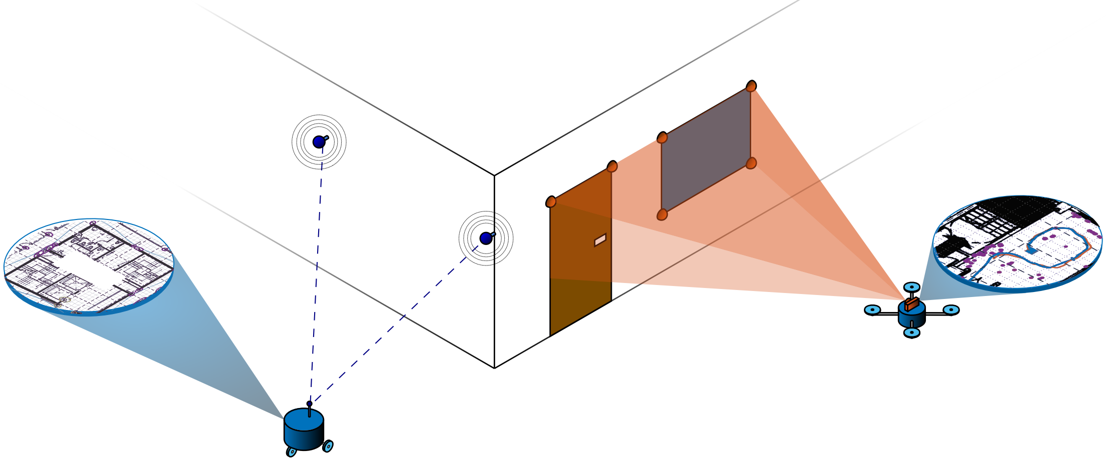
UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR TÉCNICO



Globally Convergent Simultaneous Localization and Mapping: Design Techniques, Analysis, and Implementation

Pedro António Duarte Marques Lourenço

Supervisors: Doctor Paulo Jorge Coelho Ramalho Oliveira
Doctor Pedro Tiago Martins Batista
Doctor Carlos Jorge Ferreira Silvestre

Thesis approved in public session to obtain the PhD Degree in Electrical and Computer Engineering

Jury final classification: Pass with Distinction and Honour

2019

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Globally Convergent Simultaneous Localization and Mapping: Design Techniques, Analysis, and Implementation

Pedro António Duarte Marques Lourenço

Supervisors: Doctor Paulo Jorge Coelho Ramalho Oliveira
Doctor Pedro Tiago Martins Batista
Doctor Carlos Jorge Ferreira Silvestre

Thesis approved in public session to obtain the PhD Degree in Electrical and Computer Engineering

Jury final classification: Pass with Distinction and Honour

Jury

Chairperson: Doctor José Alberto Rosado dos Santos Vitor, Instituto Superior Técnico, Universidade de Lisboa

Members of the committee:

Doctor Jorge Manuel Miranda Dias, College of Engineering, Khalifa University, United Arab Emirates

Doctor Maria Isabel Lobato de Faria Ribeiro, Instituto Superior Técnico, Universidade de Lisboa

Doctor António Paulo Gomes Mendes Moreira, Faculdade de Engenharia, Universidade do Porto

Doctor Carlos Baptista Cardeira, Instituto Superior Técnico, Universidade de Lisboa

Doctor Pedro Tiago Martins Batista, Instituto Superior Técnico, Universidade de Lisboa

Funding Institution:
Fundação para a Ciência e a Tecnologia

2019

“A mind is like a parachute.
It doesn’t work if it is not open.”
– Frank Zappa.

ABSTRACT

This dissertation addresses the problem of simultaneous localization and mapping (SLAM), with the main aim of devising strategies to achieve global convergence and stability guarantees, covering its main formulations in terms of the exteroception sensors used. It is argued that, since the pose is unobservable as all the information is relative to the vehicle, removing it from the filter and exploring the relative nature of the measurements is advantageous for that goal. Following that line of reasoning the SLAM problem is divided in two: (i) design an observer for the relative map with global convergence guarantees for its error dynamics; and (ii) find an complementary strategy that allows the computation of the pose from the sensor-based map and the initial pose. Regarding the first part of the problem, this dissertation presents a series of algorithms deeply rooted in a sensor-based approach to the SLAM problem that provide global convergence guarantees. The presented algorithms address the more usual range-and-bearing SLAM problem, in three dimensions (3-D) using an RGB-D camera, as well as the range-only (with an application to sensor networks) and bearing-only SLAM problems, using, respectively, radio/acoustic transceivers and a monocular camera. For each of these formulations a nonlinear system is designed, for which state and output transformations are considered together with augmented dynamics, in such a way that the underlying system structure can be regarded as linear time-varying for observability analysis and filter design purposes. This naturally allows for the design of Kalman filters with globally exponentially stable error dynamics, for which several simulated and experimental trials with aerial and ground vehicles are presented to demonstrate the validity of the approach and highlight the performance and consistency of the obtained filters. Finally, for the second part of the approach, an algorithm that builds on the closed form solution for an optimization problem equivalent to the orthogonal Procrustes problem is proposed. Due to the uncertainty present in the sensor-based estimation, the Procrustes problem is also studied thoroughly, and a novel uncertainty characterization of its results is presented and validated through extensive Monte Carlo simulations. Simulation and experimental results that complement those of the first part are included to illustrate the performance of the complete algorithm comprising the Earth-fixed Trajectory and Map estimation in cascade with each of the sensor-based SLAM filters under realistic conditions.

Keywords: simultaneous localization and mapping; sensor-based approach; observability; global exponential stability; Kalman filtering.

