.6) in (2.5), we can derive the state equations for the body velocities of the satellite in \( F_b \), given as \([80, \text{eq. } 1.5-16d]\)

\[
b_{CM/e} \dot{v}^b_{CM/e} = \frac{1}{m} \sum f_{\text{dis}}^b + C_{b/i}^{\text{h}} \left( g^i - \left( \omega_{b/i}^i + \omega_{e/i}^i \right) \times v_{CM/e}^i \right),
\]

(2.7)

where \( \omega_{b/i}^i \) and \( \omega_{e/i}^i \) are the angular velocities of the \( F_b \) and the \( F_e \) frames with respect to the \( F_i \) frame, expressed in the body frame, while the gravitation term \( g^i = \tilde{g}^i - \omega_{e/i}^i \times (\omega_{e/i}^i \times p_{CM/O}^i) \), where \( \tilde{g}^i \) represents the Earth’s gravity, which is obtained from an Earth gravity model discussed next.

The Earth Gravity Model

For an oblate Earth the gravity \( \tilde{g}^e \) in \( F_e \) is given as \([80, \text{eq. } 1.4-16]\)

\[
\tilde{g}^e = \frac{-GM}{p^2} \left[ \begin{array}{c}
\{1 + 1.5 J_2 (a/p)^2 (1 - 5 \sin^2 \psi_e)\} p_x / p \\
\{1 + 1.5 J_2 (a/p)^2 (1 - 5 \sin^2 \psi_e)\} p_y / p \\
\{1 + 1.5 J_2 (a/p)^2 (3 - 5 \sin^2 \psi_e)\} p_z / p
\end{array} \right],
\]

(2.8)

where \( GM \) is the Earth-mass gravitational constant, \( p := [p_x \ p_y \ p_z]^T \) is the geocentric position vector, \( p \) is the length of the position vector, \( a \) is the semi-major axis of the Earth, \( \psi_e := \sin^{-1}(p_z/p) \) is the geocentric latitude and \( J_2 \) is the gravitational harmonic constant obtained from the Earth Gravitational Model (EGM96) coefficients. The gravity in \( F_e \) is then converted to \( F_i \) to be used in (2.7).

In (2.7), we also need \( \omega_{e/i}^i \) and \( \omega_{b/i}^i \), which are calculated using the rotational dynamic equations and the transformation matrix from \( F_i \) to \( F_b \), which is calculated using the quaternion updated at each time step. The disturbance force \( f_{\text{dis}} \) includes control magnetic disturbance, drag, solar pressure, however, we have not modeled these forces.

2.3.2 Rotational Dynamics

Using Newton’s law for the rotational momentum of a satellite, i.e. \( \sum \tau = \dot{\mathbf{h}} \), where \( \sum \tau \) represents the total torque and \( \dot{\mathbf{h}} \) is the rate of change of the angular momentum of the satellite with respect to \( F_i \), we derive the equation for the body rotational rates with respect to \( F_i \), expressed in the body frame, namely

\[
b_{CM/e} \dot{\omega}^b_{b/i} = (J^b)^{-1} \left[ \sum \tau^b - \omega^b_{b/i} \times J^b \omega^b_{b/i} \right],
\]

(2.9)
where $\sum \tau^b$ is the total torque acting on the satellite and is the sum of the gravity gradient, the control and the disturbance torques. We assume an \textit{axisymmetric} satellite with two of the principal moments of inertia of approximately equal size i.e. $J_x = J_y$, and zero product of inertias, resulting in the satellite principal axis aligned with the body axis. The inertial matrix in the body frame can be written as $J^b = \text{diag}(J_x, J_y, J_z)$.

To write the state equations for the satellite body rates with respect to the orbit frame i.e. $\omega^b_{b/o} := \begin{bmatrix} P & Q & R \end{bmatrix}^T$, we first relate $\omega^b_{b/o}$ with $\omega^b_{b/i}$, both expressed in $F^b$, as

$$\omega^b_{b/o} = \omega^b_{b/i} - C^b_{b/o} \omega^o_{o/i},$$

(2.10)

where $\omega^o_{o/i} := \begin{bmatrix} 0 & -\omega_o & 0 \end{bmatrix}$ is the rate of the orbit frame with respect to $F_i$, and is constant for circular orbits. By taking time derivative of both sides of (2.10), we write the rotational state equation with respect to the orbit frame as

$$\dot{\omega}^b_{b/o} = \dot{\omega}^b_{b/i} - \left( \frac{d}{dt} C^b_{b/o} \right) \omega^o_{o/i} - C^b_{b/o} \left( \frac{d}{dt} \omega^o_{o/i} \right) + \Omega^b_{b/o} C^b_{b/o} \omega^o_{o/i}.$$  

(2.11)

Gravity Gradient Torque

The gravity gradient torque in $F^b$ is given as [77]

$$\tau^b_{gg}(t) = \frac{3GM}{p^5} \left\{ \mathbf{p}^b_{CM/O} \times J^b \mathbf{p}^b_{CM/O} \right\}.$$  

(2.13)

Simplifying (2.13) we get the expression

$$\tau^b_{gg}(t) = \frac{3GM}{p^5} \begin{bmatrix} (J_z - J_y)p_y p_z \\ (J_x - J_z)p_z p_x \\ (J_y - J_x)p_x p_y \end{bmatrix}.$$  

(2.14)
Control Torque

The control torque is dependent on the dipole moment \( \mathbf{m}(t) \) of the three orthogonal coils installed on the body of the satellite and the Earth’s magnetic field vector \( \mathbf{\beta}(t) \in \mathbb{R}^3 \) measured by a magnetometer, and is given as

\[
\tau^b_c(t) = \mathbf{m}^b(t) \times \mathbf{\beta}^b(t). \tag{2.15}
\]

2.3.3 Kinematics

The kinematic equations are required to describe the rigid body orientation, for which we are using a quaternion. The differential equations for the quaternion representing transformation from the orbit frame to the body frame, are given as [77, eq. 4.7-13]

\[
\dot{q}_{b/o} = \frac{1}{2} \begin{bmatrix}
0 & R & -Q & P \\
-R & 0 & P & Q \\
Q & -P & 0 & R \\
-P & -Q & -R & 0
\end{bmatrix}
\begin{bmatrix}
q_1 \\
q_2 \\
q_3 \\
q_4
\end{bmatrix}, \tag{2.16}
\]

Euler Angles

In the attitude estimation problem we also use the Euler angles to compare performance. We briefly describe how we can relate them with a given transformation matrix. The Euler angles about each body axis \( x_b, y_b \) and \( z_b \) are denoted by \( \phi, \theta \) and \( \psi \) respectively. For the orbit to body transformation matrix \( C_{b/o} \), these angles can be calculated using the relations [80, eq. 1.5-21] for a \( \psi \rightarrow \theta \rightarrow \phi \) rotation sequence

\[
\begin{align*}
\phi &= \text{atan2}(c_{23}, c_{33}) \\
\theta &= -\text{asin}(c_{13}) \\
\psi &= \text{atan2}(c_{12}, c_{11}),
\end{align*} \tag{2.17}
\]

where \( c_{ij} \) is the element of the transformation matrix \( C_{b/o} \).

2.4 Sensor Modeling

As discussed before, we consider two types of sensors installed on the body of the satellite for attitude estimation, i.e. magnetometers and sun sensors. Both sensors give measurements in the body frame. In this section, we do not give the physical description of the
sensors, instead we only discuss their mathematical modeling. For the Earth’s magnetic field, we give an overview of the higher order mathematical models, their low order approximations and possible errors. For sun vector, we will give description of a simplified model, along with its error analysis.

### 2.4 Earth Magnetic Field Modeling

The Earth’s magnetic field $\beta$ measured at orbital altitudes is a combination of the magnetic fields generated by different sources. Generally, the Earth’s magnetic field is described by three components

$$
\beta(\psi_e, l_e, p, t) = \beta_m(\psi_e, l_e, p, t) + \beta_c(\psi_e, l_e, p) + \beta_d(\psi_e, l_e, p, t),
$$

where $\psi_e$ and $l_e$ are the geocentric latitude and longitude and $p$ is the magnitude of the position vector. $\beta_m$ is the main field due to the core of the Earth that comprises 95% of the field strength at Earth’s surface, $\beta_c$ is due to magnetized crustal rocks and $\beta_d$ is a disturbance component that can be as much as 10% of the main field and is mainly due to sources such as currents in the ionosphere and solar wind effects on the Earth’s magnetosphere [60]. The variations in $\beta_m$ and $\beta_d$ are function of the position and time, whereas the variation in time occurs at very large time scales. $\beta_c$ only depends on the position and its contribution to the total field decreases with the increase in altitude. This component is generally neglected at orbital altitudes. Different existing Earth’s magnetic field models, such as the International Geomagnetic Reference Field (IGRF)[72], the World Magnetic Model (WMM) [60] only model $\beta_m$, causing a possibility of error in the model output and the measured Earth’s magnetic field.

Magnetic fields are mathematically represented as a negative gradient of a potential function, which can be given as an infinite series of spherical harmonics. The magnetic field vector at the location of the satellite can be given as [60; 72]

$$
\beta(\psi_e, l_e, p) := -\nabla V(\psi_e, l_e, p) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} \beta_{nm}(g_{nm}, h_{nm}, \ldots),
$$

where $\nabla$ represents differentiation with respect to the position vector, $V$ is the potential function. The coefficient $\beta_{nm}$ represents the contribution of the spherical harmonics to the total magnetic field, which depends on $g_{nm}$ and $h_{nm}$, which are the Gauss coefficients of degree $n$ and order $m$ and are obtained from the Earth’s magnetic field models.
2.4 Sensor Modeling

Figure 2.2: The Earth’s magnetic field against the geocentric latitude.

Figure 2.3: The Earth’s magnetic field against longitude.
2.4 Sensor Modeling

Figure 2.4: Comparison of the first order Earth’s magnetic field in $F_o$ with the analytical expression of (2.21).

Lower Order Models

Higher order models provide good estimate of the Earth’s magnetic field, but are computationally expensive. The lower order models, such as the simple dipole or the first order IGRF model, which are computationally less expensive, describe almost 90% of the Earth’s magnetic field as compared to the full order model [6]. There are further 5-10% variations due to external effects. The first order IGRF approximation of the Earth’s magnetic field in the Earth frame is given as [72]

$$\beta(\psi, l, p, t) \approx \beta_{10} + \beta_{11} = \left( \frac{a}{p} \right)^3 [3(\mathbf{\hat{p}} \cdot \mathbf{s})\mathbf{\hat{p}} - \mathbf{s}],$$

(2.20)

where $\mathbf{s} = g_{11}\mathbf{\hat{e}}_1 + h_{11}\mathbf{\hat{e}}_2 + g_{10}\mathbf{\hat{e}}_3$, while $g_{11}, g_{10}$ and $h_{11}$ are the IGRF coefficients. A comparison of the Earth’s magnetic field for the first order and the 10th order IGRF model is given in Figures 2.2 and 2.3, the first one shows the effect of latitude variation, while the second shows the effect of longitude variation.

The Earth’s magnetic field in the orbit frame can approximately be represented using
2.4 Sensor Modeling

the analytical expressions [71]:

\[
\begin{bmatrix}
\beta_0 \\
\beta_y \\
\beta_z
\end{bmatrix} = \frac{GM}{p^3}
\begin{bmatrix}
\cos \omega_o t \sin i_m \\
-\cos i_m \\
2 \sin \omega_o t \sin i_m
\end{bmatrix},
\tag{2.21}
\]

where \(i_m\) is the inclination of the orbit plane with respect to the geomagnetic equator and \(\omega_o\) is the orbit rate. A comparison of the Earth’s magnetic field obtained from the first order IGRF with (2.21) for one orbit is shown in Figure 2.4 for \(72^\circ\) inclination. These models are very useful for simplified analysis or on-line control algorithms, which need the Earth’s magnetic field information to calculate control laws.

2.4.2 Sun Vector Modeling

In this section we are mainly interested in a simple mathematical model which can give fairly good sun vector information for attitude determination of the satellite. Sophisticated models are also available, but are based on the data developed by NASA for geocentric ephemeris of the sun for each year. These models give sun position as a function of the Julian date, which is a continuous count of the number of days since noon (12:00 UT) on January 1, 4713BC. These models also need information of the sun’s orbit parameters. However, in this thesis, we use a simplified model, mainly to avoid data availability issues, increased complexity and the fact that the error between the sophisticated and the simplified models is small and does not significantly affect the estimation and control design process. The chosen simplified model is based on two assumptions: the Earth’s orbit around the sun is circular with an orbit time of 365 days, and the satellite is at the center of the Earth. To analyze the error introduced by these assumptions in sun direction estimation, consider a time instant in which a line from the Earth’s center to the sun’s center is parallel to the ecliptic equator of the Earth, and the satellite is located directly above the geographic north pole. This geometry is shown in Figure 2.5(a). In this figure, \(p\) is the distance of the satellite from the Earth’s center, which is the sum of the Earth’s radius (6371 km) and the satellite altitude (650 km). The mean distance of the Earth’s center to the sun’s center is denoted by \(p_s = 1.496 \times 10^8\) km. Using these vector lengths, the error angle is \(\varepsilon = \tan^{-1}(p/p_s) \approx 4.693 \times 10^{-5}\) rad. This analysis shows that by considering the satellite at the center of the Earth introduces an error of only \(0.0027^\circ\), which is small compared to the attitude of the satellite.

To find the sun vector, consider Figure 2.5(b), where \(\xi_s\) is the elevation and \(\lambda_s\) is the
azimuth. Due to the Earth rotation axis not being perpendicular to the orbital plane, $\xi_s$ varies between $\pm 23^\circ$, with a period of 365 days approximately, and is zero on the first day of spring and fall. It can be calculated as $\xi_s = \left(\frac{23\pi}{180}\right) \sin \left(\frac{2\pi T_s}{365}\right)$ radians, where $T_s$ represents the number of days elapsed since the first day of spring (i.e. when the Earth passes the vernal equinox). The azimuth $\lambda_s$ is the angle between the line from the center of the Earth towards the sun on the first day of spring, and the same line at the current position of the sun, given as $\lambda_s = \frac{2\pi T_s}{365}$ radians. Using these two parameters, the sun vector in the ECI frame can be calculated as

$$s^i = \begin{bmatrix} \cos \lambda_s \cos \xi_s \\ \sin \lambda_s \\ \cos \lambda_s \sin \xi_s \end{bmatrix}, \quad (2.22)$$
Once the sun vector is obtained in the inertial frame, it can be easily converted to the orbit or the body frame using the transformations given in Appendix B. The transformed sun vector \( s^b \) and \( s^o \) are then used in sun sensor modeling and attitude determination.

Apart from the errors introduced due to simple modeling discussed above, the physical sun sensors based on solar cells can also have significant error due to neglecting the Earth albedo effect. The Earth albedo effect is generated due to light reflected by the Earth. This effect is significant for the LEO satellites and if not properly taken care of with good quality models, can introduce an error of around 15 degrees in the attitude determination [9; 11]. In this work, however, we have not considered the Earth’s albedo models.

### 2.5 Actuator Modeling

Three magnetic actuators, also known as magnetorquers, are installed about each axis of the body frame of the satellite. As given in \((2.15)\), the torque \((\tau)\) generated by these magnetorquers is the cross product of the magnetic dipole moment \((m)\) of the torquers and the Earth’s magnetic field \((\beta)\). The dipole moment about each axis is the product of the current passing through it \((I)\) and the area of the coil \((A)\) and is measured in \(Am^2\). The area of each coil is fixed and is a function of the wire diameter and the number of turns in each coil. The controller output needs to be transformed to the current for each magnetorqer in real satellite. This current generates a magnetic dipole moment which interact with the Earth’s magnetic field to generate the control torque.

### 2.6 A Closed Loop Simulation Environment

Finally, we present a closed-loop simulation environment based on the presented mathematical models. The data used in the simulation is given in Table 2.2 and is taken from [85]. To close the loop, control system is required to bring the tumbling satellite to an equilibrium state, a state initiated by launch disturbances. A classical control scheme for de-tumbling the satellite is now discussed.

#### 2.6.1 Initial Acquisition Phase Control

Attitude control of a satellite using the Earth’s magnetic field poses many challenges, such as inherent under-actuation of the magnetorquers, time varying control authority due to varying Earth’s magnetic field, system nonlinearities, especially during initial acquisition.
Table 2.2: Data used in the nonlinear simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbit Parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius</td>
<td>$p$</td>
<td>650</td>
<td>km</td>
</tr>
<tr>
<td>Inclination</td>
<td>$i$</td>
<td>72</td>
<td>deg</td>
</tr>
<tr>
<td>Eccentricity</td>
<td>$e$</td>
<td>0</td>
<td>deg</td>
</tr>
<tr>
<td>Mass Properties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimensions $(x \times y \times z)$</td>
<td>$10 \times 10 \times 30$</td>
<td>cm$^3$</td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>$m$</td>
<td>3.0</td>
<td>kg</td>
</tr>
<tr>
<td>Moment of inertia about x-axis</td>
<td>$J_x$</td>
<td>0.023001</td>
<td>kgm$^2$</td>
</tr>
<tr>
<td>Moment of inertia about y-axis</td>
<td>$J_y$</td>
<td>0.023565</td>
<td>kgm$^2$</td>
</tr>
<tr>
<td>Moment of inertia about z-axis</td>
<td>$J_z$</td>
<td>0.004197</td>
<td>kgm$^2$</td>
</tr>
<tr>
<td>Control Parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum dipole moment</td>
<td>$m_{\text{max}}$</td>
<td>0.1</td>
<td>Am$^2$</td>
</tr>
</tbody>
</table>

During the post-launch phase, initial disturbances lead to a continuously tumbling motion.

A commonly used control technique for de-tumbling phase uses derivative of the Earth’s magnetic field, and is known as $\beta$-dot control. One main reason of its common use is that it does not require body rates and attitude estimates; instead it only needs direct measurements of the Earth’s magnetic field in the body frame obtained from magnetometers. An early application of this concept can be found in \cite{82}. The control law is given as

$$m(t) = -K\dot{\beta}(t)$$ (2.23)

where $K \in \mathbb{R}^{3\times3}$ is a gain matrix and $\dot{\beta}$ is the rate of change of the Earth’s magnetic field in the body frame. One approach to design $K$ is given in \cite{29}. In this approach, the gain selection is based on the required decay time constant of the body rates; however this approach is based on simplified assumptions and hence the designed control does not exactly follow the design reference. Consider the gain calculation to generate a control torque $\tau$ along the $z$-axis of the body frame. To generate a torque about one axis, only two magnetorquers are required. We make two simplifying assumptions for the gain calculation. Firstly, the body is rotating about the $z$-axis of the body frame with a rate $R$ and secondly, the Earth’s magnetic field is confined only in the $x - y$ plane of the body frame. With these assumptions, we can write $\beta(t) := \begin{bmatrix} \beta_o \cos Rt & -\beta_o \sin Rt & 0 \end{bmatrix}^T$ and $m(t) := \begin{bmatrix} m_x & m_y & 0 \end{bmatrix}^T$, where $\beta_o$ is the magnitude of the Earth’s magnetic field vector.
2.6 A Closed Loop Simulation Environment

Figure 2.6: Performance of the controller based on derivative of the Earth’s magnetic field.

Using these relations and the control law 
\[ m_x := -k_x \dot{\beta}_x \quad \text{and} \quad m_y := -k_y \dot{\beta}_y, \]
we get the torque about the body z axis as 
\[ \tau_z = -k \beta_o^2 R. \]
Here we take \( k = k_x = k_y \). Using the torque equation for the rotational motion i.e. 
\[ \tau_z = \dot{h}_z = J_z \dot{R} \]
and solving, we get
\[ k = \frac{J_z}{J \beta_o^2}, \quad (2.24) \]
where \( J \) is the time constant for decay of \( R \) and can be chosen as a criteria to select \( k \). Finally, we take \( K = \text{diag}(k, k, 0) \). A performance of this type of control is shown in Figure 2.6 for an initial \( R = 2 \) deg/s and using \( k = 10^4 \), which is selected for a time constant of 270 sec. It can be observed from Figure 2.6 that actual decay time is different from the designed value, which is mainly due to deviation of the actual magnetic field from the assumptions made while deriving (2.24).

A similar approach to design the control gains is to keep the dipole moment \( m \) to its maximum value and selecting only the sign of the torque using the Earth’s magnetic field direction, making it a \textit{bang-bang} control. For one axis reorientation, this control law is considered time optimal at the expense of the maximum control energy [77]. Some other approaches for de-tumbling the satellite, such as [52; 92], are based on angular velocity feedback. The control law in [92] is given as \( m(t) = (K \omega(t)) \times \beta(t) \), while in [52] is
2.6 A Closed Loop Simulation Environment

\[ \mathbf{m}(t) = -\varepsilon K_v \omega(t), \]
where \( K > 0, K_v > 0, \varepsilon > 0 \) and \( \omega \) is the body rate with respect to the orbit frame. These control laws are closely related to the \( \beta \)-dot control and show almost similar performance.

2.6.2 Normal Operation Phase Control

Once the satellite approaches an equilibrium state, the rates remain small and the nonlinear system can be approximated with a linear system. In this operational phase, many nonlinear [50; 52] as well as linear [71; 91] control design techniques can be used. An overview of all these schemes is given in the survey paper [78]. However, we do not discuss this phase, as in this thesis we are mainly interested in the initial acquisition phase.

2.6.3 Summary

This chapter has presented mathematical modeling for the satellite to define the attitude estimation and control problems. We have given the definitions of the coordinate frames to be used in the dynamic equations, followed by a brief description of the translational and rotational dynamics. Mathematical modeling for the Earth’s magnetic field and the sun’s direction vector is presented with a discussion on uncertainties in different models. The effect of these uncertainties on attitude estimation, especially in the initial acquisition phase motivates to the study of robustness in the estimation problem, which is the main topic addressed in next three chapters. Similarly, the need to bring the body rates to zero in a minimum time with minimal use of power resources motivates to use model predictive control to address the conflicting objectives, which is the topic addressed in Chapter 7.
Chapter 3

Estimation Problems with Orthogonality Constraint

This chapter starts by defining a static attitude estimation problem, which is based on a weighted least-squares approach with an orthogonality and a determinant constraint. A commonly used approach to solve this problem is also presented. Another estimation problem known as the Orthogonal Procrustes Problem (OPP), also with orthogonality constraint, is introduced. This problem is a generalization of the attitude estimation problem and has diverse applications. A brief discussion on semidefinite relaxations is also included at the end.

3.1 The Attitude Estimation Problem

In this section a static attitude estimation problem for a satellite is discussed. This type of attitude determination has been widely used, not only in satellites [25; 76], but also in aerospace, marine and automotive systems [65]. The attitude is obtained by solving an optimization problem, based on a weighted least-squares approach with nonlinear constraints, known as the Wahba problem [88] in literature. This type of attitude estimation problem, which only depends on vector information, can be called static estimation, as opposed to dynamic estimation which depends on system dynamics. The static estimation could be useful for the nonlinear systems where dynamic filters could suffer divergence due to lack of a good a priori state estimate [25]. The attitude determined using a static approach can provide a reliable state initialization in these cases, reducing the likelihood of divergence.
3.1 The Attitude Estimation Problem

3.1.1 Definition of Attitude

To compute the attitude of a satellite, i.e. its orientation in space with respect to some known reference, two coordinate frames are needed, i.e. the body frame, which is fixed to the body of the satellite, and the orbit frame, which is taken as the reference frame in this thesis. In the earth pointing equilibrium state, both the body and the orbit frames are aligned. The attitude of the satellite provides the control system with the information to calculate the distance it is from the required equilibrium. A formal definition follows: The attitude of an object is defined as the coordinate transformation that transforms the reference coordinates into the body coordinates [77].

This transformation is obtained through a orthogonal transformation matrix \( C \in \mathbb{R}^{3\times3} \) with \( \det(C) = 1 \), introduced in Chapter 2. The matrix \( C \) can be obtained by solving an optimization problem (to be introduced shortly) subject to constraints, such as the orthogonality constraint \( C^T C = I_3 \) to preserve the length of the vectors, and the constraint on the determinant of the matrix \( C \) to preserve the frame orientation in a rotation. This includes the set of all rotation matrices in the special orthogonal group of rigid rotations in \( \mathbb{R}^{3\times3} \), denoted by \( SO(3) \) [21]. Many efficient solutions of this constrained least squares problem can be found in the literature, mostly developed for satellite applications [55; 56; 62; 63; 76]. Most of these algorithms are based on a quaternion transformation [45], which transforms the Wahba problem into an eigenvalue problem [76].

3.1.2 The Optimization Problem

The classical static attitude estimation problem is based on minimizing a weighted least squares cost, first proposed by Grace Wahba [88] for satellite applications, given as

\[
\min_C \frac{1}{2} \sum_{i=1}^{n} w_i \| b_i - C r_i \|_2^2 \\
\text{subject to } C^T C = I_3, \quad \det(C) = +1,
\]

(3.1)

where \( b_i \in \mathbb{R}^3 \) represents \( i^{th} \) measurement in the body frame, for \( i = 1, \ldots, n \), \( n \) being the total number of sensors, \( r_i \in \mathbb{R}^3 \) is the corresponding vector in the reference frame obtained from some model, \( w_i \in \mathbb{R} \) are non-negative weights.
3.1.3 A Classical Solution

One common approach used to solve (3.1) is to convert it into an equivalent maximization problem. Let $B := \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}$, $R := \begin{bmatrix} r_1 & r_2 & \ldots & r_n \end{bmatrix}$, $W := \text{diag}(w_1, w_2, \ldots, w_n)$, where $B, R \in \mathbb{R}^{3 \times n}$ and $W \in \mathbb{R}^{n \times n}$. Using this compact notation and expanding the cost function used in (3.1), we get

$$\frac{1}{2} \sum_{i=1}^{n} w_i \| b_i - Cr_i \|_2^2 = \frac{1}{2} \text{tr}(WB^T B + WR^T C^T CR) - \text{tr}(WB^T CR)$$

where we have used the condition $C^T C = I_3$, which is a constraint in the optimization problem. Neglecting the constant term, since it does not change the solution of the optimization problem, an equivalent maximization problem can be written as

$$\max_C \quad \text{tr}(WB^T CR)$$
subject to $C^T C = I_3, \ \det(C) = +1$. (3.2)

To solve this maximization problem, Davenport’s q-method \cite{45, 76} is commonly used. The method transforms the optimization variable from transformation matrix $C$ to quaternion $q := \begin{bmatrix} q^T & q_4 \end{bmatrix}^T$. Two important steps of the q-method are given now, which will be helpful to solve the robust problem addressed in the thesis.

**Step 1:** Find an equivalent formulation of (3.2) in terms of a quaternion. This new formulation, first reported in \cite{45}, states that the maximization of (3.2) is equivalent to the problem (see Appendix C for derivation):

$$\max_q \quad q^T K(B, R)q$$
subject to $q^T q = 1$, (3.3)

where $K : \mathbb{R}^{3 \times n} \times \mathbb{R}^{3 \times n} \rightarrow \mathbb{R}^{4 \times 4}$ is defined as

$$K(B, R) := \begin{bmatrix} (B(B, R))^T + B(B, R) - \text{tr}(B(B, R))I_3 & \mathbf{z}(B, R) \\ \mathbf{z}(B, R)^T & \text{tr}(B(B, R)) \end{bmatrix},$$

where $B : \mathbb{R}^{3 \times n} \times \mathbb{R}^{3 \times n} \rightarrow \mathbb{R}^{3 \times 3}$ and $\mathbf{z} : \mathbb{R}^{3 \times n} \times \mathbb{R}^{3 \times n} \rightarrow \mathbb{R}^{3}$ are defined as $B(B, R) := BW R^T$ and $\mathbf{z}(B, R) := (B \times R)We$, where $e$ is a vector of ones of appropriate dimension and $B \times R$ is the column wise cross product, i.e., if $b_i$ is the $i^{th}$ column of $B$ and $r_i$ is the $i^{th}$ column of $R$, then $b_i \times r_i$ is the $i^{th}$ column of the output matrix. Moreover, $K(B, R)$
is a symmetric and indefinite matrix, indicating that the objective function in (3.3) is neither concave nor convex.

Step 2: Using the new formulation, the non-concave maximization problem can be converted into an eigenvalue problem. For this we first add the constraint $q^T q = 1$ using a Lagrange multiplier $\lambda$ in (3.3), which gives

$$f(q, \lambda) = q^T K(B, R)q - \lambda(q^T q - 1). \quad (3.5)$$

To obtain a stationary point, we solve $\partial f / \partial q = 0$ and $\partial f / \partial \lambda = 0$ and obtain the following expression, which has the same formulation as a standard eigenvalue problem

$$K(B, R)q = \lambda q, \quad (3.6)$$

where $\lambda$ represents an eigenvalue of $K(B, R)$. Four eigenvectors of the symmetric matrix $K(B, R)$ are solutions of (3.6); however, the eigenvector corresponding to the maximum eigenvalue solves (3.3) [57; 76], i.e.

$$K(B, R)q_{opt} = \lambda_{max} q_{opt}, \quad (3.7)$$

where $q_{opt}$ is the solution to (3.3) and $\lambda_{max}$ is the maximum eigenvalue of $K(B, R)$. Most of the work in the static attitude estimation is based on this result and many efficient algorithms have been proposed, such as QUEST [76], ESOQ1 [62], ESOQ2 [63], especially for satellite applications. The survey paper [57] provides a general description of many of these algorithms.

The use of a quaternion in the new formulation simplifies the optimization problem, e.g., it replaces the optimization variables from a matrix to a four element vector and also it avoids the constraint $\det(C) = +1$ of (3.1). However, the main benefit obtained is the transformation of the optimization problem into an eigenvalue problem. The quaternion representation has a drawback, namely this representation is not unique, because $+q$ or $-q$ represents the same attitude. This issue needs to be taken care of in practical control applications, where quaternions are directly used in control algorithms.

### 3.2 The Orthogonal Procrustes Problem

The Orthogonal Procrustes Problem is a well-known mathematical problem [34; 39]. It deals with finding a geometrical transformation that involves rotations or reflections with
3.2 The Orthogonal Procrustes Problem

an orthogonality constraint. In simple words, given two arbitrary real valued matrices $A$ and $B$ of the same dimension, the optimal solution of the OPP is an orthogonal matrix $X$, which can best transform one matrix to the other, such that the Frobenius norm of the error $AX - B$ is minimized. There are many formulations of this problem, which can address rotations, reflections and translations having application in various areas, such as image processing, computer vision, statistics, satellites and aerospace. The attitude estimation problem, discussed in Section 3.1 is also a sub-problem of the OPP. In this thesis, we are mainly interested in the formulations which can address rotations and reflections. In Chapter 6, we address this problem using a semidefinite relaxation approach with data uncertainties.

The Orthogonal Procrustes Problem is mathematically defined as

\[
\min_X \|AX - B\|_F^2 \\
\text{subject to } XX^T = X^TX = I,
\]

(3.8)

where $A, B \in \mathbb{R}^{m \times n}$ are given data matrices, $m \geq n$ and $X \in \mathbb{R}^{n \times n}$ is the unknown orthogonal matrix, which belongs to an orthogonal group of order $n$, i.e.

\[
\mathcal{X} := \{X : XX^T = I\}.
\]

(3.9)

The determinant of $X$ is either $+1$ or $-1$. One important subclass of the Orthogonal Procrustes Problem includes an additional nonlinear constraint $\text{det}(X) = +1$ (see for example Wahba problem [88]). This problem deals specifically with rotations and has a wide range of applications. In these applications, we are only interested in $X \in SO(3)$, a special orthogonal group of order 3, defined as

\[
\mathcal{X}_+ := \{X : X \in \mathbb{R}^{3 \times 3}, XX^T = I, \text{det}(X) = +1\}.
\]

(3.10)

One can find many solutions of this problem in the literature [40; 55]. Most solutions are generally applications specific, satisfying some special requirements, such as computational efficiency, numerical stability. We present here a numerically robust solution based on the singular value decomposition (SVD) [39].

3.2.1 An SVD-based Solution

This section describe a commonly used solution for the OPP based on singular value decomposition [40]. To find an SVD-based solution, first we write the objective function
3.2 The Orthogonal Procrustes Problem

in (3.8) as

\[ \|AX - B\|_F^2 = \text{tr} \left((AX - B)(AX - B)^T\right) \]
\[ = \text{tr} \left(AXXTA^T + BB^T - AXB^T - BXTA^T\right). \]  
\[(3.11)\]

To simplify the expression we use the constraint \(XX^T = I\) in (3.11) and write \(AXXTA^T = AA^T\). Further, removing the constant terms, i.e. \(AA^T + BB^T\), we can write the maximization problem, which is equivalent to (3.8).

\[ \max_X \text{tr}(BXTA^T) \]
\[ \text{subject to } XX^T = I. \]  
\[(3.12)\]

To solve this problem, we use the permutation property of the trace operator \([33]\) and write the cost function as

\[ \text{tr}(BXTA^T) = \text{tr}(XTA^TB), \]
\[ = \text{tr}(XTUV^T), \]
\[ = \text{tr}(VTXU^T), \]
\[ \leq \text{tr}(\Sigma) = \sum_i \sigma_i. \]  
\[(3.13)\]

Here \(U\Sigma V^T\) is the singular value decomposition of the term \(A^TB\), where \(U, V\) are unitary matrices. The inequality in (3.13) becomes an equality when \(VTXU^T = I\), i.e. \(X = UV^T\), which is the required solution. The solution is unique if \(A^TB\) is nonsingular \([40, \text{Theorem 8.6}]\).

For the solution of the OPP for rotations with additional constraint \(\det(X) = +1\) \((X \in SO(3))\), to maximize (3.13) we use some properties of the determinant operator \([33]\) and write

\[ \det(VTXU^T) = \det(U^TXV), \]
\[ = \det(U^TV) \det(X) \]
\[ = \det(U^TV) = \pm 1. \]  
\[(3.14)\]

A unified solution for the cases when \(\det(U^TV) = \pm 1\) can be given as \([55; 79]\]

\[ X = U \text{diag}(1, 1, \det(U^TV))V^T. \]  
\[(3.15)\]

The sign of the last diagonal term, i.e., \(\det(U^TV)\) ensures \(\det(X) = 1\).
3.3 Semidefinite Relaxations

The semidefinite relaxation (SDR) technique is an effective tool to obtain approximate solutions of difficult optimization problems, such as non-convex and robust optimization problems, especially with quadratic objective and constraints. For these problems the exact global solution either does not exist or is difficult to compute [16; 53; 66]. In recent years application of the relaxation technique in different areas of engineering, such as robust optimization, control, signal processing, communications, aerospace, has attracted significant attention [10; 41; 54; 67; 84]. The solution of the relaxed problem is either exact to the optimal solution or is a good approximation; however in both cases, the relaxed problem can be written as a semidefinite program (SDP), which has a linear cost and linear matrix inequality constraints and can be solved efficiently with existing interior point methods for SDP problems [7]. In this section first we give a formal definition of a semidefinite programming problem, and then review a useful result to approximate the non-convex difficult optimization problems with a SDP.

3.3.1 A Semidefinite Programming Problem

A semidefinite program is a minimization problem of the form [32, Sec. 2]

\[
\min_x \quad c^T x \\
\text{subject to} \quad \mathcal{F}(x) \succeq 0,
\]

where \( x, c \in \mathbb{R}^m \) and \( \mathcal{F}(x) \succeq 0 \) is a linear matrix inequality (LMI) constraint on the vector \( x \) of the form

\[
\mathcal{F}(x) = \mathcal{F}_0 + \sum_{i=1}^{m} x_i \mathcal{F}_i \succeq 0,
\]

where the symmetric matrices \( \mathcal{F}_i = \mathcal{F}_i^T \in \mathbb{R}^{k \times k}, i = 1, \ldots, m \) are given. The dual problem to (3.16) is

\[
\max \quad - \text{tr} \mathcal{F}_0 Z \\
\text{subject to} \quad Z \succeq 0, \\
\text{tr} \mathcal{F}_i Z = c_i,
\]

where \( Z \) is a symmetric \( k \times k \) matrix and \( c_i \) is the \( i^{th} \) component of the vector \( c \). When both problems are strictly feasible (that is, when there exists \( x, Z \) which satisfy the constraints strictly), the existence of optimal points is guaranteed, and both problems have equal optimal objectives. In this case, the optimal primal-dual pairs \((x, Z)\) are those
pairs such that $x$ is feasible for the primal problem, $Z$ is feasible for the dual one, and $F(x)Z = 0$.

### 3.3.2 The S-procedure

The S-procedure is an elegant tool to obtain relaxations for non-convex and robust optimization problems [31; 95] and has applications in systems and control theory, robust estimation and control [16; 32]. A formal definition of the S-procedure is given in the following lemma [16, Page 23].

**Lemma 3.1.** Let $F_0, \ldots, F_p$ be quadratic functions of the variable $\zeta \in \mathbb{R}^m$, i.e.

$$F_i(\zeta) := \zeta^T T_i \zeta + 2u_i^T \zeta + v_i, \quad i = 0, \ldots, p,$$

where $T_i = T_i^T$. The following condition

$$F_i(\zeta) \succeq 0 \text{ for all } \zeta \text{ such that } F_i(\zeta) \succeq 0, \quad i = 1, \ldots, p,$$

holds if there exists $\tau_1 \geq 0, \ldots, \tau_p \geq 0$ such that

$$\begin{bmatrix} T_0 & u_0 \\ u_0^T & v_0 \end{bmatrix} - \sum_{i=1}^p \tau_i \begin{bmatrix} T_i & u_i \\ u_i^T & v_i \end{bmatrix} \succeq 0. \quad (3.21)$$

When $p = 1$, the converse holds, provided that there is some $\zeta_0$ such that $F_1(\zeta_0) \succeq 0$.

### An Example Problem

To elaborate the use of the S-procedure, we derive an approximate formulation for a non-convex quadratic optimization problem with a quadratic equality constraint, which satisfies conditions (3.20) and (3.21). We present a systematic procedure to derive this relaxed formulation. Later on, we follow similar steps to find a relaxation of a robust attitude estimation problem in Chapter 5.

Consider a non-convex QCQP problem given as

$$\max_x \quad x^T H x$$

subject to $x^T x \leq 1,$

$$\text{subject to } x^T x \leq 1,$$

where $x \in \mathbb{R}^n$ is the decision variable and $H \in \mathbb{R}^{n \times n}$ is an indefinite matrix. This problem is closely related to the well-known trust region problem [30]. For this kind of
problems, semidefinite relaxation finds a minimum value of an upper bound satisfying the constraints. To find this upper bound the non-convex QCQP is transformed into an approximate, but convex optimization problem with linear cost and linear matrix inequality constraints (an SDP). To derive this approximate formulation consider \( \bar{\gamma} \) is an upper bound on the cost of \( (3.22) \). Consider the identity where left hand side is equal to the right hand side:

\[
x^T H x - \bar{\gamma} = -\mu(1 - x^T x) + x^T H x - \bar{\gamma} + \mu(1 - x^T x),
\]

\[
= -\mu(1 - x^T x) - \begin{bmatrix} x^T & 1 \end{bmatrix} \begin{bmatrix} \mu I - H & 0 \\
0 & \bar{\gamma} - \mu \end{bmatrix} \begin{bmatrix} x \\
1 \end{bmatrix},
\]

(3.23)

where \( \mu \geq 0 \). In this expression, if the constraint is satisfied, then \( 1 - x^T x \geq 0 \). Now if \( \begin{bmatrix} \mu I - H & 0 \\
0 & \bar{\gamma} - \mu \end{bmatrix} \succeq 0 \), it ensures that the left hand side is either zero or negative, indicating that \( \bar{\gamma} \) is an upper bound on the cost of \( (3.22) \). To find this upper bound, we can solve the convex optimization problem, which is a relaxation of \( (3.22) \),

\[
\min_{\mu \geq 0, \bar{\gamma}} \bar{\gamma}
\]

subject to

\[
\begin{bmatrix} \mu I - H & 0 \\
0 & \bar{\gamma} - \mu \end{bmatrix} \succeq 0.
\]

(3.24)

This semidefinite relaxation is in fact the S-procedure for quadratic forms [17].