RESUMO

Esta dissertação aborda o problema da localização e mapeamento simultâneos (SLAM), com o objetivo de desenvolver algoritmos para os quais existam garantias globais de convergência e estabilidade, e que cubram as suas principais formulações em termos dos sensores de perceção externa utilizados. Nesse sentido, argumenta-se que é vantajoso explorar a natureza relativa das medidas e remover do filtro a posição e atitude do veículo, dado que estas não são observáveis por toda a informação obtida ser relativa ao veículo. Esta abordagem pressupõe a divisão do problema de SLAM em dois: (i) projetar um observador para o mapa relativo com garantias globais de convergência para a sua dinâmica de erro; e (ii) encontrar uma estratégia complementar que permita a determinação da posição e atitude do veículo partindo do mapa relativo baseado nos sensores e da posição e atitude iniciais. No que toca à primeira parte do problema, esta dissertação apresenta uma série de algoritmos firmemente enraizados na abordagem baseada nos sensores para o problema de SLAM que fornecem garantias globais de convergência. Os algoritmos apresentados abordam várias formulações do problema, desde o mais usual SLAM com distâncias e direções (conhecido como *range-and-bearing* SLAM) em três dimensões (3-D) passando por SLAM utilizando apenas distâncias (com uma aplicação a redes de sensores) e por SLAM utilizando apenas direções, através de, respetivamente, uma câmara de profundidade RGB-D, transdutores radio/acústicos e uma câmara monocular. Em cada uma destas formulações é projetado um sistema não linear, para o qual são consideradas transformações de estado e saída em conjunto com dinâmicas aumentadas. Esta metodologia permite que a estrutura dos sistemas resultantes seja considerada linear e variante no tempo no contexto de uma análise de observabilidade construtiva e do projeto de filtros. O passo seguinte é naturalmente a síntese de filtros de Kalman com dinâmicas de erro globalmente exponencialmente estáveis, para os quais várias simulações e experiências com veículos aéreos e terrestres são apresentadas para demonstrar a validade da abordagem proposta e enfatizar o desempenho e consistência dos filtros obtidos. Finalmente, para a segunda parte desta abordagem, propõe-se um algoritmo baseado num problema de optimização equivalente ao problema ortogonal de Procrustes, que tem uma solução em forma fechada. Devido à incerteza presente nas estimativas baseadas em sensores fornecidas por cada um dos filtros sintetizados na primeira parte, o problema de Procrustes é aprofundado e uma nova caracterização da sua incerteza é apresentada e validada através de simulações de Monte Carlo. Com o objetivo de ilustrar a performance do algoritmo completo constituído pela estimação da trajetória e do mapa num referencial fixo na Terra em cascata com cada um dos filtros para SLAM baseados em sensores, são apresentados ainda resultados de simulação e experimentais que complementam os da primeira parte da dissertação.

Palavras-chave: localização e mapeamento simultâneos; abordagem baseada em sensores; observabilidade; estabilidade global exponencial; filtros de Kalman.

To my family.

ACKNOWLEDGEMENTS

My first words of utmost gratitude go to my scientific advisors, Professor Paulo Oliveira, Professor Pedro Batista, and Professor Carlos Silvestre. Since they encouraged me to start this journey, their insightful guidance, invaluable support, and thoughtful motivation were fundamental to the path that lead to this dissertation. I am very grateful for their suggestion of my research topic and for the fruitful discussions we had that helped me steer through the challenges in my path. Their everlasting availability of which I took advantage many times throughout these years is something that I appreciate very much. I would also like to thank Professor Carlos Silvestre for the opportunity to shortly visit the SCORE Lab at the Faculty of Science and Technology of the University of Macau, which provided me with high-end equipment to perform most of the experimental part of this work.

I would also like to express my gratitude to Professor António Moreira and Professor Jorge Dias for their availability to be in the jury of my thesis along with Professor Carlos Cardeira and Professor Isabel Ribeiro to whom I am indebted for being members of the jury and of my steering committee. Their feedback was invaluable in the development of this dissertation, and greatly contributed to its improvement.

I owe a profound thanks to Bruno Guerreiro for his collaboration in some of the work presented in this dissertation. Aside from kick-starting this approach, for which I am grateful, his help, support, and feedback were instrumental to my path.

I would like to thank everyone involved in the development of the software and hardware used in the data acquisition for some of the chapters, namely André Oliveira and Bruno Cardeira of the DSOR Lab and David Cabecinhas of the SCORE Lab.

During my years at the DSOR group, I was blessed to witness what is arguably a one of a kind working environment filled with kind, generous, friendly people. A special thanks goes to all the friends that shared room 8.11 with me, João Almeida, Sérgio Brás, David Cabecinhas, Pedro Casau, Tiago Gaspar, Daniel Silvestre, Daniel Viegas, and Carlos Neves who was adopted by the guys in the room and with whom I shared some of my course work. I will always cherish the uncountable good moments we had in and outside ISR, they certainly contributed to some of the best memories in these years. I would also like to extend my thanks to the rest of the DSOR group, in particular to Bruno Gomes and Manuel Rufino for our daily friendly diatribes, and to Pedro Batista, Bruno Cardeira, Rita Cunha, and Bruno Guerreiro for letting me join their enlightened discussions over lunch and for contributing to the amazing work atmosphere of the DSOR lab.

I could not have reached this point without the support and friendship of my long-time friends from Técnico and in Setúbal. For the countless laughs, and for all the experiences that we have been through together over the years: my heartfelt thank you.

I owe my deepest gratitude to my family, to which I am devoted, and most importantly, to my parents for their unconditional support and love. At the end of the day, they endured my ups and downs, and are an indelible part of who I am.

Financial support: This work was partially supported by the Fundação para a Ciência e a Tecnolologia (FCT) through ISR under LARSyS UID/EEA/50009/2013 contract. The author was partially supported by the PhD. student grant SFRH/BD/89337/2012 from the FCT.