### 3.4 Summary

In this chapter we have mainly reviewed the standard static attitude estimation problem and the Orthogonal Procrustes Problem. One commonly used solution for each problem is also presented. Lastly, we have given a brief introduction to semidefinite relaxation, a semidefinite program and the S-procedure. In next three chapters we solve robust formulations of the presented estimation problems using semidefinite relaxation.
Chapter 4

Robust Attitude Estimation

In Chapter 3 we have presented a classical attitude estimation problem based on a weighted least-squares approach. However, with large data uncertainties, the classical formulation gives large attitude errors. In this chapter we formulate a robust attitude estimation problem, which is difficult to solve due to its unfavorable convexity properties. Our main concern in this chapter is to derive an approximate formulation, which is easier to solve and gives improved performance when uncertainties are large. The discussions and results presented in this chapter are mainly based on [3; 5].

4.1 Robust Problem Description

To formulate a robust attitude determination problem, we represent an uncertain measurement vector in the body frame with \( \bar{b}_i \in \mathbb{B}(b_i) \) and an uncertain reference vector with \( \bar{\bm{r}}_i \in \mathbb{R}(r_i), i = 1, \ldots, n \), where \( \mathbb{B}(b_i), \mathbb{R}(r_i) \subseteq \mathbb{R}^3 \) are bounded uncertainty sets.

To formulate the robust problem for attitude estimation, we use the weighted least-squares approach as used in (3.1) for the nominal problem. We define the robust problem as

\[
\min_C \max_{\bar{b}_i \in \mathbb{B}(b_i), \bar{\bm{r}}_i \in \mathbb{R}(r_i), i = 1, \ldots, n} \frac{1}{2} \sum_{i=1}^{n} w_i \| \bar{b}_i - C \bar{\bm{r}}_i \|^2_2
\]

subject to

\[
C^T C = I_3, \quad \det(C) = +1.
\]

(4.1)

To take advantage of using quaternions to simplify the optimization problem, as a first step, we reformulate (4.1) introducing the quaternion \( q \) using the same approach used to derive (3.3). We define

\[
\bar{\mathbb{B}} := \left[ \begin{array}{c} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_n \end{array} \right],
\]

52
4.2 Uncertainty Model

\[ \bar{R} := [\bar{r}_1 \bar{r}_2 \ldots \bar{r}_n]. \]

Using this stacked notation, we write the cost function in (4.1) in terms of \( q \) as

\[ J(q, \bar{B}, \bar{R}) := \left\{ \frac{1}{2} \text{tr}(W\bar{B}^T\bar{B} + W\bar{R}^T\bar{R}) - q^T K(\bar{B}, \bar{R})q \right\}. \quad (4.2) \]

The robust attitude determination problem is then written as

\[ \hat{q}^* := \arg \min_q \max_{\bar{B} \in \mathcal{B}(B), \bar{R} \in \mathcal{R}(R)} J(q, \bar{B}, \bar{R}) \quad (4.3) \]

subject to \( q^T q = 1 \),

where \( \mathcal{B}(B) := \mathcal{B}(b_1) \times \mathcal{B}(b_2) \times \cdots \times \mathcal{B}(b_n) \) and \( \mathcal{R}(R) := \mathcal{R}(r_1) \times \mathcal{R}(r_2) \times \cdots \times \mathcal{R}(r_n) \). Like the matrix \( K(B, R) \) in (3.3), the matrix \( K(\bar{B}, \bar{R}) \) is also symmetric and indefinite.

4.2 Uncertainty Model

In general, uncertainties in the input vectors are of diverse nature. These vectors are obtained from different sensors and mathematical models. Sensor errors are generally attributed to measurement noise, having a stochastic interpretation, and biases and misalignments, which are fixed values. Modeling inaccuracies have generally no clear interpretation. An uncertainty model, which can fully capture all these uncertainties is complex and can make the problem intractable. Keeping in view the tractability of solving (4.3), we consider an affine parameterization of the uncertainty sets \( \mathcal{B}(b) \) and \( \mathcal{R}(r) \) [10].

Let \( \beta \in \mathbb{R}^3 \) and \( \rho \in \mathbb{R}^3 \) be vectors of perturbation variables for uncertainty parameterization and \( \gamma_b \in \mathbb{R} \) and \( \gamma_r \in \mathbb{R} \) be bounds on the uncertainty for each input vector in the body and reference frame, respectively. We assume that each input vector may have different error bound. This type of uncertainty is called an interval uncertainty and the corresponding perturbation set represents a box := \( \{ \zeta : \|\zeta\|_{\infty} \leq 1 \} \), where \( \zeta \) is the normalized perturbation vector of appropriate dimension [10]. The interval uncertainty model is suitable for such mix uncertainties and can sufficiently capture most of the realistic errors. To further elaborate this point for vector quantities in \( \mathbb{R}^3 \) consider the uncertainties introduce an error in the true value. Let the maximum error introduced in each axis be bounded by \( \pm \gamma \), then we can say that the true value lies in an interval of size \( 2\gamma \) around the measurement. This interval in each axis makes a box in \( \mathbb{R}^3 \) with each side of length \( 2\gamma \). The size of this interval, i.e., the bound \( \gamma \) for each measurement or model vector, should be chosen carefully, as unnecessarily large values may result in large offset
4.3 An Approximation in the Robust Formulation

between the non-robust and the robust solution for nominal cases. The choice of bounds depends on the specific sensor or mathematical model used. Generally, sensor noise is known in a stochastic sense, e.g. standard deviation or variance, while modelling errors are given based on previous experimentation or analysis. However, biases and offsets need to be determined for each installed sensor separately. Overall, the chosen bound should sufficiently capture all these errors. Further, we normalize each perturbation vector in the body and reference frame with the corresponding uncertainty bound and denote it as $\delta_b := \beta/\gamma_b$ and $\delta_r := \rho/\gamma_r$. Using these normalized perturbation vectors, we describe the uncertainty sets in the body and reference frame as

$$B(b) = \left\{ b + \sum_{l=1}^{3} \delta_{bl} \tilde{b}_l \mid \| \delta_b \|_\infty \leq 1 \right\},$$

$$R(r) = \left\{ r + \sum_{l=1}^{3} \delta_{rl} \tilde{r}_l \mid \| \delta_r \|_\infty \leq 1 \right\},$$

where $\delta_b := \begin{bmatrix} \delta_{b1} & \delta_{b2} & \delta_{b3} \end{bmatrix}^T$, $\delta_r := \begin{bmatrix} \delta_{r1} & \delta_{r2} & \delta_{r3} \end{bmatrix}^T$, $\tilde{b}_l := \gamma_b e_l$ and $\tilde{r}_l := \gamma_r e_l$ are fixed vectors for a given problem settings with $e_l$ being the $l^{th}$ standard basis vector in $\mathbb{R}^3$.

4.3 An Approximation in the Robust Formulation

Using the described uncertainty model, we first derive an expression equivalent to (4.3), separating the terms, which depend on the uncertainty parameterization.

**Theorem 4.1.** The formulation given in (4.3) is equivalent to

$$\hat{q}^* := \arg \min \limits_q \left\{ -q^T \mathbf{K}(B,R)q + \max \limits_{\| \delta \|_\infty \leq 1} \left( p(q,B,R)\delta + \delta^T \mathbf{Q}(q) \delta \right) \right\}$$

subject to $q^T q = 1$, 

(4.5)
4.3 An Approximation in the Robust Formulation

where \( \delta := \begin{bmatrix} \delta_{b1}^T & \delta_{r1}^T & \delta_{b2}^T & \delta_{r2}^T & \ldots & \delta_{bn}^T & \delta_{rn}^T \end{bmatrix}^T \), \( \mathbf{p} := \mathbb{R}^4 \times \mathbb{R}^{3 \times n} \times \mathbb{R}^{3 \times n} \rightarrow \mathbb{R}^{6n} \) is

\[
\mathbf{p}(q, B, R) := \begin{bmatrix}
w_1 \gamma_{b1} (b_{11} - q^T K_{r1}^1 q) \\
w_1 \gamma_{b1} (b_{12} - q^T K_{r1}^2 q) \\
w_1 \gamma_{b1} (b_{13} - q^T K_{r1}^3 q) \\
w_1 \gamma_{r1} (r_{11} - q^T K_{b1}^1 q) \\
w_1 \gamma_{r1} (r_{12} - q^T K_{b1}^2 q) \\
w_1 \gamma_{r1} (r_{13} - q^T K_{b1}^3 q) \\
\vdots \\
w_n \gamma_{rn} (r_{n1} - q^T K_{bn}^1 q) \\
w_n \gamma_{rn} (r_{n2} - q^T K_{bn}^2 q) \\
w_n \gamma_{rn} (r_{n3} - q^T K_{bn}^3 q)
\end{bmatrix}, \tag{4.6}
\]

where \( b_{ij} \) and \( r_{ij} \) are \( j^{th} \) element of the \( i^{th} \) vector. The definition of the matrices \( K_{ri}^j \) and \( K_{bi}^j \) is given in the Appendix D. The matrix \( \mathbf{Q} \in \mathbb{R}^{6n \times 6n} \) is given as

\[
\mathbf{Q}(q) := \begin{bmatrix}
\frac{1}{2} w_1 \gamma_{b1}^2 I_3 & \frac{1}{2} w_1 \gamma_{b1} \gamma_{r1} C & \ldots & 0_{3 \times 3} & 0_{3 \times 3} \\
-\frac{1}{2} w_1 \gamma_{b1} \gamma_{r1} C^T & \frac{1}{2} w_1 \gamma_{r1}^2 I_3 & \ldots & 0_{3 \times 3} & 0_{3 \times 3} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0_{3 \times 3} & 0_{3 \times 3} & \ldots & \frac{1}{2} w_1 \gamma_{bn}^2 I_3 & -\frac{1}{2} w_1 \gamma_{bn} \gamma_{rn} C \\
0_{3 \times 3} & 0_{3 \times 3} & \ldots & -\frac{1}{2} w_1 \gamma_{bn} \gamma_{rn} C^T & \frac{1}{2} w_1 \gamma_{rn}^2 I_3
\end{bmatrix}, \tag{4.7}
\]

where the transformation matrix \( C \) is a function of \( q \).

\textbf{Proof.} Using the uncertainty model (4.4) in (4.2), the first term can be written as

\[
\frac{1}{2} \left( \text{tr}(WB^T B) + \text{tr}(WR^T R) \right) = \frac{1}{2} \left( \text{tr}(WB^T B) + \text{tr}(WR^T R) \right) + \\
\text{tr}(WB^T \Delta_b) + \text{tr}(WR^T \Delta_r) + \frac{1}{2} \left( \text{tr}(W \Delta_b^T \Delta_b + \text{tr}(W \Delta_r^T \Delta_r) \right), \tag{4.8}
\]

where \( \Delta_b := \begin{bmatrix} \gamma_{b1} \delta_{b1} & \gamma_{b2} \delta_{b2} & \ldots & \gamma_{bn} \delta_{bn} \end{bmatrix} \) and \( \Delta_r := \begin{bmatrix} \gamma_{r1} \delta_{r1} & \gamma_{r2} \delta_{r2} & \ldots & \gamma_{rn} \delta_{rn} \end{bmatrix} \). To simplify the second term, we write

\[
\mathbf{K}(\bar{B}, \bar{R}) = \mathbf{K}(B, R) + \mathbf{K}(B, \Delta_r) + \mathbf{K}(\Delta_b, R) + \mathbf{K}(\Delta_b, \Delta_r), \tag{4.9}
\]

55
4.3 An Approximation in the Robust Formulation

where $K$ is defined as in (3.4), with

$$
B(\bar{B}, \bar{R}) = B(B, R) + B(B, \Delta_r) + B(\Delta_b, R) + B(\Delta_b, \Delta_r),
$$

$$
= BW_{\Delta T} + BW_{\Delta r} + \Delta_b W_{\Delta T} + \Delta_b W_{\Delta r},
$$

$$
z(\bar{B}, \bar{R}) = z(B, R) + z(B, \Delta_r) + z(\Delta_b, R) + z(\Delta_b, \Delta_r),
$$

$$
= (B \times R)W_e + (B \times \Delta_r)W_e + (\Delta_b \times R)W_e + (\Delta_b \times \Delta_r)W_e,
$$

where $e$ is a vector of ones of appropriate dimension. We simplify and rearrange (4.8) and (4.9) and then write the expressions as a function of $\delta$. Separating the terms, which are linear or quadratic in $\delta$ and using the transformation matrix in terms of $q$ [77], i.e.

$$
C = \begin{bmatrix}
q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_2q_3 - q_1q_4) \\
2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\
2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2
\end{bmatrix},
$$

we can write the required expression.

It can be observed from (4.5) that the robust problem coincides the nominal problem if no uncertainty in the input vectors is considered. Moreover, the robust problem is always feasible, because the solution of the nominal problem (3.3) always exists for at least two linearly independent measurements [76]. However, finding an optimal solution of the formulated robust problem is difficult, because of the following two main reasons. Firstly, due to the matrix $Q(q)$ being positive semidefinite, the maximization term in (4.5) is non-concave in $\delta$, hence making it difficult to find a unique optimal maximum, and secondly, because of the matrix $K(B, R)$ being indefinite, the objective function is non-convex in $q$. To develop a tractable method of solving this difficult min-max problem, as a first step, we determine an upper bound on the maximum value of $p(q, B, R)^T \delta + \delta^T Q(q) \delta$ over $\delta$. The result is given in the following lemma, however the dependence of $p(q, B, R)$ and $Q(q)$ on $B, R$ and $q$ has been omitted for notational simplification.

**Lemma 4.1.** An upper bound on the ‘max’ term in (4.5) is

$$
0 \leq \max_{\|\delta\|_{\infty} \leq 1} (p^T \delta + \delta^T Q \delta) \leq \|p\|_1 + 6n\lambda_{\max}(Q).
$$

**Proof.** We start with the following inequality

$$
\max_{\|\delta\|_{\infty} \leq 1} (p^T \delta + \delta^T Q \delta) \leq \max_{\|\delta\|_{\infty} \leq 1} p^T \delta + \max_{\|\delta\|_{\infty} \leq 1} \delta^T Q \delta
$$

(4.11)
4.3 An Approximation in the Robust Formulation

Using the Hölder dual norm [18], the first term on the right hand side of (4.11) is given as

$$\max_{\|\delta\|_\infty \leq 1} p^T \delta = \|p\|_1.$$  \hspace{1cm} (4.12)

For the second term appearing in (4.11), since $Q$ is a symmetric matrix, we write the maximum eigenvalue of $Q$ as [18]

$$\lambda_{\text{max}}(Q) = \sup_{\|\delta\|_2 \leq 1} \delta^T Q \delta.$$  \hspace{1cm} (4.13)

Hence, we first replace the $\infty$-norm in the second term on the right hand side of (4.11) with the 2-norm using the inequality $\|\delta\|_2 \leq \sqrt{6n} \|\delta\|_\infty$ for $\delta \in \mathbb{R}^{6n}[33]$. We can write

$$\max_{\|\delta\|_\infty \leq 1} \delta^T Q \delta \leq \max_{\|\delta\|_2 \leq \sqrt{6n}} \delta^T Q \delta \leq 6n\lambda_{\text{max}}(Q),$$  \hspace{1cm} (4.14)

Using (4.12) and (4.14), we can write (4.10).

**Lemma 4.2.** The maximum eigenvalue of the block diagonal matrix $Q(q)$ does not depend on $q$.

**Proof.** The eigenvalues of the block diagonal matrix $Q(q) = \text{diag}(Q_1, Q_2, \ldots, Q_k)$ are simply those of $Q_i, i = 1, 2, \ldots, k$. To find the eigenvalues of each block $Q_i, i = 1, 2, \ldots, k$ we need to solve the characteristic equation $\det(Q_i - \lambda I_6) = 0$. For the case $i = 1$ we write the characteristic equation as

$$\det(Q_1 - \lambda I_6) = \det\begin{bmatrix} \lambda(\lambda - a) & 0 & 0 \\ 0 & \lambda(\lambda - a) \\ 0 & 0 & \lambda(\lambda - a) \end{bmatrix} = \lambda^3(\lambda - a)^3 = 0,$$

where $a := \frac{1}{2}w_1(\gamma_{61}^2 + \gamma_{71}^2)$. The six roots of the characteristic equation are $0, 0, 0, a, a, a$, showing that the eigenvalues are independent of $q$. Since all blocks have similar structure, the above analysis is applicable to all blocks of $Q(q)$. Hence, the maximum eigenvalue of $Q(q)$ is also independent of $q$. \hfill $\Box$
4.3.1 Comparison of the Analytical Upper Bound

This section discusses the tightness of the upper bound given in (4.10). Since the ‘max’ term in (4.5) is convex in $\delta$, it is hard to find the optimum. For such problems fairly tight bounds can be obtained using semidefinite relaxation [36; 54; 67]. We now find an upper bound using semidefinite relaxation and compare it with the analytical bound. We do this analysis for known $B, R$ and $q$. For these known values, we have $p = p(q, B, R)$ and $Q = Q(q)$.

Consider the ‘max’ term in (4.3). Suppose $\bar{\gamma}$ is an upper bound on this term satisfying the constraint $\|\delta\|_{\infty} \leq 1$, which can also be written as $-e \leq \delta \leq e$, where $e \in \mathbb{R}^{6n}$ is a vector of ones. Let $D \in \mathbb{R}^{6n \times 6n}$ be a diagonal matrix, then the following identity is always true

$$
\delta^T Q \delta + p^T \delta - \bar{\gamma} = -(e + \delta)^T D (e - \delta) - \begin{bmatrix} D - Q & -\frac{p}{2} \\ -\frac{p}{2} & \bar{\gamma} - e^T De \end{bmatrix} \begin{bmatrix} \delta \\ 1 \end{bmatrix}.
$$

(4.15)

In this expression, we know from the constraint $\|\delta\|_{\infty} \leq 1$, that both $(e + \delta) \geq 0$ and $(e - \delta) \geq 0$. Now the minimum value of $\bar{\gamma}$ represents an upper bound on the ‘max’ term if the diagonal matrix $D \succeq 0$, and matrix $\mathcal{F}(D, \bar{\gamma}) := \begin{bmatrix} D - Q & -\frac{p}{2} \\ -\frac{p}{2} & \bar{\gamma} - e^T De \end{bmatrix} \succeq 0$, i.e.

$$
\max_{\|\delta\|_{\infty} \leq 1} (\delta^T Q \delta + p^T \delta) \leq \min_{D, \bar{\gamma}} \{ \bar{\gamma} \mid D \succeq 0, \mathcal{F}(D, \bar{\gamma}) \succeq 0 \};
$$

(4.16)

which is a semidefinite relaxation (SDR) of the ‘max’ term of (4.3) for a given value of $q, B$ and $R$. A comparison of the analytical bound and the bound obtained using the SDR is given in Section 4.6.1, which shows that the relative error between the two bounds is small. Thus, the use of the analytical bound can give computational simplification, but at the cost of less accurate solution of (4.5), although the analysis shows that the gap is not much assuming that the (4.5) and its SDR are close.

At this stage, one might think of using (4.16), instead of the analytical upper bound (4.10) to simplify the original problem (4.3), as it is tighter than the analytical bound. Note that the formulation (4.16) is based on the assumption that $q$ is known. However, if $q$ is unknown, which is the case in actual problems, we may not get much computational benefit, because the terms of the matrix inequality $\mathcal{F}(D, \bar{\gamma}) \succeq 0$ are nonlinear in $q$. In this work, however, we use the analytical upper bound (4.10) to obtain a tractable, but suboptimal solution of (4.5).
4.4 Addition of a Regularization Term

Use of the analytical upper bound in the robust optimization problem (4.3) introduces an approximation, although it is motivated by the computational benefit and the fact that the unique solution of the inner loop maximization may not be possible. However, due to this approximation, the algorithm may not always give good results in terms of robustness. To improve performance, we introduce a new type of regularization in the objective function.

The optimization variable in the estimation problem, i.e., the quaternion \( q \), represents a coordinate transformation as a consequence of the Euler theorem of rotation \[ \alpha \] and is defined as

\[
q = \begin{bmatrix} q_0 \\ q_4 \end{bmatrix} := \begin{bmatrix} \hat{e} \sin(\alpha/2) \\ \cos(\alpha/2) \end{bmatrix}, \tag{4.17}
\]

where \( \hat{e} \in \mathbb{R}^3 \) is the axis of rotation and \( \alpha \) is the angle of rotation. We propose minimizing an additional term \(-\eta q_4^2\) along with the primary objective function. This regularization term is similar in concept to the Tikhonov regularization in linear least square problems \[ \ref{18} \]. In such problems, the regularization term is normally a function of the length of the solution vector. While minimizing the regularized least square cost, the added term enforces a trade-off between the primary objective and the length of the solution vector. However, in our case, as the norm of the solution vector is 1 \( (q^T q = 1) \), the regularization term is made a function of \( q_4 \), which corresponds to the angle of rotation for a given quaternion. Hence, the added term \( \eta q_4^2 \) as a second objective, enforces finding a quaternion \( q \), which minimizes a weighted combination of both objectives. In the added regularization term \( \eta > 0 \) is a tuning parameter. A large value of this tuning parameter makes the optimal solution of the regularized problem stiff to perturbations with a large offset in the nominal case. Simulations have shown that in the considered environment, \( \eta = 0.1 - 0.5 \) has shown good performance with a smaller offset and reasonably large robust performance margin.

4.5 Approximate Robust Formulation

Using Lemmas 4.1 and 4.2 and the regularization term, we now present the simplified but approximate formulation for the robust problem, which is easier to solve than the min-max problem (4.3). In Chapter 6, we use semidefinite relaxation to solve this problem. Due to this reason, the approximate robust problem is presented as a maximization
4.6 Performance Analysis of the Robust Formulation

problem, which is a form suitable for application of semidefinite relaxation.

**Corollary 4.1.** The problem (4.5) can be approximated by the maximization problem

\[
(q^*, u^*) = \arg \max_{q, u} q^T K_r(B, R)q - u^T e
\]

subject to

\[
q^T q = 1,
\]

\[-u \leq p(q, B, R) \leq u,
\]

(4.18)

where \(u_i \geq 0, i = 1, 2, \ldots, 6n\), \(K_r(B, R) := K(B, R) + \eta S\) and \(S = \text{diag}(0, 0, 0, 1)\).

**Proof.** In the formulation (4.5) given in Theorem 4.1, we replace the ‘max’ term with the upper bound given in Lemma 4.1. We neglect the term involving \(6n\lambda_{\text{max}}(Q(q))\), as according to Lemma 4.2 it does not depend on the decision variable \(q\), hence have no effect on the solution. We represent the regularization term \(\eta q^2\) as \(\eta q^T S q\). The non-differentiable term \(\|p(q, B, R)\|_1\) in the cost function is replace by the differentiable term \(u^T e\), where \(e\) is a vector of ones, with additional constraints \(-u_i \leq p_i \leq u_i, i = 1, 2, \ldots, 6n\), where \(p_i\) are elements of the vector \(p(q, B, R)\) [18, page 294]. Finally, expressing as a maximization problem, we write (4.18). We have neglected all constant terms in this expression having no effect on the solution of the optimization problem. These terms need to be added in the bound obtained from the SDR to find an exact value. \(\Box\)

4.6 Performance Analysis of the Robust Formulation

We consider the problem of attitude determination for a low cost CubeSat, as described in Chapter 2. For attitude determination, we use only two measurements, namely the Earth’s magnetic field and the sun vector. It has been discussed in Chapter 2 that there are two magnetometers, one inside the satellite, which is mainly used in the initial acquisition phase, while the other is installed on an extended boom, which is deployed once the satellite has achieved an equilibrium. A pair of sun sensors is used to measure the sun vector. Both of these measurements are in the body frame. In the reference frame, the Earth’s magnetic field is obtained from the first order IGRF model [72], while the sun vector is obtained using the simplified sun model discussed in Section 2.4.2. Both sensor measurements and reference vectors are not accurate due to measurement and modeling errors, as discussed in Section 2.4. Especially in the post-launch tumbling phase, the measurement noise further increases due to the use of an internal magnetometer installed
4.6 Performance Analysis of the Robust Formulation

Table 4.1: Vector data without errors. Vectors in the body frames are represented by $b_i$ and in the reference frames by $r_i$. The index $i$ is 1 for Earth magnetic field measurement and 2 is for the sun sensor data.

<table>
<thead>
<tr>
<th>Vectors in the body frame</th>
<th>Vectors in the reference frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1 = [-0.5422, -0.3158, 0.7787]^T$</td>
<td>$r_1 = [-0.5289, -0.3349, 0.7798]^T$</td>
</tr>
<tr>
<td>$b_2 = [-0.6732, 0.0212, 0.7392]^T$</td>
<td>$r_2 = [-0.6656, 0.00037, 0.7463]^T$</td>
</tr>
</tbody>
</table>

on-board the satellite, which suffers from coupling with the magnetic field generated by the magnetic actuators and the surrounding electronics. Similarly, the reference vectors are also not exact because they are obtained from mathematical models, which are normally based on low-order approximations for simplification and computational benefits. We are considering all such errors as $\infty$-norm bounded uncertainties, and for simulation purpose we set a bound of 30% of the norm of vectors in the body and the reference frame.

4.6.1 Tightness of the Analytical Upper Bound

In this section a comparison of the analytical upper bound (4.10) with the bound obtained from the semidefinite relaxation (4.16), as described in Section 4.3.1 is presented. For this comparison, we used two pairs of unit vectors, one in each pair is in the body frame and the other in the reference frame, as given in Table 4.1.

A uniformly distributed random error in the range $\pm \gamma_{bi}$ and $\pm \gamma_{ri}$ is introduced in the corresponding vectors for each simulation run. A comparison of both bounds and their relative errors for 100 simulations with random data is given in Figure 4.1. The plot shows that the relative error is less than 2% on average and less than 5% in the worst cases. This error corresponds to the error introduced by the analytical upper-bound in the true solution of (4.3). This analysis reveals that the price paid for transforming the min-max problem into a minimization problem, i.e., the error introduced by the use of the analytical bound, is small if the SDR bound is very close to the actual maximum value.
4.6 Performance Analysis of the Robust Formulation

Figure 4.1: Comparison of the analytical bound given in (4.10) and the bound obtained by solving the semidefinite program (4.16). The bound difference is given as a percentage of relative error between the SDR and the analytical bounds.

4.6.2 Robust Performance Comparison for Input Data at One Time Instant

The effect of uncertainty on attitude determination using the robust and non-robust solutions is presented for a given set of data for one time instant. A number of tests have been performed by adding uncertainty in the input vectors within the given bounds. To obtain the data, we added a uniformly distributed random error in the range of $\pm \gamma_{bi}$ and $\pm \gamma_{ri}$ to each measured and model vectors obtained from a simulation without error. The set of test vectors is given in Table 4.2.

For this data, the non-robust solution is obtained solving (3.3), while the robust solution is obtained solving (4.18). Figure 4.2 presents this comparison, where a histogram of the distribution of the cost (4.1) for different cases of uncertainty is shown. Here the x-axis represents the cost and the y-axis shows the number of tests. It can be observed
4.6 Performance Analysis of the Robust Formulation

\[ J = \frac{1}{2} \sum_{i=1}^{n} w_i \| \bar{b}_i - C\bar{r}_i \|_2^2 \]

Figure 4.2: Histogram showing the distribution of the cost for the robust and non-robust solutions. The plot shows data for 2000 runs. In each run, uniformly distributed random error is added in the test vectors within the given uncertainty bounds.

Table 4.2: Vector data with added error. Vectors in the body frames are represented by \( b_i \) and in the reference frames by \( r_i \).

<table>
<thead>
<tr>
<th>Vectors in the body frame</th>
<th>Vectors in the reference frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_1 = \begin{bmatrix} -0.7757 &amp; -0.4611 &amp; 0.431 \end{bmatrix}^T )</td>
<td>( r_1 = \begin{bmatrix} -0.5411 &amp; -0.3263 &amp; 0.775 \end{bmatrix}^T )</td>
</tr>
<tr>
<td>( b_2 = \begin{bmatrix} -0.9272 &amp; 0.0114 &amp; 0.3743 \end{bmatrix}^T )</td>
<td>( r_2 = \begin{bmatrix} -0.6725 &amp; 0.000133 &amp; 0.74 \end{bmatrix}^T )</td>
</tr>
</tbody>
</table>
4.6 Performance Analysis of the Robust Formulation

\[ J = \frac{1}{2} \sum_{i=1}^{n} w_i |\bar{b}_i - C_i\|^2 \]

Figure 4.3: The effect of the change of the tuning parameter \( \eta \) on robust performance.

that spread of the cost using the nominal solution is much more than that using the robust solution, as the uncertainty is varied within the defined box, showing the usefulness of the robust approach.

4.6.3 The Effect of the Regularization Term

This section analyzes the effect of the regularization term added in the cost function of (4.18). Figure 4.3 shows how the performance of the robust solution changes against uncertainties with different values of the tuning parameter \( \eta \). In this analysis, we added error in the input vectors in a systematic way by parameterizing it with a single uncertainty parameter varying in the range -1 to 1. On the y-axis, we plot the cost \( J \). It can be observed that the solution with regularization term improves the robustness against uncertainties.
4.6 Performance Analysis of the Robust Formulation

In this section, we present performance comparison of the robust and the non-robust approaches in the presence of uncertainties, using in-orbit data obtained from a nonlinear closed-loop simulation for the satellite. The ideal data has been corrupted by adding uniformly distributed random errors in the range of $\pm \gamma_{bi}$ and $\pm \gamma_{ri}$ in the corresponding vectors. We present attitude determination results for 25 minutes of flight data obtained with a sample time of 1 second. The simulation is initialized with roll, pitch and yaw body rates of 0.5, 0.5 and 0.1 deg/s and roll, pitch and yaw angles of $10^\circ$, $0^\circ$ and $0^\circ$, respectively. We have solved the robust problem formulated in (4.18) using the nonlinear optimization solver `fmincon` of MATLAB. The performance benefit of the robust approach is illustrated by in Figure 4.4, showing an improvement over the non-robust approach in the presence of uncertainties. It can be observed that due to the

---

Figure 4.4: A comparison of the attitude angles obtained using the robust and non-robust algorithms. The dotted line shows the original data without errors while the other two cases include errors within the chosen uncertainty bound.

---

4.6.4 Robust Performance Comparison for In-Orbit Simulation Data
uncertainties in the sensors and model vector information, the non-robust approach can
give large errors in the estimated attitude angles, while the robust approach gives much
better performance, limiting the maximum attitude error to a smaller band.

4.7 Summary

We have presented a robust min-max attitude estimation problem with nonlinear con-
straints. The robust problem where the ‘max’ term is convex has been approximated
by a minimization problem using an upper bound on the ‘max’ term. The approximate
formulation is non-convex with quadratic objective function and constraints (a QCQP).
A regularization term has also been introduced to improve the robust performance. The
performance of the robust solution has been evaluated in comparison with the nominal
solution of Chapter 3. It has been observed that the approximate robust formulation
gives improved performance than the nominal formulation in the presence of data uncer-
tainties. However, solving this nonlinear non-convex QCQP may still be computationally
expensive. Finding an efficient solution of this problem is the topic addressed in the next
chapter.
Chapter 5

Solution of the Robust Estimation Problem using Semidefinite Relaxation

The robust estimation problem presented in Chapter 4 has a non-convex quadratic cost with quadratic equality and inequality constraints. Finding a unique optimal solution of this non-convex problem is computationally intractable. In this chapter, we use semidefinite relaxation to find an upper-bound on this non-convex problem using a semidefinite program, which can efficiently be solved in practice. This chapter is mainly based on [4; 5]

5.1 Semidefinite Relaxation for Non-convex QCQP Problems

As discussed in Chapter 3, semi-definite relaxation has proved to be a very useful tool to approximately solve difficult optimization problems, such as non-convex QCQP problems. Since the formulated robust estimation problem is also of this kind, we use semidefinite relaxation, as discussed in Chapter 3, Section 3.3, to solve this problem. Using this relaxation, an upper bound is found on the non-convex cost by solving a semidefinite programming problem (SDP), for which very efficient polynomial time interior-point algorithms exist [8; 53]. When applying relaxations, it is very important to analyze the gap between the actual and the relaxed problem. This analysis helps in judging the quality of the obtained solution in comparison to the optimal solution.
5.2 Semidefinite Relaxation for the Robust Estimation Problem

In this section we apply semidefinite relaxation on the robust estimation problem (4.18) [4]. Suppose $\bar{\gamma}$ is an upper bound for the objective function of (4.18). Using a similar approach, as used in deriving (4.15), we obtain the expression. (we drop the dependence on $B, R$ and $q$ for notational simplification, except where necessary)

$$q^T K_r q - u^T e - \bar{\gamma} = -\mu_1 (1 - q^T q) - \mu_2 (u_1 - p_1) - \mu_3 (u_1 + p_1) - \mu_4 (u_2 - p_2) - \mu_5 (u_2 + p_2) - \ldots - \mu_{12n} (u_{6n} - p_{6n}) - \mu_{12n+1} (u_{6n} + p_{6n}) - x^T \mathcal{L}(\mu, B, R)x,$$  \hspace{1cm} (5.1)

where

$$x := \begin{bmatrix} q^T & u^T & 1 \end{bmatrix}^T,$$

$$\mu := \begin{bmatrix} \mu_1 & \mu_2 & \ldots & \mu_{12n+1} \end{bmatrix}^T,$$

$$p(q, B, R) := \begin{bmatrix} p_1 & p_2 & \ldots & p_{6n} \end{bmatrix}^T,$$

$$\mathcal{L}(\mu, B, R) := \begin{bmatrix} \mathcal{L}_{1,1}(\mu, B, R) & 0_{4 \times 1} & \ldots & 0_{4 \times 1} & 0_{4 \times 1} \\ 0_{1 \times 4} & 0 & \ldots & 0 & \frac{1}{2} (\mu_2 - \mu_3) \\ 0_{1 \times 4} & 0 & \ldots & 0 & \frac{1}{2} (\mu_4 - \mu_5) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{1 \times 4} & 0 & \ldots & 0 & \frac{1}{2} (\mu_{12n} - \mu_{12n+1}) \\ 0_{1 \times 4} & \frac{1}{2} (\mu_2 - \mu_3) & \ldots & \frac{1}{2} (\mu_{12n} - \mu_{12n+1}) & \ell_{j,j}(\mu, B, R) \end{bmatrix}.$$  

$$\mathcal{L}_{1,1}(\mu, B, R) := \mu_1 I_4 - (\mu_2 - \mu_3) w_1 \gamma_{b_1} K_{r_1}^1 - (\mu_4 - \mu_5) w_1 \gamma_{b_1} K_{r_1}^2 - (\mu_6 - \mu_7) w_1 \gamma_{b_1} K_{r_1}^3 - (\mu_8 - \mu_9) w_1 \gamma_{r_1} K_{b_1}^1 - \cdots - (\mu_{12n} - \mu_{12n+1}) w_n \gamma_{r_1} K_{b_n}^3 - K_r(B, R),$$

$$\ell_{j,j}(\mu, B, R) := \bar{\gamma} = -\mu_1 + \sum_{l=1}^{6n} (\mu_{2l} - \mu_{2l+1}) c_l(B, R).$$

Here $j$ is the size of $x$ and $c(B, R) := \begin{bmatrix} w_1 \gamma_{b_1} b_1^T & w_1 \gamma_{r_1} r_1^T & \ldots & w_n \gamma_{b_n} b_n^T & w_n \gamma_{r_n} r_n^T \end{bmatrix}^T.$

If the right hand side of (5.1) is either zero or negative, we can say that $\bar{\gamma}$ is an upper bound on the cost of (4.18). Using this relaxation, we write an optimization problem to find the minimum value of this upper bound ensuring the right hand side is either zero.
5.2 Semidefinite Relaxation for the Robust Estimation Problem

or negative, given by

\[
(\bar{\gamma}, \mu^*) := \arg\min_{\hat{\gamma}, \mu} \{ \hat{\gamma} \mid L(\mu, B, R) \succeq 0, \mu_i \geq 0, i = 2, 3, \ldots, 12n + 1 \},
\]  

(5.2)

where \(L(\mu, B, R)\) satisfies the conditions of the S-procedure, given in Section 3.3.2. Note that a few diagonal entries of the matrix \(L(\mu, B, R)\) are zero. For this matrix to be positive semidefinite, we can force the corresponding non-diagonal terms to be zero. This does not only result in a reduced set of optimization variables, but also avoids numerical issue arising due to the zero diagonal entries.

**Theorem 5.1.** Using a reduced set of optimization variables \(\mu_r := [\mu_1 \ \mu_2 \ \mu_4 \ \mu_6 \ \ldots \ \mu_{12n}]^T\) an equivalent formulation of (5.2) is

\[
\mu_r^* = \arg\min_{\mu_r} \mu_1 - \sum_{l=1}^{6n} (2\mu_{2l} - 1) c_l(B, R) \\
\text{subject to} \quad 0 \leq \mu_i \leq 1, \ i = 2, 4, \ldots, 12n, \\
L_{1,1}(\mu_r, B, R) \succeq 0,
\]  

(5.3)

where \(L_{1,1}(\mu_r, B, R)\) is given by

\[
L_{1,1}(\mu_r, B, R) := \mu_1 I_4 - 2\mu_2 w_1 \gamma b_1 K_{r1}^1 - 2\mu_4 w_1 \gamma b_1 K_{r1}^2 \\
- 2\mu_6 w_1 \gamma b_1 K_{r1}^3 - 2\mu_8 w_1 \gamma b_1 K_{b1}^1 - \ldots \\
- 2\mu_{12n} w_1 \gamma b_n K_{r1}^3 + w_1 \gamma b_1 K_{r1}^1 + w_1 \gamma b_1 K_{r1}^2 + w_1 \gamma b_1 K_{r1}^3 - K_r(B, R).
\]  

(5.4)

**Proof.** Note that in (5.2) the symmetric matrix \(L(\mu, B, R)\) has zero diagonal elements. For \(L(\mu, B, R)\) to be positive semidefinite, as required in (5.2), all row/column elements corresponding to the zero diagonal entries must also be zero [33, Thm 4.2.6], i.e. \(1 - \mu_2 - \mu_3 = 0, 1 - \mu_4 - \mu_5 = 0, 1 - \mu_6 - \mu_7 = 0\) and so on. Using this property, we can force these elements to be zero by eliminating \(\mu_3, \mu_5, \ldots, \mu_{12n+1}\) from (5.2) with additional constraints \(1 - \mu_2 \geq 0, 1 - \mu_4 \geq 0, \ldots, 1 - \mu_{12n} \geq 0\). Moreover, the minimum possible value of the scalar \(\hat{\gamma}\) that satisfies the constraint \(L(\mu_r, B, R) \succeq 0\) is one which results in \(l_{j,j}(\mu_r, B, R) = 0\), giving

\[
\hat{\gamma} = \mu_1 - \sum_{l=1}^{6n} (2\mu_{2l} - 1) c_l(B, R).
\]  

(5.5)
With these modifications, instead of $\mathcal{L}(\mu, B, R) \succeq 0$, we only need $\mathcal{L}_{1,1}(\mu_r, B, R) \succeq 0$, hence can write (5.3) using a reduced number of optimization variables, which is equivalent to solving (5.2) for a minimum upper bound on (4.18).

\[\square\]

### 5.3 Finding the Robust Quaternion ($q^*$)

Although the solution of the semidefinite program (5.3) gives a minimum upper bound on the robust estimation problem (4.18), our main interest is to find $q^*$ that could maximize the cost (4.18). Now the question arises, can we find $q^*$ using the solution $\mu_r^*$ of (5.3)? Suppose $\mu_r^*$ results in a zero value of the right hand side of (5.1), then $\hat{\gamma}^*$, i.e. the minimum value of the cost (5.3), is equal to the maximum cost of (4.18), and the corresponding $q$ is the required $q^*$.

In this regard, as a first step, we establish whether there exists a $q$ that can make

\[q^T \mathcal{L}_{1,1}^* q = 0, \quad \text{where} \quad \mathcal{L}_{1,1}^* := \mathcal{L}_{1,1}(\mu_r^*, B, R).\]

If such a $q$ exists, it further ensures $x^T \mathcal{L}^* x = 0$, where $\mathcal{L}^* := \mathcal{L}(\mu^*, B, R)$ and $\mu^*$ can be obtained from $\mu_r^*$.