CONTENTS

Abstract **vii**

Resumo **ix**

Acknowledgements **xiii**

Contents **xv**

List of Figures **xxi**

List of Tables **xxvii**

Acronyms **xxix**

1 Introduction **1**
- 1.1 Motivation 2
- 1.2 What is SLAM? 3
  - 1.2.1 Different flavours 3
  - 1.2.2 Associated concepts 7
- 1.3 Theoretical challenges 8
  - 1.3.1 Consistency of SLAM algorithms 8
  - 1.3.2 Convergence of SLAM algorithms 9
- 1.4 The sensor-based approach 10
- 1.5 Problem statement 11
- 1.6 Proposed solution 12
  - 1.6.1 Part I: Sensor-based SLAM 12
  - 1.6.2 Part II: Earth-fixed Trajectory and Map 13
  - 1.6.3 Applications 14
- 1.7 Summary of contributions and publications 15
- 1.8 Outline 16
- 1.9 Notation 19

2 Designing a sensor-based SLAM algorithm **21**
- 2.1 The sensor-based framework 22
  - 2.1.1 Coordinate frames 22
  - 2.1.2 Nonlinear System Design 23
- 2.2 Observability 26
  - 2.2.1 Preliminary definitions and results 26

2.2.2 The analysis 27
2.3 Filter design 29
2.3.1 Convergence Analysis 29
2.3.2 Discrete implementation 31
2.4 Earth-fixed Trajectory and Map 33

Part I. Sensor-based SLAM 37

3 Range-and-bearing SLAM with RGB-D Vision 39
3.1 Introduction 39
3.2 Problem formulation 40
3.2.1 Nonlinear system dynamics 41
3.2.2 Problem statement 42
3.3 Observability analysis 42
3.4 SLAM filter design 55
3.4.1 Update step 55
3.4.2 State maintenance 57
3.5 Experimental results 58
3.5.1 Run #1 59
3.5.2 Run #2 62
3.6 Conclusions 64

4 Range-only SLAM 67
4.1 Introduction 67
4.2 Problem statement and System dynamics 68
4.2.1 Problem statement 68
4.2.2 Augmented system dynamics 69
4.3 Observability analysis 71
4.4 Filter design and implementation 83
4.5 Simulation results 84
4.6 Experimental results 86
4.6.1 Setup 86
4.6.2 Results 87
4.7 Conclusions 90

5 Range-only SLAM in Sensor Networks 93
5.1 Introduction 93
5.2 SLAM in sensor networks 94
5.2.1 Problem statement 94

5.2.2 Filtering concept . . . 95
5.3 System design . . . 97
5.3.1 SLAM in Sensor Networks . . . 97
5.3.2 Tracking . . . 99
5.3.3 Observability analysis . . . 100
5.4 Filter design . . . 106
5.5 Simulation results . . . 110
5.6 Conclusions . . . 113

6 Bearing-only SLAM with monocular vision 115
6.1 Introduction . . . 115
6.2 The bearing-only SLAM problem . . . 116
6.2.1 The sensor-based approach . . . 116
6.2.2 Problem statement . . . 117
6.3 Proposed solution: GES BO-SLAM . . . 117
6.3.1 State augmentation and output transformation . . . 118
6.3.2 Observability analysis . . . 120
6.3.3 Filter design . . . 124
6.4 Algorithm implementation . . . 127
6.4.1 Obtaining bearing measurements from a camera . . . 128
6.4.2 Landmark association . . . 129
6.4.3 Map maintenance . . . 130
6.4.4 Complexity reduction . . . 131
6.5 Simulation results and discussion . . . 132
6.5.1 Setup . . . 132
6.5.2 Typical run results . . . 132
6.5.3 Different input noise results . . . 134
6.6 Experimental results . . . 135
6.6.1 Rawseeds dataset . . . 136
6.6.2 The algorithm at work . . . 138
6.7 Conclusions . . . 145

Part II. Earth-fixed Trajectory and Map 147

7 Uncertainty Characterization of the Orthogonal Procrustes Problem 149
7.1 Introduction . . . 149
7.2 Preliminary definitions . . . 152
7.3 Procrustes optimization problem and closed-form solution . . . 154
7.4 Uncertainty characterization . . . 155

7.4.1 Rotation uncertainty 156
7.4.2 Translation uncertainty 161
7.4.3 Cross translation-rotation uncertainty 162
7.5 Algorithm validation 162
7.6 Conclusions 174

8 Earth-fixed Trajectory and Map online estimation **175**
8.1 Introduction 175
8.2 Overview 176
8.3 Uncertainty characterization 178
8.3.1 Earth-fixed map uncertainty 179
8.4 Simulation results: Range-only simultaneous localization and mapping (SLAM) 180
8.5 Experimental results 185
8.5.1 Range-and-bearing SLAM 185
8.5.2 Bearing-only SLAM 188
8.6 Conclusions 191

9 Conclusions and Future Work **195**
9.1 Conclusions 195
9.2 Directions for Future Work 199

A Supplementary Definitions and Results **201**
A.1 Results and definitions necessary for observability analysis 201
A.2 Lemmas 201

B Proofs of minor results **203**
B.1 Proof of Lemma 6.5 203
B.2 Proof of Lemma 7.1, Properties of the anti-commutation matrix 204
B.3 The error/perturbation model for orthogonal matrices 207

C Simulated Environment **211**
C.1 The map 211
C.2 Motion model 212
C.3 Trajectory design 212

D Experimental Setup: SCORE Lab **217**
D.1 *VICON* Motion Capture System 218
D.2 The instrumented *AscTec Pelican* 218