**Lemma 5.1.** Let $\mu_r^*$ be a minimizer for the SDR problem (5.3), then $\lambda_{\min}(\mathcal{L}_{1,1}^*) = 0$.

**Proof.** Using $\mu_r$, the objective function (5.3) can be written as $J := \mu_1 - d$, where $d$ is the sum of all remaining terms. Whatever the sign of $d$ is, the cost function $J$ is minimum when $\mu_1$ is minimum. However, at the same time, we need $\mathcal{L}_{1,1}(\mu_r, B, R) \succeq 0$. We can also write $\mathcal{L}_{1,1}(\mu_r, B, R) = \mu_1 I_4 - K_{\mu}$, where $K_{\mu}$ is the sum of all other terms in the expression. This is a symmetric matrix with real eigenvalues $\lambda_1, \ldots, \lambda_4$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. Then, $\mu_1 I_4 - K_{\mu}$ has eigenvalues $\mu_1 - \lambda_1, \mu_1 - \lambda_2, \mu_1 - \lambda_3, \mu_1 - \lambda_4$. Now, $\mu_1 = 1$ is the smallest possible value that can make $\mathcal{L}_{1,1}(\mu_r, B, R) \succeq 0$. This optimal value of $\mu_1$, i.e. $\mu_1^*$ also ensures that $\lambda_{\min}(\mathcal{L}_{1,1}^*)$ is equal to zero. \[\square\]

**Remark 5.1.** As stated in Lemma 5.1, the matrix $\mathcal{L}_{1,1}^*$ has at least one eigenvalue equal to zero. Suppose there is only one eigenvalue equal to zero and $\tilde{q}$ is an eigenvector of $\mathcal{L}_{1,1}^*$ corresponding to the zero eigenvalue, then this $\tilde{q}$ results in both $\mathcal{L}_{1,1}^* \tilde{q} = 0_{4 \times 1}$ and $\tilde{q}^T \mathcal{L}_{1,1}^* \tilde{q} = 0$, because $\tilde{q}$ belongs to the null space of $\mathcal{L}_{1,1}^*$. From this we can deduce that $\tilde{x}^T \mathcal{L}^* \tilde{x} = 0$, where $\tilde{x} := \begin{bmatrix} \tilde{q}^T & \tilde{u}^T & 1 \end{bmatrix}^T$, although we have no knowledge of $\tilde{u}$ at this stage. This is possible because all elements of the matrix $\mathcal{L}^*$ are zero, except the sub-matrix $\mathcal{L}_{1,1}^*$.

The vector $\tilde{q}$ can be a candidate for the robust quaternion. If there is no gap between the cost of (4.18) and (5.3), then $\tilde{q}$ is the required optimal solution of (4.18), i.e. the
5.3 Finding the Robust Quaternion ($q^*$)

robust quaternion.

5.3.1 Relaxation Gap

This section discusses the gap between the approximate robust attitude determination problem (4.18) and its semidefinite relaxation (5.3). Firstly, we present a result which is used to quantify this gap. The result and its proof follows similar lines as discussed in [54]. In the onward discussion, we have dropped the dependence on $B$ and $R$ for simplicity.

Lemma 5.2. Let $\mu^*_r$ be a minimizer for the SDR problem in (5.3), such that $h = \dim N(L_{1,1}(\mu^*_r)) \geq 1$, and

$$L_{1,1}(\mu^*_r) = \begin{bmatrix} V & V_+ \end{bmatrix} \begin{bmatrix} 0_h & 0 \\ 0 & \Lambda_+ \end{bmatrix} \begin{bmatrix} V^T \\ V_+^T \end{bmatrix}$$  \hspace{1cm} (5.6)

be a spectral decomposition of $L_{1,1}(\mu^*_r)$ for some orthogonal $[V \ V_+]$ and $\Lambda_+ \succ 0$. Let the optimal cost of (5.3) be $J(\mu^*_r)$. Let $z = [z_1, z_2, z_4, z_6, \ldots, z_{12n}]^T \geq 0$, where $z_i \in \mathbb{R}, i = 1, 2, 4, 6, \ldots, 12n$, then there does not exist a $z$ such that

1. $J(\mu^*_r - z) = J(\mu^*_r)$, i.e. $J(z) = 0$

2. $\mu^*_{2i} \geq z_{2i}, i = 1, \ldots, 6n,$

3. $1 - \mu^*_{2i} + z_{2i} \geq 0, i = 1, \ldots, 6n,$

4. $V^T L_0(z)V < 0$, where $L_0(z) = L_{1,1}(\mu^*_r) - L_{1,1}(\mu^*_r - z)$.

Proof. Suppose such a $z$ exists. We choose a small value of $\epsilon > 0$ such that $\mu^*_r - \epsilon z$ is another solution to (5.3), satisfying all above points. We evaluate $L_{1,1}(\mu^*_r - \epsilon z) = L_{1,1}(\mu^*_r) - \epsilon L_0(z)$, and write

$$\begin{bmatrix} V^T \\ V_+^T \end{bmatrix} L_{1,1}(\mu^*_r - \epsilon z) \begin{bmatrix} V & V_+ \end{bmatrix} = \begin{bmatrix} V^T \\ V_+^T \end{bmatrix} \left( L_{1,1}(\mu^*_r) - \epsilon L_0(z) \right) \begin{bmatrix} V & V_+ \end{bmatrix}.$$

Using (5.6), we can write

$$\begin{bmatrix} V^T \\ V_+^T \end{bmatrix} L_{1,1}(\mu^*_r - \epsilon z) \begin{bmatrix} V & V_+ \end{bmatrix} = \begin{bmatrix} -\epsilon V^T L_0(z)V & -\epsilon V^T L_0(z)V_+ \\ -\epsilon V^T L_0(z)V & -\epsilon V^T L_0(z)V_+ \end{bmatrix} \Lambda_+ - \epsilon V^T L_0(z)V_+.$$

Now, from point 4, we know that $-V^T L_0(z)V > 0$ and

$$-V^T L_0(z)V - \epsilon V^T L_0(z)V \left( \Lambda_+ - \epsilon V^T L_0(z)V_+ \right)^{-1} V_+^T L_0(z)V > 0,$$
5.3 Finding the Robust Quaternion \((q^*)\)

because \(\Lambda_+ > 0\) and we can choose \(\epsilon > 0\) such that \(\Lambda_+ - \epsilon V^T_+ \mathcal{L}_0(z) V_+ > 0\) and the above is true. Using Schur complement condition for positive definiteness, the above implies that \(\mathcal{L}_{1,1}(\mu^*_r - \epsilon z) > 0\). However, this contradicts Lemma 5.1.

Next, we present our main result regarding the gap between the SDR and (4.18) and also relate the vector \(\tilde{q}\) determined in Remark 5.1 and \(q^*\), i.e. the solution of (4.18).

**Theorem 5.2.** Assume \(h = 1\) the vector \(\tilde{q}\), which makes \(\tilde{q}^T \mathcal{L}^*_{1,1} \tilde{q} = 0\) gives no relaxation gap between the approximate problem (4.18) and its semidefinite relaxation (5.3), hence \(q^* = \tilde{q}\).

**Proof.** For no gap, we need to prove each term on the right hand side of (5.1) to be zero. We use \(\tilde{q}\) obtained from Remark 5.1, satisfying \(\tilde{q}^T \tilde{q} = 1\) and \(\mathcal{L}_{1,1}(\mu^*_r) \tilde{q} = 0\).

1. Satisfying \(\tilde{q}^T \tilde{q} = 1\) implies \(\mu_1 (1 - \tilde{q}^T \tilde{q}) = 0\).
2. Satisfying \(\mathcal{L}_{1,1}(\mu^*_r) \tilde{q} = 0\) implies \(\tilde{x}^T \mathcal{L}(\mu^*) \tilde{x} = 0\).
3. To prove that the remaining terms are zero, we first show that

   (a) if \(\mu_{2i} \neq 0\), then \(p_i \geq 0\), \(i = 1, \ldots, 6n\).
   (b) if \(\mu_{2i+1} \neq 0\), then \(p_i \leq 0\), \(i = 1, \ldots, 6n\).

To prove (a) we write the optimal cost function is (5.3) in terms of \(p_i\). For this, pre and post multiplying both sides of (5.4) by \(\tilde{q}^T\) and \(\tilde{q}\), and using the fact that \(\tilde{q}^T \mathcal{L}_{1,1}(\mu^*_r) \tilde{q} = 0\) yields

\[
0 = \mu_1 - (2\mu_2 - 1)w_1 \gamma_{b1} \tilde{q}^T K^1_{11} \tilde{q} - (2\mu_4 - 1)w_1 \gamma_{b1} \tilde{q}^T K^2_{11} \tilde{q} - (2\mu_6 - 1)w_1 \gamma_{b1} \tilde{q}^T K^3_{11} \tilde{q} - (2\mu_8 - 1)w_1 \gamma_{r1} \tilde{q}^T K^1_{b1} \tilde{q} - \ldots
\]

Finally, subtracting (5.7) from (5.5) yields

\[
\gamma = \tilde{q}^T \mathcal{K}_r(B,R) \tilde{q} - (2\mu^*_2 - 1)p_1 - (2\mu^*_4 - 1)p_2 - \cdots - (2\mu^*_1 - 1)p_{6n}, \quad (5.8)
\]

Now, consider the case \(i = 1\). If \(p_1 < 0\), we need to prove \(\mu^*_2 = 0\). Let we contradict by saying that \(\mu^*_2 > 0\). Then there exist \(z_2 > 0\), such that \(\mu^*_2 \geq z_2\). We assume \(z_4, \ldots, z_{12n}\) to be zero. Using Lemma 5.2 (point 1), we have \(z_1 = 2z_{2c1}\) i.e. the
new $z$ satisfies points 1-3 of Lemma 5.2. Then, using $\tilde{q}^T L_{1,1}(\mu^*_r) \tilde{q} = 0$ we have

$$
\tilde{q}^T L_{1,1}(\mu^*_r - z) \tilde{q} = \tilde{q}^T (-z_1 I_4 + 2z_2 w_1 \gamma b_1 \tilde{q}^T K^1_{r_1}) \tilde{q},
$$
$$
\tilde{q}^T L_0(z) \tilde{q} = 2z_2 p_1.
$$

Here, as $p_1 < 0$ and $z_2 > 0$, we have $\tilde{q}^T L_0 \tilde{q} < 0$, which contradicts Lemma 5.2. Hence we conclude that such a $z_2$ is not feasible and $\mu^*_2 = 0$. Using a similar approach, we obtain the result for all values of $i$, proving part (a). Similarly, for part (b), we need to show that if $p_i > 0$, then $\mu_{2i+1} = 0$ or, in the reduced variable settings, $1 - \mu_{2i} = 0$, using the condition $\mu_{2i} + \mu_{2i+1} = 1$ and the constraint $1 - \mu_{2i} \geq 0$. Now, following a similar approach as in part (a), we can write for $i = 1$

$$
\tilde{q}^T L_0(z) \tilde{q} = -2(1 - z_2)p_1.
$$

Since $p_1 > 0$ and $1 - z_2 > 0$, hence $\tilde{q}^T L_0 \tilde{q} < 0$, which is not possible from Lemma 5.2, proving part (b).

Finally, we prove that there exists $u \geq 0$, such that the remaining terms in (5.1) are zero. Since $\mu_{2i} + \mu_{2i+1} = 1$ and $\mu_{2i} \geq 0, \mu_{2i+1} \geq 0$ and $|p_i| = u_i, i = 1, 2, \ldots, 6n$, there are three possibilities.

i) $\mu_{2i} = 1, \mu_{2i+1} = 0$: From (a), we know that in this case $p_i \geq 0$ and we define $u_i = p_i$.

ii) $\mu_{2i} = 0, \mu_{2i+1} = 1$: From (b), we know that in this case $p_i \leq 0$ and we define $u_i = -p_i$.

iii) $\mu_{2i} \neq 0, \mu_{2i+1} \neq 0$: From (a) and (b), we know that in this case $p_i = 0$ and we define $u_i = 0$.

It has been observed in the numerical simulations that the chance that the case $h > 1$ occurs is rare. However, we cannot rule out the possibility of its occurrence. If such a case occurs, then there are more than one eigenvectors corresponding to the zero eigenvalues. For any of these solutions, zero relaxation gap cannot be guaranteed. The possibility of a solution with zero relaxation gap for $h > 1$ needs to be further explored and could be a possible future research direction.
5.4 Numerical Results

To analyze the performance of the semidefinite relaxation, we mainly use the in-orbit simulation data, previously used to evaluate the performance of the robust formulation in Chapter 4.

5.4.1 Comparison of the Quaternion Obtained from the Approximate Problem and the SDR

Here, we present a quantitative comparison of the optimal quaternion obtained from (4.18) using MATLAB’s `fmincon` (with interior-point algorithm, tolerance of $10^{-12}$ and an initial guess of the eigenvector of the $K$ matrix corresponding to the largest eigenvalue i.e. the quaternion for the non-robust solution) and the solution of (5.3) using MATLAB’s robust control toolbox command `mincx` (with the same tolerance). We used used the perturbed vector data given in Table 4.2. A comparison of the two quaternions is given in Table 5.1. Note that $q^*$ is obtained using Remark 5.1. The error between the two quaternion obtained by solving a nonlinear optimization problem and its relaxed semidefinite program is negligible, which may be attributed towards numerical errors.
5.4 Numerical Results

5.4.2 Comparison of the Solutions for In-Orbit Simulation Data

In this section we present comparison of the robust solutions obtained through the nonlinear optimization algorithm with the solution obtained from the semidefinite relaxation for in-orbit data, as discussed in Chapter 4. We have solved the robust problem formulated in (4.18) using the nonlinear optimization solver fmincon of MATLAB, while the problem formulated using the semidefinite relaxation in (5.3) was solved using the Robust Control toolbox command mincx. Results are shown in Figure 5.4. It can be observed that both solutions gives similar performance and the attitude error is less than ±0.1°, which can be attributed to the numerical errors while transforming quaternion to Euler angles.

5.4.3 Illustration of Sections 5.2 and 5.3: Theoretical Results

In this section we illustrate the theoretical results presented in Section 5.2 and 5.3. Firstly, the eigenvalues of $L_{1,1}^*$ are plotted in Figure 5.2 for the in-orbit flight data, showing that the smallest eigenvalue is zero for all cases, validating Remark 5.1. Figures 5.3 and 5.4 support Theorem 5.2. Figure 5.3 shows the relaxation gap between the robust problem
and its semidefinite relaxation i.e. $q^T K_r q - u^T e - \bar{\gamma}$, where the first part $q^T K_r q - u^T e$ is calculated using the results obtained from `fmincon` and $\bar{\gamma}$ is obtained from `mincx`. It can be observed that the gap is zero for all time instances of the simulation. Figure 5.4 shows the difference between the quaternion obtained from the two solutions. For the SDR case, the quaternion are obtained using Remark 5.1. It can be observed that both the relaxation gap and the error in corresponding components of quaternion is almost zero for all time instances.

### 5.5 Summary

Semidefinite relaxation has been used to transform the non-convex QCQP formulated in Chapter 4 for robust attitude estimation into a semidefinite program with a linear cost and linear matrix inequality constraints. The relaxed problem can efficiently be solved using any SDP solver. It has also been shown how to extract the robust attitude information from the solution of the relaxed problem. Further, we proved that the gap between the formulation (4.18) and its relaxation (5.3) is zero for the case $h =$
Table 5.1: Comparison of the elements of quaternion $q_i, i = 1, 2, \ldots, 4$ obtained from (4.18) and (5.3) for the vector set given in Table 4.2

| $q_i^*$ | $\tilde{q}_i$ | $|q_i^* - \tilde{q}_i|$ |
|---------|--------------|------------------|
| 0.0761303170 | 0.0761303011 | 1.59928936×10⁻⁸ |
| 0.0444603409 | 0.0444603345 | 6.39020509×10⁻⁹ |
| -0.0305429683 | -0.0305429452 | -2.31056824×10⁻⁸ |
| 0.9956377755 | 0.9956377777 | -2.21610053×10⁻⁹ |

1, showing that the extracted quaternion is the solution of the nonlinear optimization problem (4.18). Numerical results have been presented to validate the theoretical results and the discussions presented in different sections of the chapter.

Figure 5.4: Difference between the quaternions obtained by solving (4.18) and its semidefinite relaxation.
Chapter 6

Orthogonal Procrustes Problem with Data Uncertainties

This chapter extends the discussions presented in Chapters 4 and 5 to a more general problem, known as the Orthogonal Procrustes Problem. Different formulations of this problem address reflections, rotations and translations and have diverse applications. The attitude estimation problem is also a sub-class of the Orthogonal Procrustes Problem. It is discussed how semidefinite relaxations can be used to solve different formulations of the Orthogonal Procrustes Problem with and without data uncertainties. The discussions presented in this chapter are mainly based on [2].

6.1 Semidefinite Relaxation for the Orthogonal Procrustes Problem

The Orthogonal Procrustes Problem has been introduced in Chapter 3, however, its mathematical definition is being repeated here for convenience, namely

$$\min_X \|AX - B\|_F^2,$$
subject to
$$XX^T = X^TX = I,$$  \hspace{1cm} (6.1)

where $A, B \in \mathbb{R}^{m \times n}$ are given data matrices with $m \geq n$, and $X \in \mathbb{R}^{n \times n}$ is the unknown orthogonal matrix which belongs to an orthogonal group of order $n$. The determinant of $X$ is $+1$ or $-1$. In the OPP for rotations $X \in \mathbb{R}^{3 \times 3}$ with an additional nonlinear constraint, i.e., $\det(X) = +1$. In this case, the solution now belongs to $SO(3)$.

In this section we use semidefinite relaxations to solve this problem. Although very
6.1 Semidefinite Relaxation for the Orthogonal Procrustes Problem

Efficient and numerically robust solutions of this problem already exist, such as the one presented in Section 3.2.1, our main objective is to study how SDR can be applied to solve this problem and how this approach can be extended to solve the robust problem. We present a relaxed formulation for the standard OPP (3.8), as well as the OPP for rotations. Both constraints $XX^T = I$ and $\det(X) = +1$ are non-convex and cannot be directly handled in the SDR framework. In the relaxed formulations, the non-convex constraints are either relaxed or replaced with convex approximations.

6.1.1 Relaxation of the Standard OPP

To derive a semidefinite relaxation of the standard OPP, we use (3.11), repeated here as

$$\|AX - B\|^2_F = \text{tr} (AXX^T A^T + BB^T - AXB^T - BX^T A^T),$$

(6.2)

and simplify the expression using the constraint $XX^T = I$. By introducing a linear objective, i.e., the trace of an unknown symmetric matrix $M \in \mathbb{R}^{m \times m}$, we write the optimization problem, which is equivalent to (6.1),

$$\min_{X,M} \quad \text{tr}(M),$$

subject to

$$M - AA^T - BB^T + AXX^T A^T \preceq 0,$$

$$XX^T = I.$$  

(6.3)

Since the orthogonality constraint $XX^T = I$ is not convex, we relax it to a convex quadratic inequality $XX^T \preceq I$. Using this relaxation, we write an approximate problem, which has a linear cost and linear matrix inequality constraints, namely

$$\min_{X,M} \quad \text{tr}(M),$$

subject to

$$M \succeq 0,$$

$$X \succeq 0,$$

$$XX^T = I.$$  

(6.4)

where $M = \begin{bmatrix} M + AXB^T + BX^T A^T & A & B \\ A^T & I & 0 \\ B^T & 0 & I \end{bmatrix}$ and $X = \begin{bmatrix} I & X \\ X^T & I \end{bmatrix}$ are Schur complement [33] for the first and the relaxed second constraint in (6.3).

We first present following results which will be used to prove the result regarding relaxation gap.
6.1 Semidefinite Relaxation for the Orthogonal Procrustes Problem

Lemma 6.1. For any matrix $Y \in \mathbb{R}^{n \times n}$

1. $YY^T \preceq I$ implies $|y_{ii}| \leq 1$, $i = 1, 2, \ldots, n$, where $y_{ii}$ is the $i^{th}$ diagonal element of $Y$.

2. $YY^T \preceq I$ and $y_{ii} = 1$, $i = 1, 2, \ldots, n$ implies $y_{ij} = 0$, $i \neq j, i, j = 1, 2, \ldots, n$.

Proof.

1. From $YY^T \preceq I$ we write $I - YY^T \succeq 0$. If $I - YY^T \succeq 0$, then all diagonal elements of $I - YY^T$ are greater or equal to zero [33, Section 4.2]. Hence $y_{ii}^2 \leq 1$ and $|y_{ii}| \leq 1$, $i = 1, 2, \ldots, n$.

2. For $I - YY^T \succeq 0$, using the same argument as above, it follows that

$$
y_{i1}^2 + y_{i2}^2 + \cdots + y_{in}^2 \leq 1, \quad i = 1, 2, \ldots, n.
$$

Equation (6.5) and $y_{ii} = 1$ implies $y_{ij} = 0$, $i \neq j$.

Lemma 6.2. For any matrix $Y \in \mathbb{R}^{n \times n}$ such that $YY^T \preceq I$ and a diagonal matrix $\Sigma \in \mathbb{R}^{n \times n}$ with $\sigma_i \geq 0$, $i = 1, 2, \ldots, n$, where $\sigma_i$ is the $i^{th}$ diagonal element of $\Sigma$,

$$
\text{tr}(Y \Sigma) \leq \text{tr}(\Sigma).
$$

Proof. To prove the inequality, we write

$$
\text{tr}(Y \Sigma) = \sum_{i=1}^{n} y_{ii} \sigma_i \leq \sum_{i=1}^{n} \sigma_i,
$$

which is true if $y_{ii} \leq 1$. Using the first point of Lemma 6.1, $YY^T \preceq I$ implies $|y_{ii}| \leq 1$, proving the claim.

Lemma 6.3. For the matrices $Y$ and $\Sigma$ as defined in Lemma 6.2 and $\sigma_i > 0$, $i = 1, 2, \ldots, n$,

$$
\text{tr}(Y \Sigma) = \text{tr}(\Sigma),
$$

if and only if $Y = I$.

Proof.

1. For $Y = I$, it trivially follows that $\text{tr}(Y \Sigma) = \text{tr}(\Sigma)$. 

80
2. It follows from the 1st point in Lemma 6.1 that \(-1 \leq y_{ii} \leq 1\), \(i = 1, 2, \ldots, n\). If

\[
\text{tr}(Y\Sigma) = \text{tr}(\Sigma),
\]

then

\[
\sum_{i=1}^{n} y_{ii}\sigma_i = \sum_{i=1}^{n} \sigma_i.
\]  

(6.10)

Since \(-1 \leq y_{ii} \leq 1\) and \(\sigma_i > 0\), \(i = 1, 2, \ldots, n\), it follows that (6.10) is true if and only if \(y_{ii} = 1\), \(i = 1, 2, \ldots, n\). It follows from the 2nd point in Lemma 6.1 that \(y_{ij} = 0\), \(i \neq j\), \(i, j = 1, 2, \ldots, n\), hence \(Y = I\).

\[
\square
\]

The following result regarding the gap between problem (6.1) and its relaxation (6.4) holds.

**Theorem 6.1.** There is no gap between problem (6.1) and its relaxation (6.4) and the solution of the relaxation is the optimal solution of (6.1) if \(A^TB\) is nonsingular.

**Proof.** To prove the theorem, following Section 3.2.1, we write an equivalent problem to (6.1) as

\[
\gamma := \max_{XX^T = I} \text{tr}(BX^TA^T),
\]

(6.11)

which has a unique solution when \(A^TB\) is nonsingular [40, Theorem 8.6]. Relaxing the constraint \(XX^T = I\) to \(XX^T \preceq I\), we write the relaxed convex problem as

\[
\gamma := \max_{XX^T \preceq I} \text{tr}(BX^TA^T).
\]

(6.12)

To prove that there is no gap, we need to show that

\[
\gamma = \gamma.
\]

(6.13)

Suppose \(X\) is a maximizer for (6.12) such that \(XX^T \preceq I\). We show that \(XX^T = I\) proving absence of the gap. To find the max in (6.12), we follow similar steps used to derive the solution of (6.11) in Section 3.2.1 and write

\[
\text{tr}(BX^TA^T) = \text{tr}(X^TA^TB),
\]

\[
= \text{tr}(X^TU\Sigma V^T),
\]

\[
= \text{tr}(V^TX^TU\Sigma),
\]

\[
= \text{tr}(Y\Sigma),
\]

(6.14)
where $Y := V^T X^T U$. Since $XX^T = X^T X \preceq I$ and $U$ and $V$ are unitary matrices obtained from the singular value decomposition of $A^T B$, it follows that $V^T X^T X V \preceq I$, which gives $YY^T = V^T X^T X V \preceq I$. From Lemma 6.2 it follows that

$$\text{tr}(Y \Sigma) \leq \text{tr}(\Sigma). \tag{6.15}$$

Since $A^T B$ is nonsingular, all singular values of $A^T B$ in $\Sigma$ are non-zero. From Lemma 6.3 it follows that $\text{tr}(Y \Sigma) = \text{tr}(\Sigma)$ if and only if $Y = I$, i.e. $V^T X^T U = I$. This is true if and only if $X = U^T V$. Note now that $XX^T = I$, hence proving the claim.

6.1.2 Relaxation of the OPP for Rotations

The OPP formulation for rotations needs to handle the additional nonlinear constraint \( \det(X) = +1 \), which cannot be directly handled in the SDR framework. The OPP for rotations is of much interest from a practical point of view for many applications [34; 37; 88].

One approach to handle the rotation problem in the SDP framework is proposed by [68]. We also refer to [73, Proposition 4.1] for more details. We present the main point in a simplified form.

Consider a symmetric matrix $Z \in \mathbb{R}^{4 \times 4}$, then an exact SDP representation of the convex hull of $SO(3)$ is given by

$$X(Z) = \begin{bmatrix} z_{11} + z_{22} - z_{33} - z_{44} & 2z_{23} - 2z_{14} & 2z_{24} + 2z_{13} \\ 2z_{23} + 2z_{14} & z_{11} - 2z_{22} + z_{33} - z_{44} & 2z_{34} - 2z_{12} \\ 2z_{24} - 2z_{13} & 2z_{34} + 2z_{12} & z_{11} - z_{22} - z_{33} + z_{44} \end{bmatrix}, \tag{6.16}$$

if and only if the matrix $Z$ satisfies the following constraints

$$Z \succeq 0,$$

$$\text{trace}(Z) = 1. \tag{6.17}$$

In the above expression $z_{ij}$ is the element of the matrix $Z$. To understand this fact, consider that positive semidefinite $4 \times 4$ matrices with both trace and rank 1 are parameterized by

$$Z = \frac{1}{a^2 + b^2 + c^2 + d^2} \begin{bmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & cd \\ ad & bd & cd & d^2 \end{bmatrix}. \tag{6.18}$$
such that $a^2 + b^2 + c^2 + d^2 \neq 0$. The image of this rank 1 matrix under the linear map (6.16) is precisely the group $SO(3)$. This parameterization is known as the Cayley transform [73].

Embedding the trace constraint, i.e. $z_{11} + z_{22} + z_{33} + z_{44} = 1$ within the definition of $Z$, we write an exact SDP representation for the rotation problem as

$$\begin{align*}
\min_{Z,M} & \quad \text{tr}(M) \\
\text{subject to} & \quad M \succeq 0, \\
& \quad Z \succeq 0,
\end{align*}$$

(6.19)

where

$$M = \begin{bmatrix} M + AX(Z)B^T + BX(Z)^T A^T & A & B \\ A^T & I & 0 \\ B^T & 0 & I \end{bmatrix}.$$  

(6.20)

The new $X(Z)$ with embedded trace constraint is given as

$$X(Z) = \begin{bmatrix} 2z_{11} + 2z_{22} - 1 & 2z_{23} - 2z_{14} & 2z_{24} + 2z_{13} \\ 2z_{23} + 2z_{14} & 2z_{11} + 2z_{33} - 1 & 2z_{34} - 2z_{12} \\ 2z_{24} - 2z_{13} & 2z_{34} + 2z_{12} & -2z_{22} - 2z_{33} + 1 \end{bmatrix}.$$  

(6.21)

The solution of this problem gives the matrix $Z$ satisfying the constraints given in (6.17). The matrix $Z$ is then used to calculate $X$, which is the optimal solution of the rotation problem. We use this transformation in Section 6.2 to solve the OPP for rotations with data uncertainties.

6.2 The OPP with Data Uncertainties

In various applications of the OPP, the data matrices $A$ and $B$ are generally obtained from different sources, e.g., some camera, sensors, mathematical models. This input information has always some sort of uncertainty, which could be large in some operational conditions. It is well-known that in the presence of large uncertainties the accuracy of the obtained solution is questionable due to its sensitivity to data errors. Clearly, under the worst case uncertainties, the error in the solution may be large.

To reduce the sensitivity of the solution to data uncertainties, in this section, we formulate and solve a robust Orthogonal Procrustes Problem.
6.2 The OPP with Data Uncertainties

6.2.1 The Robust Problem

We define the robust problem

$$\begin{align*}
\min_X \max_\Delta & \|\bar{A}X - \bar{B}\|_F^2 \\
\text{subject to} & \quad XX^T = I, \\
& \quad \Delta\Delta^T \preceq I,
\end{align*}$$

(6.22)

where $\bar{A}$ and $\bar{B}$ are uncertain data matrices and $\Delta$ is the uncertainty.

6.2.2 Uncertainty Representation in the Data Matrices

We consider the uncertainty structure in the data matrices described by

$$\begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} = \begin{bmatrix} A & B \end{bmatrix} + E\Delta \begin{bmatrix} F_1 & F_2 \end{bmatrix},$$

(6.23)

where $A$ and $B$ represent the nominal data, $E, F_1$ and $F_2$ are known matrices and $\Delta$ is the uncertainty matrix such that $\Delta\Delta^T \preceq I$.

Perturbation model of this form is common in robust estimation, filtering and control [32; 74; 83]. By a suitable selection of $E, F_1, F_2$ and $\Delta$, this model can represent both structured and unstructured uncertainty. For example, a norm-bounded full $\Delta$ represents unstructured uncertainty, while a norm-bounded diagonal $\Delta$ with a suitable choice of other matrices represents structured uncertainty. A suitable choice of $E, F_1$ and $F_2$ specifies both the components of $A$ and $B$ affected by the uncertainty $\Delta$ and also the amount of uncertainty.

In this work we use this general uncertainty structure to formulate the robust problem. The choice of the constant matrices defines the uncertainty to be structured or unstructured. Further, we consider 2-norm limit on $\Delta$, i.e. $\Delta\Delta^T \preceq I$, which represents a ball uncertainty.

6.2.3 Semidefinite Relaxation of the Robust OPP

To solve the min-max problem we use the SDR approach, following similar steps followed while solving the nominal problem (6.4). The main challenge is to handle uncertainties...
6.2 The OPP with Data Uncertainties

in the max problem. By expanding the cost in (6.22) we define
\[
\begin{align*}
\text{tr}(J(X, \Delta)) &= \|\tilde{A}X - \tilde{B}\|^2_F \\
&= \text{tr}
\begin{pmatrix}
(A + E\Delta F_1)XX^T(A + E\Delta F_1)^T - (A + E\Delta F_1)X(B + E\Delta F_2)^T \\
- (B + E\Delta F_2)^T(A + E\Delta F_1)^T + (B + E\Delta F_2)(B + E\Delta F_2)^T
\end{pmatrix}.
\end{align*}
\]

The constraint \(XX^T = I\) is used to further simplify the cost. To transform this problem into a tractable LMI formulation we first replace the cost in (6.22) with a linear objective function, which is trace of an unknown matrix \(M\), and write an equivalent problem

\[
\begin{align*}
\min_{X, M, \Delta} & \quad \text{tr}(M), \\
\text{subject to} & \quad M - J(X, \Delta) \succeq 0, \quad \Delta \Delta^T \preceq I, \\
& \quad XX^T = I.
\end{align*}
\]

We simplify the optimization problem by relaxing the first constraint and making it independent of \(\Delta\). For this, we use the identity

\[
\begin{pmatrix}
I & E\Delta
\end{pmatrix}
\begin{pmatrix}
M - J_n(X) - \lambda E E^T & T_2(X) \\
(T_2(X))^T & \lambda I - J_\Delta(X)
\end{pmatrix}
\begin{pmatrix}
I \\
\Delta^T E^T
\end{pmatrix},
\]

where \(J_n(X), T_2(X)\) and \(J_\Delta(X)\) are defined as

\[
\begin{align*}
J_n(X) &= AA^T + BB^T - AXB^T - BX^T A^T, \\
T_2(X) &= BX^TF_1^T + AXF_2^T - AF_1^T - BF_2^T, \\
J_\Delta(X) &= F_1F_1^T + F_2F_2^T - F_1XF_2^T - F_2XF_1^T.
\end{align*}
\]

Further we define the matrix in the second term on the right hand side of (6.26) as

\[
\mathcal{T}(M, X, \lambda) = 
\begin{pmatrix}
M - J_n(X) - \lambda E E^T & T_2(X) \\
(T_2(X))^T & \lambda I - J_\Delta(X)
\end{pmatrix}.
\]

The right hand side of (6.26) is either zero or positive because \(I - \Delta \Delta^T \succeq 0\) and we impose \(\lambda \geq 0\) and \(\mathcal{T}(M, X, \lambda) \succeq 0\), ensuring that \(M - J(X, \Delta) \succeq 0\), i.e., \(\text{tr}(M)\) is an
upper bound on the cost in (6.22). Finally we write a relaxation of (6.22) as
\[
\min_{X,M,\lambda} \quad \text{tr}(M), \\
\text{subject to} \quad T(M,X,\lambda) \succeq 0, \\
X \succeq 0, \\
\lambda \geq 0,
\]
(6.28)
where $X$ is as defined in (6.4).

**Remark 6.1.** Unlike Remark 6.1, it is observed through numerical simulations that the gap between the robust problem (6.22) and its semidefinite relaxation (6.28) is not zero in general.

### 6.2.4 Orthogonalization of $X$

While solving the robust problem, we need an $X$ which not only minimizes the worst case cost, but also satisfies the orthogonality constraint, which is relaxed in the SDR formulation. When the gap between the robust problem and its semidefinite relaxation is not zero, the $X$ does not satisfy the orthogonality constraint, i.e., $XX^T \neq I$. For such cases, we find the nearest orthogonal $X$ in the Frobenius norm sense. Such an $X$ can be obtained by solving the optimization problem
\[
\min_{X_o} \quad \|X_o - X\|_F^2, \\
\text{subject to} \quad X_oX_o^T = I,
\]
(6.29)
where the decision matrix $X_o$ is orthogonal. This problem is the same as the nominal Orthogonal Procrustes Problem (3.8) with $A = I$ and $B = X$. The solution is $X_o = U_sV_s^T$, where $U_s \Sigma_s V_s^T = X$ is the singular value decomposition of $X$.

### 6.2.5 Effect on Robust Performance

It can be argued that $X_o$ is not the optimal solution of (6.28). However, note that if the optimal solution $X$ of (6.28) is not orthogonal, it is not feasible for the robust problem (6.22). The matrix $X_o$ is the nearest orthogonal matrix, which also results in minimum cost variation than the non-orthogonal matrix $X$. To evaluate the cost variation, let $e_1 = \bar{A}X_o - \bar{B}$ and $e_2 = \bar{A}X - \bar{B}$, then $\|e_1 - e_2\| = \|\bar{A}(X_o - X)\| \leq \|\bar{A}\|\|X_o - X\|$. This shows that $\|\bar{A}\|\|X_o - X\|$ is an upper-bound on the cost difference. The smallest value
of the upper-bound ensures that the orthogonal $X_o$ results in a minimum cost variation from the SDR solution. However, the robustness properties of $X_o$ may not be the same as that of the robust matrix $X$. An analysis of the performance is presented in Section 6.3, however, the robustness properties of $X_o$ need further analysis.

### 6.2.6 The Robust Problem for Rotations

Using the same transformation as discussed in Section 6.1.2, we write the relaxed robust problem for rotations as

$$
\min_{Z,M,\lambda} \quad \text{tr}(M),
\quad \text{subject to} \quad T(M, Z, \lambda) \succeq 0, \quad Z \succeq 0, \quad \lambda \geq 0,
$$

(6.30)

where $T(M, Z, \lambda)$ is as in (6.27), with $X$ replaced by $X(Z)$.

Although the transformed formulation in the case of the OPP with rotations without uncertainties (6.19) is exact; however, in the presence of uncertainties, the obtained solution is not always orthogonal: we need to use the orthogonalization discussed in Section 6.2.4. Based on the numerical experience with the robust OPP, as given in Section 6.3, we have the following remark.

**Remark 6.2.** It is observed in the numerical simulations for both the standard OPP and the OPP for rotations that by limiting the size of the maximum uncertainty in the robust problem (6.28) significantly reduces the number of cases in which the solution $X$ is not orthogonal.

### 6.3 Simulation Results

This section presents numerical simulations to evaluate the performance of the presented relaxation approaches and to support presented discussions and results.

#### 6.3.1 Analysis of the Standard OPP Relaxation

Firstly we compare the solution of (6.4) with the SVD solution. For this comparison, we have generated random matrices $A$ and $B$ of size $10 \times 10$ using the MATLAB command `randn`. The results of the simulation are given in Figure 6.1. The graph presents and cost $\|AX - B\|^2_F$ using both the SVD and the SDR based solutions. It also gives the
6.3 Simulation Results

Figure 6.1: Simulation results for the semidefinite relaxation of the standard OPP (6.4).

The gap between the relaxed and the original problem, which is practically zero, supporting Remark 6.1. The last sub-plot shows the determinant of $X$, which is $\pm 1$ indicating that the solution belongs to orthogonal matrices of order $n$.

6.3.2 Analysis of the OPP Relaxation for Rotations

The performance of the OPP for rotations is illustrated in Figure 6.2. The first subplot presents the relaxation gap, which is zero, validating the exactness of (6.19) for rotations. The other subplots show some parameters of the $Z$ and $X$ matrices. It can be observed that the parameters, such as the trace and the rank of $Z$, and the determinant of $X$, are as desired for all random cases.

6.3.3 Robust Performance Evaluation

In this section we evaluate the performance of the solutions obtained by solving the approximate formulations of the standard and the robust problems for a set of bounded uncertainties in the input matrices. For this analysis, we considered a structured uncertainty description. To obtained data for this test we generated a random matrix $A$ using the MATLAB’s command `randn`, and an orthogonal matrix $X$ using the MATLAB’s
command \texttt{orth}, both in $\mathbb{R}^{3 \times 3}$. Using this orthogonal matrix, we calculate the matrix $B$. This pair $A, B$ represents an exact data set, i.e. the matrix $A$ can be exactly transformed to $B$ using $X$. To this pair we add uniformly distributed random error within a range of 30\% of the size of the elements of the true matrices, to obtain $A$ and $B$ with errors. We then solve both the nominal problem (3.13) and the robust problem (6.28). A large number of $\tilde{A}$ and $\tilde{B}$ have been generated by adding uniformly distributed random errors into the nominal data within the set of uncertainty bounds. The cost value has been evaluated both for the nominal and the robust solution. The histogram of the test results is given in Figure 6.3, where the $x$–axis represents the cost value, while the $y$–axis represents the number of test with the same cost value. It can be observed that the dispersion of the cost value using the nominal solution is much larger than the dispersion of the robust solution. This benefit is more obvious for the worst case scenarios. However, for nominal cases the robust solution suffers from an offset as compared to the nominal solution.
6.3.4 Orthogonalization Step

Finally, we analyze the orthogonalization step discussed in Section 6.2.4 and its effect on the optimal cost. We have performed simulations by varying the maximum uncertainty level in the nominal data and plotted the number of cases (in percentage) which give an orthogonal $X$. The results are shown in the first subplot of Figure 6.4. The analysis shows that the number of tests having orthogonal $X$, i.e., when no orthogonalization is required, increase significantly, when the size of the uncertainty is small, supporting Remark 2. Further, we compare the change in the cost using the optimal solution of the approximate robust problem (6.28) and its orthogonalized solution for 200 runs of random data within bounded uncertainty, showing that the cost variation due to the orthogonalization is not large.