Bibliography **221**

List of Publications 233

LIST OF FIGURES

2.1 Schematic of the reference frames $\{B\}$ and $\{E\}$, and the landmark-based environment. 23

2.2 For RB-SLAM (Chapter 3), sensors measure the position of a landmark relative to the vehicle, for RO-SLAM (Chapters 4 and 5), the distance to a landmark, and for BO-SLAM (Chapter 6), the relative direction to a landmark. 25

2.3 Schematic description of the process of designing a globally convergent sensor-based SLAM filter. 28

2.4 The generic SLAM algorithm proposed in this dissertation, with references to the relevant chapters, possible exteroception sensors (red-green-blue-depth (RGB-D) camera, LIDAR, radio/acoustic beacons, and red-green-blue (RGB) camera), and possible proprioception sensors (optical flow, wheel encoders, and IMU). 33

2.5 Summary of the work performed with the objective of designing the Earth-fixed Trajectory and Map estimation algorithm which is the focus of Part II of this dissertation, showing the focus on the uncertainty of the Procrustes problem in Chapter 7 and the complete algorithm described in Chapter 8. 34

3.1 In range-and-bearing SLAM, the vehicle is equipped with sensors that are capable of measuring the position of a landmark relative to the vehicle. For example, a LIDAR in 2-D or an RGB-D camera for 3-D. In any of these cases, a feature detection/extraction tool has to be ran on the 2-D scans or the RGB images to obtain usable landmarks. 40

3.2 Geometrical interpretation of the observability conditions for range-and-bearing SLAM in 3-D. Three landmarks form a plane in one observation (left), two observations of two landmarks form a plane (center) or three observations of one landmark form a plane (right). 47

3.3 The experimental setup. 59

3.4 Run #1 – Landmark map of the environment with the estimated and real trajectory. 59

3.5 Run #1 – Time evolution of the estimation error of the Earth-fixed estimates. . . 60

3.6 Run #1 – Time evolution of the sensor-based estimates with 2σ bounds. 61

3.7 Run #1 – Time evolution of landmark-related variables. 61

3.8 Run #2 – RGB pictures (left) with detected features and coloured pointclouds (right) obtained by the *Kinect* at the beginning of the experiment (top) and while in the corridor (bottom). The size of the circles indicates the magnitude of the Hessian of each feature and the direction of the vertical radius represents the Laplacian (positive-up, negative-down). Green circles represent features extracted but not matched in the point-cloud, blue circles are detected landmarks, and red circles are used as measurements. 63

3.9 Run #2 – Top view of the Earth-fixed map with vehicle trajectory and ground truth against the floor blueprint. 64

3.10 Run #2 – Time evolution of the attitude and body-fixed linear velocity estimates against the ground truth. 64

4.1 In range-only SLAM, the vehicle is equipped with sensors that are capable of measuring the distance to a landmark. For example, acoustic beacons such as the Cricket system. [PCB00] 68

4.2 Trilateration for positioning a landmark in 2-D. 76

4.3 Picture of the estimated map rotated and translated using the true transformation at $t = 300$ s. The real trajectory is the blue line, the red ellipsoids represent the 3σ uncertainty of the currently visible landmark estimates, the purple ellipsoids represent the uncertainty of the non-visible landmarks, and the green ellipsoid is the landmark selected for closer inspection on Figure 4.5. 84

4.4 Evolution of the estimation of 5 landmarks in time. 84

4.5 The estimation error of a single landmark for all coordinates with 3σ uncertainty bounds. The yellow bars indicate the moments when the landmark was observed. 85

4.6 The flow of information in the proposed algorithm. The process employed to obtain linear velocity measurements is also explained. 86

4.7 The experimental setup. The Cricket beacons can be seen on the left spread around the room. The Cricket receiver is mounted on the quadrotor on the right. 87

4.8 The sensor-based landmark #1 estimate against ground truth with 3σ uncertainty bounds and observation instants. 88

4.9 The sensor-based landmark #7 estimate against ground truth with 3σ uncertainty bounds and observation instants. 89

4.10 The position of the vehicle in time. 89

4.11 Top view of the estimated sensor-based map at $t = 175$ s. 90

5.1 Schematic of a possible network scenario. Circles indicate static nodes, the wheeled circle is the vehicle and the hexagon represents the target. The movement of the vehicle allows for the connection of isolated nodes or groups of nodes, which otherwise would not communicate with the rest of the network. . 95

5.2 The concepts behind the algorithm proposed in this chapter, showcasing the strategy used for design, as well as the relation with Chapter 4. 96

5.3 Evolution of the estimation of 5 beacons in time. 110

5.4 On top, a picture of the estimated map rotated and translated using the true transformation, with 3σ ellipsoids, and the path of the target, both real (blue) and estimated (red). Non-visible beacons in green; Visible beacons in purple; Vehicle path in yellow. Below, the estimation error and 3σ uncertainty bounds of the target position and velocity when it follows a straight line. Solid lines indicate the estimation error for each coordinate and dashed lines the uncertainty bounds. The vertical yellow bars indicate the time intervals where the vehicle is directly observing the target. 111

5.5 The estimation error and 3σ uncertainty bounds of the target position and velocity when it follows a random trajectory. The vertical yellow bars indicate the time intervals where the vehicle is directly observing the target. 112

6.1 In bearing-only SLAM, the vehicle is equipped with sensors that are capable of measuring the relative direction to a landmark. For example, a single camera. . 117

6.2 Triangulation for positioning a landmark in 2-D. 122

6.3 A schematic of the construction of the initial covariance. R_m and R_M are the minimum and maximum range of the camera, α is the noise error, and σ_i^2 are the eigenvalues of the new covariance. 130

6.4 Top view of the estimated map in a typical simulation. The ellipsoids represent the 3σ bounds of the current estimated map. The lines inside the ellipsoids represent the distance from the current estimate to the true position of each landmark. 133

6.5 Results of a typical simulation. 134

6.6 Sensitivity of the estimation error with respect to noise in the linear velocity and angular velocity measurements. The red horizontal lines represent the median, the blue boxes delimits the 25th and 75th percentiles, the black whiskers are extended at 2.7σ, and the red crosses are all the individual occurrences beyond that bound. On the left, angular velocity noise standard deviation is fixed at $0.15°/s$, and on the right the linear velocity noise standard deviation is fixed at 0.01 m/s. 135

6.7 Inertial measurement unit and odometry-based angular velocities error evolution. 136

6.8 Information on the dataset. Schematic of the area and the trajectory, the robot used in the experiments and examples of places visited by the robot. 137

6.9 Results of dataset benchmark using Inverse Depth EKF and 1-point RANSAC. Monocular SLAM estimation from the combination of monocular camera plus wheel odometry (thin red) and ground-truth trajectory (thick blue). Reproduced from [CGM09]. 138

6.10 The noise in angular velocity measurements for the experiments in Section 3.5 (left) and in Section 4.6 (right). 139

6.11 The estimated map at the end of the experiment, along with the dead-reckoned path and extended ground-truth. 139

6.12 The estimated map throughout the whole of the experiment, with the spatial evolution of each landmark shown in intervals of 20 seconds. 140

6.13 Algorithm performance statistics. 141

6.14 At the top: loop-closed landmarks in the first loop closing peaks of Figure 6.13(a), with loop-closed landmarks highlighted in red. The middle and bottom pictures are frames of the camera stream with detected (red), selected (green) and associated (blue) features. The bottom one contains the most loop-closed landmarks of the corresponding time slot and the top one shows the one where more loop-closed landmarks where first observed. 142

6.15 Maps corresponding to loop closing peaks in Figure 6.13(a), with loop-closed landmarks highlighted in red. 143

7.1 The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and nominal input covariances. 166

7.2 The cumulative distribution of the likelihood ratios for the 3-D case including all the covariances built with each particular simulation values and nominal input covariances. 167

7.3 The cumulative distribution of the likelihood ratios for the 7-D case including all the covariances built with each particular simulation values and nominal input covariances. 168

7.4 The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 2 input covariances. 169

7.5 The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 3 input covariances. 170

7.6 The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 4 input covariances. 171

7.7 Translation error and uncertainty for a series of combinations of baseline-number of points. Translation error Monte Carlo samples in red, 99% bound in dashed black, 3σ simulation covariance ellipse in dashed green, and the 3σ covariance ellipse resulting from the uncertainty characterization with the median likelihood ratio. The baseline ($\Delta = \max\|\mathbf{b}_i - \mathbf{b}_j\|$ for all $i \neq j,\ i,j \in \mathcal{S}$) increases in each figure from left to right, whereas the number of points used in the optimization increases from top to down. Note that the scaling changes from line to line as the number of points increases (ellipses in the bottom are smaller than those at the top). 172

7.8 Translation error and uncertainty for a series of combinations of (rotation, translation, covariance shape). Translation error Monte Carlo samples in red, 99% bound in dashed black, 3σ simulation covariance ellipse in dashed green, and the 3σ covariance ellipse resulting from the uncertainty characterization with the median likelihood ratio. The first four figures have narrow covariances, and the last ones have rounder covariances. The translation and rotation combinations are, in each group, respectively, $\{[0,0]^T, \mathbf{I}\}$, $\{[2,2]^T, \mathbf{I}\}$, $\{[0,0]^T, \mathbf{R}(180°)\}$, and$\{[2,2]^T, \mathbf{R}(180°)\}$. 173

8.1 RO-SLAM: Evolution of the vehicle position estimation error with uncertainty bounds. 181

8.2 RO-SLAM: Evolution of the vehicle attitude estimation error with uncertainty bounds. 181

8.3 RO-SLAM: One particular landmark position error and uncertainty, with two different runs of the algorithm. 182

8.4 RO-SLAM: The spatial evolution of the map covariances. On the left, overlapped snapshots of the Earth-fixed map taken every second. On the center, details of the evolution of the input sensor-based covariance for one of the landmarks in the map, taken every 100 seconds and rotated to the Earth-fixed frame. On the right, the corresponding output Earth-fixed covariance every second, respectively. In this figure, the darker the colour of the ellipse the later in the run it appears. 182

8.5 RO-SLAM: Earth-fixed map and trajectory with uncertainty bounds for both landmarks and equally time-spaced poses. Top-view of the map with the floor blueprint in the background. 184

8.6 Information on the experimental setup of each trial described in this section. The top column depicts RB-SLAM and the bottom one depicts BO-SLAM. On the left, the data acquisition platforms for each algorithm. On the bottom right, the locations of the acquisition of each picture/LiDAR scan. 186
8.7 RB-SLAM: Run #1 – Evolution of the position estimates with ground truth and uncertainty bounds. Horizontal trajectory (top two figures), Vertical trajectory (bottom). 187
8.8 RB-SLAM: Run #1 – Evolution of the estimation error of the Euler angles with uncertainty bounds. From top to bottom: roll, pitch, and yaw. 187
8.9 RB-SLAM: Run #1 – Top-view of the Earth-fixed map and trajectory with ground truth and uncertainty bounds. 188
8.10 RB-SLAM: Run #2 – Top-view of the Earth-fixed map and trajectory with ground truth, uncertainty bounds, and floor blueprint. 188
8.11 RB-SLAM: Run #2 – Evolution of the position estimates with ground truth (available during the first 60 seconds) and uncertainty bounds. Horizontal trajectory (top two figures), Vertical trajectory (bottom). 189
8.12 RB-SLAM: Run #2 – Evolution of the Euler angle estimates with uncertainty bounds against the ground truth (available during the first 60 seconds). From top to bottom: roll, pitch, and yaw. 189
8.13 BO-SLAM: Evolution of the estimation error of the vehicle pose. On top, the two horizontal coordinates of the position estimation error, and on the bottom, the attitude estimation error in the form of the yaw error. 190
8.14 BO-SLAM: The two-dimensional Earth-fixed trajectory and map, with ground truth, executive drawings, and uncertainty bounds. 191