6.4 Summary

We have presented a unified approach based on semidefinite relaxations to solve different formulations of the Orthogonal Procrustes Problem. It has been demonstrated that the
relaxation of the nominal problem results in no gap between the actual and the relaxed problems. The SDR framework allows to handle uncertainties in the data matrices. It has been further demonstrated that while considering uncertainties, the gap is not necessarily zero and, in such cases, the obtained solution needs to be orthogonalized.
Chapter 7

Sub-optimal Predictive Control for Satellite De-tumbling

This chapter deals with control of the satellite in the initial acquisition phase. In this phase, the launch disturbances induce high body rates resulting in a tumbling motion. The control system is required to minimize the de-tumbling time by damping the high body rates with minimal use of on-board power resources. We present performance analysis of three nonlinear model predictive control (NMPC) schemes to address these conflicting objectives. The first two formulations are based on existing results, which guarantee closed-loop stability; however they have high computational requirements. The computational burden becomes prohibitively high as compared to the available resources in such satellites, when the initial rates are large. To address this issue, a third formulation is proposed, which sacrifices stability guarantees and allows an early termination of the optimizer by imposing an additional constraint on the cost.

7.1 Introduction

The computational burden to optimally solve a nonlinear model predictive control (NMPC) problem is a major concern for practical applications. In this chapter, the main focus is to analyze the performance of NMPC for the satellite rate damping problem and to study the possibility of real-time implementation, which should take into account the system’s computational limitations. The use of NMPC is motivated to address the conflicting objectives, such as reducing the de-tumbling time using minimal on-board power resources in the presence of dynamic nonlinearities due to high body rates, under-action of magnetic actuators and time-varying control due to the change in the Earth’s magnetic
field at different locations of the orbit.

We analyze the performance of two existing NMPC schemes in comparison to $\beta$–dot control. The issues that arise in practice while solving a NMPC problem are highlighted, especially when the initial body rates are large. The main issues faced are infeasibility and the large number of iterations that the optimization solver takes to reach optimality. Limiting the maximum number of iterations may result in degraded performance due to the possibility of the cost increasing at each sample time. To overcome this problem, we propose a new sub-optimal formulation with an additional constraint on cost reduction. Although the imposed constraint adds slightly to the computational burden of the optimizer, we demonstrate through extensive numerical simulations that, by limiting the maximum iterations of the optimizer to a small number, the cost at each sample time decreases. The imposed cost reduction also gives significant performance improvement by reducing the de-tumbling time compared to the other NMPC formulations.

7.2 Optimal NMPC Formulations

In the MPC literature, most of the research is dominated by establishing stability guarantees, and as a result conditions have been developed for different formulations of MPC to ensure stability [59]. The classical MPC setup uses a terminal equality constraint to guarantee stability [46]. However, such strict stability constraints add to the computational burden of the optimizer. Sometimes satisfaction of the exact equality constraint cannot be achieved in finite computational time, especially for nonlinear systems [58; 75]. Research has been done in relaxing these constraints. One common approach is to relax the strict terminal equality constraint to an inequality, where the terminal state is required to be in a terminal region. A terminal cost is also added to penalize the states at the end of the prediction horizon [22]. The terminal region needs to be invariant for the nonlinear system controlled by a local state feedback control. A similar scheme, known as dual-mode MPC, is proposed in [61]. The reference [75] discusses MPC schemes with a strict terminal equality constraint and a scheme similar to the dual-mode MPC for nonlinear systems. Furthermore, in [43; 44] it has been shown that careful selection of either a terminal cost or horizon length can ensure stability of MPC. They have proposed selection strategies for both the terminal cost and horizon length in a continuous-time setting. In [48], a stabilizing discrete-time NMPC is discussed without terminal constraints, but used a terminal cost with a suitable selection of weighing matrices. Similar considerations for discrete-time systems have been recently reviewed in [35].
In this paper we assume that all states are measured that there is no mismatch between the model and the plant and that the control and prediction horizons are the same. Most of the discussions on the performance of the presented NMPC formulations are based on our numerical experience using a specific optimization solver, known as IPOPT \[87\], which uses a primal-dual interior-point algorithm to solve the optimization problem.

### 7.2.1 Formulation with Terminal Constraint

For the continuous-time nonlinear satellite attitude dynamics given in Section 7.4, we choose the following NMPC formulation, which ensures stability, subject to some standard regularity assumptions on the cost function, dynamics and input constraint set \[22\]

\[
\begin{aligned}
\min_{u(\cdot)} & \quad V(x(\cdot), u(\cdot), t) \\
\text{subject to} & \\
\qquad x(s) &= f(x(s), u(s), s), \quad \forall s \in [t, t + T] \\
\qquad u(s) &\in U, \\
\qquad x(t + T) &\in \mathbb{X}_f,
\end{aligned}
\]

where \(V(\cdot)\) is the cost function, \(f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n\) with \(f(0, 0, s) = 0\) \(\forall s \in (t, t + T)\), the initial state at time \(t\) is \(x(t)\) which is assumed to be known, \(U\) is a compact subset of \(\mathbb{R}^m\) containing the origin in its interior, \(x(t + T) := x^u(t + T; x(t), t)\), where \(x^u(\cdot; x(t), t)\) is the state trajectory with a given control trajectory \(u(\cdot)\) and initial state \(x(t)\) at time \(t\) and \(\mathbb{X}_f\) is a terminal constraint set. The NMPC cost function is given by

\[
V(x(\cdot), u(\cdot), t) := \int_t^{t+T} \ell(x(s), u(s)) ds + F(x(t + T)),
\]

where \(\ell(\cdot)\) is the stage cost, \(T\) is the prediction horizon and \(F(\cdot)\) is a terminal cost. For the stage cost, a common choice is a quadratic cost penalizing states and inputs, given as

\[
\ell(x(s), u(s)) := \frac{1}{2} (\|x(s)\|^2_Q + \|u(s)\|^2_R), \quad \forall s \in [t, t + T],
\]

where \(Q \succeq 0\) and \(R \succ 0\) are symmetric weighting matrices. For the satellite rate damping problem we have used this quadratic stage cost with a suitable choice of \(Q\) and \(R\). The state vector used in the stage cost comprises only the body rates.

While evaluating the MPC performance, we consider the case in which \(\mathbb{X}_f = \{0\}\) and
7.2 Optimal NMPC Formulations

$F(x(t + T)) = 0$. On the other hand, if an inequality terminal constraint is considered, the set $X_f$ needs to be an invariant set. We also need to find a suitable terminal cost and a stabilizing controller for the terminal region. In this work, since we are interested in bringing the satellite to rest, we can use rotational kinetic energy of the satellite as a terminal cost, similar to the one suggested in [15], i.e. $F(x(t + T)) = \frac{1}{2} \omega(t + T)^T J \omega(t + T)$, where $J$ is the inertia matrix of the satellite and $\omega$ represents the satellite body rates.

**Warm Starting**

To speed-up computations while solving the problem $F_1$, we use a warm start strategy [75], i.e. at time $t + \delta$, where $\delta$ is the sample time, we initialize the input and state trajectories with the shifted control and state trajectories obtained at time $t$. For the last interval $[t+T, t+T+\delta]$, we repeat the control trajectories of the interval $[t+T-\delta, t+T]$. At time $t$ with horizon length $T$, let the optimal (or early terminated) control trajectory be represented by $u^*(\cdot) := u^*(\cdot; x(t), t)$ and, using this control trajectory, the state trajectory be represented as $x^{u*}(\cdot; x(t), t)$. Using this information, we write the shifted trajectories at time $t + \delta$ for the time interval $[t + \delta, t + \delta + T]$ as

$$u_{\text{shifted}}(\cdot) := \begin{cases} u^*(s; x(t), t) & \forall s \in [t + \delta, t + T] \\ u^*(s - \delta; x(t), t) & \forall s \in [t + T, t + T - \delta] \\ 0 & \forall s \in [t + T + \delta, \infty) \end{cases}$$

(7.4)

$$x_{\text{shifted}}(\cdot) := x^{u_{\text{shifted}}}(\cdot; x(t + \delta), t + \delta).$$

The warm starting technique improves the optimizer’s performance by reaching an optimal point in less number of iterations than the case without warm starting.

**7.2.2 Formulation without Terminal Constraint**

The problem $F_1$ given in (7.1) performs well with small initial body rates ($\leq \pm 5$ deg/s) as shown in Section 7.5. However, when these rates increase, the number of iterations of the optimizer increases considerably. Infeasibility can also arise in these cases due to terminal equality constraint, which is mainly imposed to ensure stability.

Since we are not considering any other state constraint, the terminal constraint is the only one that depends on the predicted state of the system. Thus, removal of the terminal constraint makes the problem much easier to solve [48]. To ensure stability, we use the fact that for a sufficiently long horizon, the MPC scheme is stabilizing even without terminal cost and constraints. Following the results presented in [44, Theorem
7.3 Sub-optimal NMPC Formulation

For initial body rates less than \( \pm 5 \) deg/s, both problems \( F_1 \) and \( F_2 \) give similar performance if the horizon is sufficiently long (see Figure 7.3). However, during the de-tumbling phase of the satellite, the initial body rates can be much higher. When initial rates are larger than \( \pm 5 \) deg/s, performance of these formulations degrade significantly. The main reason of the degraded performance with \( F_1 \) is infeasibility, i.e. with limited control torque, it becomes very difficult for the optimizer to find a control sequence, which satisfies all constraints. With \( F_2 \), due to the removal of the terminal constraint, infeasibility issues are avoided, however, the optimizer still needs a large number of iterations. This enhanced computational load is not acceptable for the computational resources normally available in small satellites. With limited computational resources, one possible approach is to fix the maximum number of iterations for the optimizer. However, by limiting the maximum number of iterations, we obtain a sub-optimal point that may result in an increase in the cost at each sample time, thereby loosing stability guarantees. If the early terminated points result in a continuous increase in the cost during first few sample times, we may also loose the benefit of reduced de-tumbling time. This issue is further discussed in Section 7.5.

Sub-optimality in MPC has been discussed in the literature; see for example [48; 75]. Both the references impose an additional constraint on the monotonic decrease in the cost
7.3 Sub-optimal NMPC Formulation

at each sample time and concentrate on feasibility, instead of optimality. They propose early termination of the optimization process when a feasible point is obtained. This feasible point, which ensures a monotonic decrease in the cost, also ensures closed-loop stability. The references [89; 96] discuss sub-optimality in terms of early termination of the optimizer based on the available time or maximum number of iterations. Both references treat linear systems and propose using primal barrier interior-point methods to solve the optimization problem. The reference [89] has demonstrated through numerical simulations that sub-optimal results with very small number of iterations (e.g. 5-10) perform surprisingly well. The reference [96], however, impose an extra constraint on the cost reduction. With the availability of a feasible solution at the start, the imposed constraint ensures stability for any available computational time for linear systems, when no disturbances are considered. In the presence of disturbances, they have proposed a robust real-time technique using tube-based MPC, which ensures input-to-state stability. Most of these methods assume availability of a feasible solution at the start to ensure stability; however, this assumption is not always true in the satellite de-tumbling phase when initial body rates are large. Moreover, the proposals to terminate the optimization based on number of iterations or computational time have been tested with linear systems only.

Following similar ideas, we derive a formulation which gives an acceptable performance with early terminated sub-optimal points for the tumbling satellite. We modify problem $F_2$ and add an additional constraint on cost reduction. To pose the cost constraint we define the value function at time $t - \delta$ as $V^*(x(t - \delta), t - \delta)$, which uses input sequence $u^*(\cdot, x(t - \delta), t - \delta)$ obtained at last time step. The proposed problem is given as

$$
\begin{align*}
\text{min}_{u(\cdot)} & \quad V(x(\cdot), u(\cdot), t) \\
\text{subject to} & \quad \dot{x}(s) = f(x(s), u(s), s), \quad \forall s \in [t, t + T] \\
& \quad u(s) \in U, \\
& \quad V(x(\cdot), u(\cdot), t) \leq V^*(x(t - \delta), t - \delta) - \gamma \|x(t)\|^2_2,
\end{align*}
$$

(7.7)

where $V(x(\cdot), u(\cdot), t)$ is the same as defined for the formulation $F_2$ and $\gamma > 0$ is a parameter to choose, which defines the required decrease in cost. We have kept the decrease in the cost as a function of state only, which is a milder requirement than the stage cost, as proposed in [75].
Algorithm 1: A Sub-optimal MPC Algorithm

Input/Parameters
- Dynamic equations $f(x(s), u(s), s) \forall s \in [t, t + T]$ and input constraint set $U$, both satisfying conditions given in Section 7.2.1.
- $Q \succeq 0$ and $R > 0$ for the quadratic stage cost (7.3).
- Initial state and input at time $t$, i.e., $x(t), u(t)$.
- Tuning parameter $\gamma > 0$, maximum iterations $\text{iter}_{\text{max}}$ and cost tolerance $\epsilon_{\text{cost}} > 0$.

Output
- Optimal or sub-optimal input trajectory $u^*(\cdot)$.

Algorithm

1: At initial time $t$, solve $F_2$ for $i \leq \text{iter}_{\text{max}}$, where $i$ represents the number of iterations, using the given initial state and input and find an optimal or sub-optimal control trajectory $u^*(\cdot)$. Use the first value of $u^*(\cdot)$, i.e., $u^*(t; x(t), t)$ for control.

2: Increment time by $\delta$. Use the warm start scheme (7.4) to obtain initial state and input trajectories. Solve $F_3$ optimally or sub-optimally and find $u^*(\cdot)$. Use the first value of the obtained control trajectory.

3: If the value function at last time step $V^*(\cdot) \geq \epsilon_{\text{cost}}$, repeat step 2, otherwise solve $F_2$ instead of $F_3$ in step 2.

Using the warm start scheme (7.4) and problem $F_3$, a sub-optimal MPC strategy is given in Algorithm 1. The sub-optimal algorithm with a limit on the maximum number of allowed optimizer iterations has shown improved performance in terms of less detumbling time compared to other two problems, especially when initial rates are larger than $\pm 5$ deg/s. In these cases, $F_1$ and $F_2$ either face infeasibility or require large number of iterations to reach a solution. Some numerical simulation results to demonstrate the performance are given in Section 7.5. The improved performance with a limited number of optimizer iterations is mainly due to the imposed cost reduction constraint, which is respected by the optimizer despite sub-optimality. The $3^{rd}$ step in Algorithm 1 also gives computational benefit by removing the cost constraint when $V(x(\cdot), u(\cdot), t) < \epsilon_{\text{cost}}$. This transition may give an increase in the cost after a small number of iterations; however since the value of the cost is already small, it does not affect the performance.

7.4 Satellite Dynamics

In Chapter 2 both translational and rotational dynamic equations have been presented. The translational dynamics are mainly required to predict the Earth’s magnetic field, which depends on the current position of the satellite. The MPC requires this information to calculate the control action at the current time. However, keeping in view the
For rotational dynamics, the state vector $x$ and the control vector $m$ are given as

$$
x := \left[ (\omega_{b/o}^b)^T \quad (q_{b/o}^b)^T \right]^T, \quad m := \left[ m_x \quad m_y \quad m_z \right]^T,
$$

(7.8)

where $\omega_{b/o}^b := \left[ P \quad Q \quad R \right]^T$ is the angular velocity of $F_b$ with respect to $F_o$ and $q_{b/o}^b \in \mathbb{R}^4$ satisfying $q_{b/o}^T q_{b/o} = 1$ represents the quaternion vector for transformation from $F_o$ to $F_b$. The control vector $m$ represents the dipole moment of the magnetic actuators installed about the corresponding axis of the body frame.

The state equations, given in Chapter 2 in (2.11) and (2.16), are recalled, namely

$$
\dot{b}\omega_{b/o}^b = (J^b)^{-1} \left[ \sum b^b \tau - \omega_{b/i} b \times J^b \omega_{b/i} b \right] + \Omega_{b/o}^b C_{b/i} \omega_{b/o}^b.
$$

(7.9)

Figure 7.1: Performance comparison between solving $F_1$ and $F_2$. 

complexity of NMPC, we only use the rotational dynamics, while to estimate the Earth’s magnetic field at different orbit locations we use a very simple approximation given in (2.21). In this section we recall the rotational dynamic equations for completeness.
7.5 Simulation Results

To evaluate the performance of the NMPC schemes we use the nonlinear simulation setup for the CubeSat discussed in Section 2.6, which is based on the data given in Table 2.2 [85]. The main objective of these simulations is to study and compare the performance of the presented NMPC schemes to damp the initial body rates of the tumbling satellite to an equilibrium in a minimum time with minimal use of control. The control command \( m \) at each time step is the first element of the optimal or sub-optimal control trajectory \( u(\cdot) \), computed using the presented NMPC formulations.

Nonlinear optimization solver

To solve the NMPC problem we use a software package for large scale nonlinear optimization problems, IPOPT (Interior Point OPTimizer) [87], along with a MATLAB based toolbox ICLOCS (Imperial College London Optimal Control Software) [94]. IPOPT implements a primal-dual interior point method for nonlinear optimization. ICLOCS is used to transcribe the infinite-dimensional optimal control problem into a finite-dimensional approximation. It offers three transcription methods, i.e. discrete, multiple-shooting and direct collocation. We have used a direct collocation method to discretize the continuous-time system dynamics using one of the three available integration methods in ICLOCS, namely, Euler, Trapezoidal and Hermite. The results presented in this section are based on simulations that use the trapezoidal method.

7.5.1 Performance Comparison between Solving Problems \( F_1 \) and \( F_2 \)

A performance comparison of solving problems \( F_1 \) and \( F_2 \) for small initial body rates (less than ±5 deg/s) is given in Figure 7.1. A horizon length of 480 seconds has been chosen with a sample time of 3 seconds. The results indicate that the chosen horizon length is sufficiently long and as a result solving problem \( F_2 \) gives almost similar performance as solving \( F_1 \). It can be observed from the second subplot, where the number of iterations is plotted against time, that solving problem \( F_1 \) requires around 250 iterations for the
7.5 Simulation Results

first sample time, while $F_2$ requires around 80. However, once a feasible point is obtained to warm start the state and control trajectories at the next sample time, the number of iterations drops significantly and is almost the same for both formulations.

7.5.2 Comparison with $\beta$-dot Control

One of the most commonly used control schemes for satellite rate damping is known as $\beta$-dot control [29], which uses the rate of change of the Earth’s magnetic field and is based on the principle of reducing the rotational kinetic energy of the tumbling satellite. The main reason for the popularity of this scheme is simple measurement requirements, i.e., the rate of change of the Earth’s magnetic field with respect to the body frame. The Earth’s magnetic field in the body frame is measured by magnetometers and its derivative is calculated numerically. A comparison between $\beta$-dot control and solving NMPC problem $F_1$ is shown in Figure 7.2. For $\beta$-dot control, a well-tuned controller gain is selected. For initial gain selection, the method given in Section 2.6 is used, which is further tuned using nonlinear simulations. The initial rates, gain used for $\beta$-dot controller and some important parameters for the NMPC are given in Table 7.1. It is observed that $\beta$-dot control damps the body rates to zero in a time around half of the orbit time ($\approx 3000$ seconds), while the NMPC schemes generate an optimal control command, keeping in view future response of the satellite over the chosen horizon. It can be observed from Figure 7.2 that the reduction in the rotational kinetic energy of the satellite is much faster with the NMPC as compared to $\beta$-dot control. For control energy, it can be observed from Figure 7.2 that the NMPC controller use more control at start; however, it quickly settles to small values, while due to longer de-tumbling time, the $\beta$-dot controller use more control.

Table 7.1: Simulation parameters for MPC and $\beta$-dot performance comparison.

<table>
<thead>
<tr>
<th>Parameter Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial body rates</td>
<td>$P = 0.04 , \text{rad/s}$, $Q = 0.02 , \text{rad/s}$, $R = -0.03 , \text{rad/s}$</td>
</tr>
<tr>
<td>$\beta$-dot gains</td>
<td>$K_x = 5 \times 10^5$, $K_y = 5 \times 10^5$, $K_z = 5 \times 10^5$</td>
</tr>
<tr>
<td>MPC parameters</td>
<td>$T = 480s$, $\delta = 3s$, $Q_c = I$, $R_c = 1 \times 10^{-6} I$</td>
</tr>
</tbody>
</table>
7.5 Simulation Results

7.5.3 Suboptimal MPC Performance

With large initial body rates, problem $F_1$ faces infeasibility at the start. To demonstrate this issue we performed two tests with initial rates $P = 6.9$ deg/s, $Q = -9.7$ deg/s and $R = 6.3$ deg/s. In the first test the maximum number of IPOPT iterations has been fixed to 300, while in the second test, to 50. The comparison is given in Figure 7.3. It can be observed that with 300 iterations, the problem $F_1$ remains infeasible for the first 150 seconds. The problem $F_2$ without terminal equality constraint also faces infeasibility during the first few iterations, mainly because of numerical reasons; however, the solution becomes feasible early on. However, during the time when the solution is sub-optimal due to early termination, the cost value increases to almost double the initial value. Afterwards, the decrease in the cost is very slow, which results in a degraded rate damping response. A similar response is observed when the maximum number of iterations is limited to 50. However, by imposing the additional cost reduction constraint, a significant improvement in performance is observed. A comparison of the performance given by

Figure 7.2: Rate damping performance between solving $F_1$ and using $\beta$-dot control.
7.5 Simulation Results

solving problem $F_2$ via Algorithm 1 is shown in Figure 7.4. Due to the additional constraint in Algorithm 1, the cost decreases even when the optimizer terminates early, i.e., the iterate is sub-optimal.