C.1 The top-view of the Earth-fixed map. 211
C.2 Bidimensional path followed by the vehicle during simulation. 213

D.1 The SCORE Lab. 217
D.2 The *VICON* system of the SCORE Lab. 218
D.3 The *AscTec Pelican* equipped with a *Microsoft Kinect*, a *Microstrain 3DM-GX3-25* and *VICON* markers. 219
D.4 A diagram of the technologies in the *Microsoft Kinect*. Reproduced from a slide in Microsoft's E3 Conference. 219
D.5 A *Crossbow Cricket* receiver. Reproduced from `http://cricket.csail.mit.edu/`. 220

LIST OF TABLES

1.1 Scenarios used within the thesis. The column on the left indicates the measurements available. 11

2.1 Summary of the information of the dynamic systems that underlie each SLAM formulation. The last row summarizes the state augmentation and output transformation details used to obtain LTV-like systems necessary for the following sections. 25

3.1 RB-SLAM filter parameters. 58

4.1 RO-SLAM Kalman filter parameters 84

7.1 Performance of the algorithm in terms of the distance to the true values. 164
7.2 Worst-case covariance likelihood ratio tests for $n = 2$. 165
7.3 Worst-case covariance likelihood ratio tests for $n = 3$. 165
7.4 Worst-case covariance likelihood ratio tests for $n = 7$. 165

8.1 Error norms for the range-only simultaneous localization and mapping (RO-SLAM) simulation results (mean±standard deviation) 181
8.2 Error norms for the experimental results (mean±standard deviation) 185

C.1 Boundary conditions for motion in corners. 214
C.2 Boundary conditions for motion in the take-off phase. 214

ACRONYMS

n-D

n-dimensional. 16, 151–153, 159, 174, 198

2-D

two-dimensional. xix–xxiv, 5, 11, 40, 55, 60, 76, 91, 111, 122, 128, 152, 163, 166, 168–171, 186, 188, 198

3-D

three-dimensional. v, vii, xix, xxii, xxiv, 6, 10–13, 15, 17, 21, 22, 39, 40, 47, 55, 62, 67, 91, 93–95, 115, 116, 129, 130, 133, 134, 144, 145, 150, 152, 163, 166–168, 184–189, 195, 196, 200

7-D

seven-dimensional. xxii, 162, 163, 166, 168

AHRS

attitude and heading reference system. 13, 180

B2B

beacon-to-beacon. 25, 95, 98, 100, 102

B2T

beacon-to-target. 25, 95, 99, 102, 104, 106, 109

BA

bundle adjustment. 7, 8

BO-SLAM

bearing-only simultaneous localization and mapping. xix, xxiv, 4, 6–8, 12, 14, 15, 21, 25, 29, 32, 115–117, 119, 120, 122, 124, 125, 131, 132, 135, 136, 141, 144, 145, 176, 185, 186, 190, 191, 196, 199, 204

CCDA

combined constrained data association. 7

DSOR

Dynamical Systems and Ocean Robotics. 220

EKF

extended Kalman filter. 3, 5–10, 39, 58

ETM

Earth-fixed trajectory and map. 13, 14, 18, 35, 58, 65, 176–178, 180, 181, 183–185, 189, 191

FOV

field-of-view. 7, 57, 61, 62, 85, 97, 100, 129, 138

GES

globally exponentially stable. 13, 15–17, 39, 40, 49, 67, 79, 93, 94, 106, 115, 124, 175, 178, 197

GPS

global positioning system. 2, 3, 200

IMU

inertial measurement unit. xix, 2, 4, 11, 12, 17, 33, 55, 86, 116, 136, 138, 196, 197

IST

Instituto Superior Técnico. 132, 181, 211

JCBB

joint compatibility branch and bound. 7

KF

Kalman filter. 11

LBL

long baseline. 94, 96, 197

LIDAR

light detection and ranging system. xix, 2–4, 33, 40, 200

LTV

linear time-varying. xxv, 11, 13, 16, 21, 25, 27–29, 31, 32, 42, 43, 47–50, 54, 65, 72, 74–79, 81, 83, 97, 101, 102, 104, 106, 110, 116, 119–125, 127, 145, 196–198

NIS

normalized innovation square. 56, 61, 62, 144

NN

nearest neighbor. 7

O

orthogonal group (see Section 1.9). 19, 157, 208, 209

ORB

oriented FAST and rotated BRIEF. 4, 6

RB-SLAM

range-and-bearing simultaneous localization and mapping. xix, xxiv, xxv, 4, 12, 14, 15, 21, 25, 29, 32, 39–43, 47, 49, 50, 55, 57, 58, 115, 176, 185–189, 199, 217

RF

radio frequency. 86, 220

RGB

red-green-blue. xix, xx, 12, 33, 39, 40, 55, 59, 63, 86

RGB-D

red-green-blue-depth. v, vii, xix, 3, 4, 11, 14, 17, 33, 39, 40, 49, 55, 62, 63, 196

RO-SLAM

range-only simultaneous localization and mapping. xix, xxiii, xxv, 4, 5, 8, 14, 15, 21, 25, 29, 32, 67, 69, 71, 72, 76, 78, 79, 83–86, 88, 93, 94, 96, 110, 112, 113, 117, 127, 176, 181, 182, 184, 196–198, 217