![Simulation Results](image)

Figure 7.3: Performance of the IPOPT solver while solving NMPC problems $F_1$ and $F_2$ with initial rates $P = 6.9$ deg/s, $Q = -9.7$ deg/s and $R = 6.3$ deg/s. Comparison is given for two cases, i.e., one with maximum iterations set to 300 and second with 50.

Two more tests have been performed to further analyze the performance. Firstly, the effect of reducing the number of maximum iterations was studied by limiting the maximum number of iterations to 50, 20 and 10. Secondly, the effect of $\gamma$ in (7.7) is studied. The results are shown in Figure 7.5. It can be observed that reducing the maximum number of iterations to a very small number may degrade performance. It can be observed that reducing the maximum number of iterations to a very small number may degrade performance. With decreasing $\gamma$, the rate of cost decrease is not as expected, which shows that for large $\gamma$ the constraint is not being satisfied.

Lastly, to show the effect of the cost reduction constraint, a comparison of decrease in the rotational energy of the satellite using $\beta$-dot control, solving problem $F_2$ and
Figure 7.4: Performance comparison between solving $F_2$ and $F_3$ for maximum iterations set to 50.
7.6 Summary

This chapter has analyzed the performance of three nonlinear model predictive control (NMPC) schemes to reduce the de-tumbling time of the satellite. Firstly, two standard formulations with guaranteed stability have been presented, which give improved rate

Algorithm 1 is presented in Figure 7.6. For this test, initial body rates are kept less than ±5 deg/s. With these initial rates problem $F_2$ gives an optimal point for all time steps, while in Algorithm 1 the maximum number of iterations is fixed to 20. It can be observed from the plot that due to the cost reduction constraint, Algorithm 1 results in a faster reduction in the rotational kinetic energy compared to the problem $F_1$ and $\beta$-dot control, although this is obtained at the expense of more control energy. However, since the states and control go to zero much faster, the overall power consumption is less.

7.6 Summary

This chapter has analyzed the performance of three nonlinear model predictive control (NMPC) schemes to reduce the de-tumbling time of the satellite. Firstly, two standard formulations with guaranteed stability have been presented, which give improved rate
damping performance when compared to $\beta$-dot control; however they have high computational requirements, especially when the initial rates are high. To address the issue of high computational requirements for a satellite with limited computational and control resources, a third formulation has been proposed, which scarify stability guarantees and use an early termination in the optimizer. An additional constraint has been imposed on the cost. Due to this additional constraint, it has been observed in numerical simulations that with early optimizer termination, the cost decreases despite sub-optimality. A suitable choice of $\gamma$ gives an additional benefit by reducing the body rates to zero faster than other two formulations, reducing the de-tumbling time. These results motivate to study establishing theoretical guarantees for the performance, which is a possible future research direction.

Figure 7.6: Rotational energy reduction of a well-tuned $\beta$-dot controller compared to $F_2$ and Algorithm 1.
Chapter 8

Conclusions

This chapter summarizes the main contributions of the thesis. Possible future research directions are also discussed.

8.1 Main Contributions

This research has focused on robust attitude estimation for a small satellite in the initial acquisition phase. A satellite rate damping problem has also been addressed using a sub-optimal predictive control strategy. The main contributions of this research are summarized now.

- In Chapter 4, have we formulated a robust static attitude estimation problem, based on a weighted least squares approach with norm bounded data uncertainties. An uncertainty model has been proposed. This model sufficiently captures most of the realistic error sources for the considered application. In the robust formulation, the optimization variable is a matrix in $SO(3)$, i.e., it is orthogonal with unit determinant. Using the proposed uncertainty model and a quaternion transformation, we have simplified the robust problem by replacing the matrix variable with a quaternion vector with only an orthogonality constraint.

- Finding an optimal solution of the simplified robust min-max problem is difficult. The main reason for this difficulty is the fact that the maximization is convex in the uncertainty parameter, making it difficult to guarantee the existence of a unique solution and therefore its computation. To develop a tractable method for solving this robust optimization problem, we have determined an analytical upper bound on the maximization term and transformed the min-max problem into a suboptimal
minimization problem. We have also proposed a regularization to improve the performance. The approximate formulation, however has a non-convex quadratic cost and quadratic constraints.

- Due to non-convexity of the approximate formulation, there is no guarantee to find a global solution. Moreover, the computational cost of solving the non-convex QCQP may make it unsuitable for practical applications. To efficiently solve this problem, in Chapter 5, we have used a semidefinite relaxation (SDR) approach and transformed the non-convex QCQP into a semidefinite program (SDP) with a linear cost and linear matrix inequality constraints. The resulting SDP can efficiently be solved using existing interior point methods for semidefinite programming. It has also been shown how to extract the attitude information from the solution of the relaxed formulation.

- We have analyzed the new relaxed formulation and studied its optimality properties. We have shown that there is no gap between the non-convex QCQP and its semidefinite relaxation if only one of the eigenvalues of the LMI constraint matrix is zero, showing that the quaternion extracted from the SDR is the solution of the approximate robust problem formulated in Chapter 4.

- In Chapter 5, we have used semidefinite relaxations as an effective tool for solving the robust attitude estimation problem. In Chapter 6 we have extended this idea to a more general class of problems, known as the Orthogonal Procrustes Problem (OPP). We have presented relaxed formulations for the standard and the rotation OPP. In the relaxed formulations the associated non-convex constraints have been relaxed to convex approximations. It has been shown that both relaxed formulations are exact, giving an optimal solution.

- We have formulated robust problems for the standard and rotation OPP and applied the SDR framework used for the nominal OPP to obtain a tractable LMI formulation. It has been observed that in the presence of uncertainties the gap between the solution of the actual and the relaxed problem is not necessarily zero and in many cases the solution matrix is not orthogonal, making it infeasible for the original robust problem. This happens due to relaxation of the non-convex orthogonality constraint. In such cases, re-orthogonalization has been proposed to obtain the nearest orthogonal matrix in the Frobenius norm sense. Simulations
have shown that the number of occurrences when the solution is not orthogonal can be significantly reduced when the uncertainty bounds are small.

- In Chapter 7 we have studied the satellite rate damping problem using magnetic actuators. We have analyzed the performance of two nonlinear model predictive control schemes with guaranteed closed-loop stability. Numerical results have shown that the NMPC schemes give improved performance in terms of reducing detumbling time when compared with a commonly used control technique based on the derivative of the Earth’s magnetic field. However, computational requirements of the NMPC schemes are much higher.

- With large body rates, the computational burden of the NMPC schemes becomes prohibitively expensive. For these situations we have presented an algorithm that allows an early termination of the optimizer by imposing an additional constraint on cost reduction. However, due to sub-optimality it is difficult to give stability guarantees. The presented algorithm significantly reduces the de-tumbling time due to the imposed reduction in cost.

## 8.2 Future Research Directions

Some ideas for future research are listed below.

- In Chapter 4, we have derived an approximation of the robust attitude estimation problem. Future work can look at improving the approximations. For example, a tighter upper bound for the convex maximization can be sought. Similarly, a systematic procedure to find the optimal tuning parameter $\eta$ in the regularization term can be investigated. These investigations may improve the performance of the robust attitude estimation.

- It would also be useful to perform the computational complexity analysis of an algorithm to solve the approximate nonlinear robust formulation (4.18) and the SDP formulation (5.3) using state of the art optimization methods. It may be possible that a customized algorithm, which exploits specific properties of the non-convex optimization problem, may give a comparable or even better performance than the SDP. For such nonlinear optimization problems with an orthogonality constraint, apart from standard nonlinear optimization methods, algorithms based on manifolds, such as those mentioned in [1], can be investigated.
8.2 Future Research Directions

- In Chapter 5, we have analyzed the relaxation gap for the cases when the matrix $L_{11}$ in (5.3) has only one eigenvalue equal to zero. It would be useful to study the relaxation gap when the number of zero eigenvalues of $L_{11}$ are greater than one. Moreover, the problem to extract the optimal quaternion with more than one eigenvalues equal to zero, still needs to be addressed.

- The application of semidefinite relaxation to solve the robust Orthogonal Procrustes Problem in Chapter 6 needs further analysis of the relaxation gap. It will be useful to either modify the uncertainty structure or propose some conditions in the optimization problem to ensure an orthogonal solution, which would be feasible for the robust problem.

- In Chapter 7 it has been demonstrated through numerical simulations that the proposed sub-optimal NMPC scheme gives good performance for reducing the detumbling time. It has been observed that even limiting the number of maximum iterations to a reasonably small number does not violate the cost reduction constraint most of the time despite sub-optimality. As a result, the cost and hence the rates go to zero much faster than both the classical control schemes, such as $\beta$-dot control as well as the standard NMPC schemes presented in Chapter 7. However, due to sub-optimality, it is difficult to give performance and stability guarantees. An investigation of the conditions which help in establishing these guarantees could be a possible future direction.

- Lastly, the effect of the state estimation errors and disturbances has not be studied while evaluating the performance of NMPC schemes. It would be useful to take into account these uncertainties and study how the robustness margins of the NMPC is affected. This study can follow similar lines as those proposed recently in [42].
Appendix A

Definition of the Satellite Orbit Parameters

This appendix introduces some definitions of the orbit parameters, which are used in Chapter 2 while defining coordinate frames, satellite dynamics and the Earth’s magnetic field modeling. These definitions are based on [77]

Vernal Equinox (\(\gamma\))

The vernal equinox is the direction of intersection of the Earth’s equatorial plane and the plane of the Earth’s orbit around the sun (ecliptic), when sun crosses the equator from south to north in an apparent annual motion along the ecliptic, as shown in Figure A-1.

Orbit Inclination (\(i\))

The orbit inclination is the angle between the plane of the satellite’s orbit and the Earth’s equatorial plane.

Right Ascension of the Ascending Node (\(\Omega\))

The right ascension of the ascending node is the angle in the equatorial plane from the vernal equinox to the ascending node line, which is the direction of the intersection of the orbit’s plane and the equatorial plane, on the ascending side of the orbit [77]. This parameter is also known as longitude of ascending node.
Figure A-1: Geometry of the Earth and the satellite orbits

**Argument of Latitude** \( (u) \)

The argument of latitude is the angle from the ascending node, measured in the orbit’s plane, to the satellite’s location in the orbit.

**Satellite Orbit Rate** \( (\omega_o) \)

The satellite orbit rate is the rate with which the satellite moves around the Earth center of mass. As only circular orbits are considered in this thesis, the orbit rate is a constant given as \( \omega_o = 1.0732e - 3 \text{ rad/s} \).

**Earth Rotation Rate** \( (\omega_e) \)

The Earth rotation rate is the rate with which the Earth moves around its spin axis and is given as \( \omega_e = 7.2921 \times 10^{-5} \text{ rad/s} \).

**Latitude** \( (\psi_e, \phi_e) \)

The latitude is an angular measurement ranging from 0° at the equator to 90° at the poles. Two types of latitudes are used. The geocentric latitude \( (\psi_e) \) is the angle between the equatorial plane and the line joining the point to and the Earth’s center, while the
Geodetic latitude ($\phi_e$) is the angle between the equatorial plane and the line passing through the point that is normal to the surface of the Earth. Both latitudes are equal at the equator and the poles.

**Longitude ($l_e, \lambda_e$)**

The longitude is measured in the equatorial plane and is defined as the angular distance of a point’s meridian from the Prime Meridian. Sometimes, a celestial longitude ($\lambda_e$) is also used, which is the angular distance of a point’s meridian from the x-axis of the inertial frame $x_i$. 
Appendix B

Coordinate Transformations

This appendix gives the definitions of some coordinate transformation matrices which are used in the derivation of the dynamic equations in Chapter 2. The transformation matrix from the coordinate system $a$ to the coordinate system $b$ is represented by $C_{b/a}$ and orthogonality of the transformation matrix allows to write $C_{a/b} = C_{b/a}^T$.

Orbit to Inertial frame ($C_{i/o}$)

This transformation can be obtained by three rotations. The first positive rotation is about the $y_o$ axis of the orbit frame of the angle $\kappa_1 = 90 + u$ where $u$ is the argument of latitude at the satellite’s current position. The second positive rotation is about the rotated axis $x'_o$ of the angle $\kappa_2 = 90 - i$, where $i$ is the orbit inclination. The third rotation is about the rotated $z''_o$ axis of the angle $\kappa_3 = -\Omega$, where $\Omega$ is the angle between the line passing through the vernal equinox and the line passing through the intersection of the orbit’s plane and the Earth’s equatorial plane in the upward orbit motion, also known as the right ascension of the ascending node (RAAN). The transformation matrix from the orbit to the inertial frame is given by

\[
C_{i/o} = \begin{bmatrix}
\cos(\kappa_3) & \sin(\kappa_3) & 0 \\
-\sin(\kappa_3) & \cos(\kappa_3) & 0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0 \\
0 & \cos(\kappa_2) & \sin(\kappa_2) \\
0 & -\sin(\kappa_2) & \cos(\kappa_2)
\end{bmatrix} \begin{bmatrix}
\cos(\kappa_1) & 0 & -\sin(\kappa_1) \\
0 & 1 & 0 \\
\sin(\kappa_1) & 0 & \cos(\kappa_1)
\end{bmatrix}
\]  

(B-1)
Inertial to Earth Frame \((C_{e/i})\)

At any instant during the satellite motion, the x-axis of the rotating Earth’s frame \((x_e)\) is at an angle \(\kappa_4 = \lambda_{e0} + \omega_e t\) with respect to the x-axis of the inertial frame, where \(\lambda_{e0}\) is the celestial longitude of the satellite with respect to vernal equinox line at the time of satellite launch, \(\omega_e\) is the Earth’s rotation rate and \(t\) is the time since launch of the satellite. For the transformation from the inertial to the Earth’s frame, a single rotation about the \(z_i\)-axis equal to \(\kappa_4\) is required. The transformation matrix is given by

\[
C_{e/i} = \begin{bmatrix}
\cos(\kappa_4) & \sin(\kappa_4) & 0 \\
-\sin(\kappa_4) & \cos(\kappa_4) & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

(B-2)

Inertial to Body Frame \((C_{b/i})\)

The transformation from the ECI to the body frame can be obtained using \(C_{b/o}\) and \(C_{i/o}\) using the relation

\[
C_{b/i} = C_{b/o} C_{o/i},
\]

(B-3)

where \(C_{b/o}\) is obtained by integrating the kinematic equations at each time instant.
Appendix C

Davenport Transformation

The Devenport transformation is used in Section 3.1.3. All terminology used in this appendix is also defined in Chapter 3. To derive the Davenport transformation consider the cost function given in (3.2), i.e., $\text{tr}(WB^T CR)$. We use two properties of the trace. The trace is invariant under cyclic permutations, and $\text{tr}(\sum_i A_i) = \sum_i \text{tr}(A_i) \forall A \in \mathbb{R}^{n \times n}$. Using these properties we write

$$\text{tr}(WB^T CR) = \text{tr}(CRW^T B^T) = \text{tr}(CB^T(B,R)), \quad \text{(C-1)}$$

where $B^T(B,R) = (BWR^T)^T = RWB^T$. Now we represent $C$ using the quaternion $q := [q^T \ q_4]^T$, where $q := [q_1 \ q_2 \ q_3]^T$, written as [76]

$$C = (q_4^2 - q^T q)I + 2qq^T + 2q_4 Q. \quad \text{(C-2)}$$

In the above equation $Q := -q \times$, where $\times$ represents the vector cross product, given as

$$Q = \begin{bmatrix} 0 & q_3 & -q_2 \\ -q_3 & 0 & q_1 \\ q_2 & -q_1 & 0 \end{bmatrix}. \quad \text{(C-3)}$$

Substituting (C-2) in (C-1) yields

$$\begin{align*}
\text{tr}(WB^T CR) & = \left(q_4^2 - q^T q\right)\text{tr}(B^T(B,R)) + 2\text{tr}(qq^T B^T(B,R)) \\
& \quad + 2q_4 \text{tr}(Q B^T(B,R)). \quad \text{(C-4)}
\end{align*}$$
The second term on the right hand side of the above equation can be written as

\[ 2 \text{tr}(qq^T B^T(B, R)) = q^T(B^T(B, R) + B(B, R))q. \]  

(C-5)

The last term can be written as

\[ 2q_4 \text{tr}(QB^T(B, R)) = q_4(q^Tz(B, R) + z^T(B, R)q), \]  

(C-6)

where \( z^T(B, R) = (B \times R)W \). Substituting (C-5) and (C-6) in (C-4) and dropping arguments of \( B(B, R) \) and \( z(B, R) \) for simplification, we get

\[
\text{tr}(WB^T CR) = \begin{bmatrix} q \\ q_4 \end{bmatrix}^T \begin{bmatrix} B^T + B - \text{tr}(B)I & z \\ z^T & \text{tr}(B) \end{bmatrix} \begin{bmatrix} q \\ q_4 \end{bmatrix} = q^T K(B, R)q, \]  

(C-7)

which is the required form, where \( K(B, R) \) is a symmetric and indefinite matrix defined in (3.4).
Appendix D

Definition of a few Matrices used in Equation 4.6

This appendix gives definitions of two sets of matrices used in (4.6), i.e., $K^l_{ri}$ and $K^l_{bi}$, where $l = 1, 2, 3$ and $bi$ and $ri$ represent the $i^{th}$ measurement and model vector, respectively.

$K^l_{ri}$

\[
K^1_{ri} = \begin{bmatrix}
  r_{i1} & r_{i2} & r_{i3} & 0 \\
  r_{i2} & -r_{i1} & 0 & -r_{i3} \\
  r_{i3} & 0 & -r_{i1} & r_{i2} \\
  0 & -r_{i3} & r_{i2} & r_{i1}
\end{bmatrix},
\]

\[
K^2_{ri} = \begin{bmatrix}
  -r_{i2} & r_{i1} & 0 & r_{i3} \\
  r_{i1} & r_{i2} & r_{i3} & 0 \\
  0 & r_{i3} & -r_{i2} & -r_{i1} \\
  r_{i3} & 0 & -r_{i1} & r_{i2}
\end{bmatrix},
\]

\[
K^3_{ri} = \begin{bmatrix}
  -r_{i3} & 0 & r_{i1} & -r_{i2} \\
  0 & -r_{i3} & r_{i2} & r_{i1} \\
  r_{i1} & r_{i2} & r_{i3} & 0 \\
  -r_{i2} & r_{i1} & 0 & r_{i3}
\end{bmatrix}.
\]
\[ K_{bi}^l \]

\[
K_{bi}^1 = \begin{bmatrix}
b_{i1} & b_{i2} & b_{i3} & 0 \\
b_{i2} & -b_{i1} & 0 & r_{i3} \\
b_{i3} & 0 & -b_{i1} & -b_{i2} \\
0 & b_{i3} & -b_{i2} & b_{i1}
\end{bmatrix},
\]

\[
K_{bi}^2 = \begin{bmatrix}
b_{i1} & b_{i2} & b_{i3} & 0 \\
b_{i1} & b_{i2} & b_{i3} & 0 \\
0 & b_{i3} & -r_{i2} & r_{i1} \\
-b_{i3} & 0 & b_{i1} & b_{i2}
\end{bmatrix},
\]

\[
K_{bi}^3 = \begin{bmatrix}
b_{i1} & b_{i2} & b_{i3} & 0 \\
b_{i1} & b_{i2} & b_{i3} & 0 \\
0 & -b_{i3} & b_{i2} & -b_{i1} \\
b_{i2} & -b_{i1} & 0 & b_{i3}
\end{bmatrix}.
\]
References


REFERENCES


121
REFERENCES


REFERENCES


REFERENCES