ROS

Robot Operating System. 220

SCORE

Sensor-based Cooperative Robotics Research. 18, 58, 86, 217

SFM

structure-from-motion. 7

SIFT

scale-invariant feature transform. 4

SLAM

simultaneous localization and mapping. v, vii, xvi, xix–xxi, xxv, 2–18, 21, 22, 24–29, 31–34, 39, 40, 42, 47, 55, 57, 58, 60, 64, 65, 68, 93, 95, 102, 115, 117, 131, 138, 141, 144, 149, 152, 175, 176, 178, 180, 181, 183, 185, 189–192, 195–200

SN-SLAM

simultaneous localization and mapping in sensor networks. 17, 101

SN-SLAMMOT

simultaneous localization and mapping in sensor networks with moving object tracking. 17, 94, 102, 104, 106

SO

special orthogonal group (see Definition 2.1). 19, 22, 23, 150, 152, 156, 157, 176, 209

so

special orthogonal Lie algebra (see Definition 7.1). 19, 23, 152, 153, 157, 208

SONAR

sound navigation and ranging system. 3, 4, 136

SURF

speeded-up robust features. 4, 6, 55–57, 62, 86, 128, 129

UAV

unmanned aerial vehicle. 1, 39

UCO

uniformly completely observable. 28

V2B

vehicle-to-beacon. 25, 95

V2T

vehicle-to-target. 25, 95, 99, 104

1

INTRODUCTION

Autonomous systems have been on the minds of researchers even before they were actually considered as such. While the term *robot* was only coined in 1920 by Karel Čapek in his play *Rossum's Universal Robots* [ČN04], as far back in time as Ancient Greece, the mathematician and engineer Heron of Alexandria described in his work *Pneumatica et Automata* [SoA99] over a hundred machines and automata, including machines powered by air, steam or water pressure, as well as mechanical or pneumatic means of performing wonders in Greek temples. The thrill of creation, which has always permeated Man's actions, led to the objective of recreating Nature's achievements: an example of an early attempt to do so is Leonardo da Vinci's humanoid robot [Ros06] designed in the XV century. However well intended accomplishments such as these were, it was the pressure of war that helped the idea of autonomous vehicles flourish. In the XIX century, Giovanni Luppis envisioned a floating unmanned device, remotely controlled through a wired link, for attacking and destroying ships - the first self-propelled torpedo [Mus10, Origin of the Whitehead torpedo, p. 6]. That century saw the appearance of several such weapons, and Nikola Tesla presented his own, with the novelty of it being wirelessly controlled by a radio link [Tes98].

The dawn of electronics and computers in the XX century paved the way to the thriving field of autonomous robotics that has been the subject of intensive research in the last decades, as autonomous vehicles pose the very interesting possibility of being able to perform tasks without the continuous guidance and supervision of humans. Search and rescue, surveillance, and the automatic inspection of critical infrastructures and buildings, such as bridges, electric power lines, dams, and construction sites, have been recently acknowledged as challenging and promising application scenarios for the use of, for instance, unmanned aerial vehicles (UAVs) (see [QSS⁺16, KBH08, GMG⁺08, CSK⁺17, MCMdD⁺06, KPL10, LBF⁺12, Ken12] and references therein). This calls for the capability to precisely determine the position and orientation of the vehicle, an ancient field of study known as navigation developed with the exploration of the seas by seafaring peoples. While in early periods sailors navigated without instruments (by staying in sight of land, for instance) and used nautical charts or sailing directions [Bow02], soon more intrepid expeditions started using rudimentary instruments such as the Phoenician sounding weight (which allowed the determination of the deepness of the sea and consequently the distance from land). This trend led to the advent of navigation with instruments, in particular for triangulation using directions.

1.1 Motivation

Navigation using directions to known sources has been in use for centuries. The observation of the sun and constellations of stars to infer position and heading are among the most ancient techniques for localization and navigation [Tay71], which are still presently in use for spacecraft navigation [XWL10]. Initially, in marine applications, several tools to measure the elevation of stars such as sextants and mariner's astrolabes were employed to derive the position of ships, and lighthouses were used in near-shore triangulation techniques. In the last century, aviation brought into use more advanced technologies supported on bearings (azimuth and/or elevation) readings: the automatic direction finder (ADF), the VHF omnidirectional range (VOR), and the instrument landing system (ILS) are the most common still in use today (see [Adm16, Chapter 16: Navigation] for an introduction to these technologies). After a great focus on bearings for navigation, range-based technologies have since become standard. These are divided in local methodologies, such as the distance measuring equipment (DME) used in aviation in conjunction with the VOR, and global, satellite-based, solutions. The appearance of global positioning systems (GPSs) has gradually replaced the use of the previous techniques, but in GPS-denied environments, aided relative algorithms are still necessary to navigate unmanned vehicles. Alternatives to these strategies of navigation include odometry-based mechanisms, the (mechanical) gyroscope, and, more recently, the inertial measurement unit (IMU) sensors, that ubiquitously equip any modern vehicle or smartphone. When the available sensors on board a vehicle are not sufficient to provide an accurate estimate of the pose of that vehicle, one solution is to use aided localization algorithms which make use of known characteristics of the environment such as maps or beacons typically employing acoustic or radio ranging, as in the case of GPS, light detection and ranging systems (LIDARs), or vision sensors. However, when navigating in an unknown environment, the mapping of that environment and the localization within that map have been shown to be dependent on each other, and a more intricate solution has to be considered: probabilistic simultaneous localization and mapping (SLAM), a concept introduced in the scientific community in the 1980's [SC86] and first coined in [LDW91].

1.2 What is SLAM?

Simultaneous localization and mapping (SLAM) is the problem of navigating a vehicle in an unknown environment, by building a map of the area and using this map to deduce its location, without the need for a priori knowledge of location. The solution to this problem is of great importance to the field of autonomous robots operating in GPS-denied environments, and therefore, since its introduction in the scientific community, SLAM has been the subject of many research efforts. These are connected to the vast diversity of research topics in SLAM, all stemming from the plethora of filtering techniques, mapping sensors, and mission profiles associated with it. There are several detailed surveys on most of the used techniques, sensors, and applications. See [DWB06] and [BDW06] for a thorough survey on the algorithms proposed in the first decades of SLAM research, [FPRARM15] for a specialized review of visual SLAM, [HD16] for a more up to date survey focused in the recent theoretical achievements and [CCC$^+$16] for an overarching survey on the history and remaining present and future challenges of SLAM, e.g., robustness and scalability, and so forth.

The research community has devoted significant effort to the study of this problem, with an initial focus on building probabilistically consistent maps. The works that established the statistical foundations for describing the relationships between landmarks and their correlations include [SC86], [DW88], [SSC90], and [LDW91]. Further research showed that a full and consistent solution to this problem would require all the vehicle pose and map variables to be jointly considered, which renders the problem intrinsically nonlinear due to the correlations between their estimates. From that initial discussion, and building upon these seminal works, a myriad of technical solutions have emerged from this challenge. These include the extended Kalman filter (EKF) SLAM [CUDW96], graph-based SLAM [LM97, TM06], the use of Rao-Blackwellized particle filters as in FastSLAM [MTKW02], or the use of information filters [TLK$^+$04]. Apart from varying in concept, the SLAM approaches can also be specially tailored to account for different sensors, depending on the mission environment. Sound navigation and ranging systems (SONARs) or other acoustic ranging systems, most common for underwater missions [BWPJ11], were used in ground vehicles in the early days of SLAM [TNNL02], whilst LIDAR, as in [MTKW02, CMCTN07, RH10], and cameras (monocular, stereo [SLL02], and red-green-blue-depth (RGB-D) cameras [EHE$^+$12]) are the main sensing devices for ground and aerial vehicles. These sensors involve obtaining range and/or bearing information of the environment, and usually demand the existence of a data association algorithm, due to the unknown correspondence between the reality and the estimated map.

1.2.1 Different flavours

The SLAM problem can also be characterized by the fundamental type of measurements available for filtering, usually referred to as landmarks. When the landmark measurements have a lower dimension than the considered mapping space (a single noise-free observation provides only a line or surface as an estimate for the relative position of the landmark), the resulting sub-problems are usually divided into range-only simultaneous localization and mapping (RO-SLAM) and bearing-only simultaneous localization and mapping (BO-SLAM), while the more usual SLAM problem is sometimes referred to as range-and-bearing simultaneous localization and mapping (RB-SLAM) to underline the case where all the relative coordinates of measured landmarks are readily available.

Range-and-bearing SLAM When the research on SLAM took off, the community mostly focused on the latter case, and both SONAR- and LIDAR-based algorithms were the norm. The applications of vision to SLAM have flourished in recent years, after much attention and effort was put into making the most of the potentialities of laser ranging sensors. Although both cameras and LIDARs have become more and more affordable, the latter are still considerably more expensive than the former. That fact, combined with very good results obtained in feature detection and image processing (for which the scale-invariant feature transform (SIFT) [Low99], speeded-up robust features (SURF) [BETG08], and oriented FAST and rotated BRIEF (ORB) [RRKB11] are an example), led to the increasing interest in vision-based SLAM. There are four main approaches in the literature differing in the number or type of cameras: mono [WAL+13], stereo [SLL02], time-of-flight [MDH+09], and RGB-D – a camera with a depth sensor [EHE+12]. All these algorithms utilize feature extraction and association to detect interest points in the collected images and associate them in two different frames (either from different cameras or at different time instants). Some use the great number of extractable features in two different time instants to compute the camera poses in a visual-inertial nonlinear optimization process, as [LFR+13], while others, such as [SLL02] and [DRMS07], employ a filtering framework to track landmarks. The combined use of cameras and sensors such as IMUs raises the need for proper calibration of camera parameters as well as rate-gyros and accelerometer bias. This topic is addressed in [JS11] and [KS11] as visual-inertial sensor fusion using extended or unscented Kalman filters. Since range-and-bearing measurements provide the most information in terms of external perception of all the formulations, its main challenges reside in how to scale the algorithms for long term mapping, data association, convergence, consistency, among other topics. Some of these will be addressed later in this text.

Range-only SLAM Although localization using distances to beacons is a very well known subject, the number of SLAM algorithms using only ranges is relatively small, especially when compared with the widespread use of algorithms working on range and bearing, or on bearings-only. On one hand, the range-only simultaneous localization and mapping (RO-SLAM) problem is not prone to association errors, as are other SLAM formulations (more on this later in this chapter). RO-SLAM bypasses this error source, owing to a fundamental aspect of the RO-SLAM problem, as the information carried by the ranging signals usually allows the unambiguous association of the measurement and the corresponding state at all times. On the other hand, one of the main challenges in RO-SLAM is the initialization of the algorithm, either due to the absence of global convergence results in EKF solutions such as [OLT06], or the computational burden of having a sufficiently representative prior belief, in particle filter solutions [BFMG08]. Most of the RO-SLAM solutions rely on some form of initializing procedure in order to create a new landmark in the state. These include trilateration with ranges from different instants to obtain a first estimate, usually through least squares, such as what was proposed in [AHWD11]. Also, due to the sparse information extracted from ranging, RO-SLAM algorithms are commonly designed for two-dimensional (2-D) environments, e.g., a ground robot and landmarks at the same height, see [MZZ$^+$09].

The common RO-SLAM formulation has similarities with the problem of Sensor Networks (SN), in the sense that there is an agent receiving signals from a network of sensors, and, therefore, the two ideas have been used in conjunction in works such as [DS12] and [TGDO14], where, along with agent-to-sensor ranges, sensor-to-sensor ranges are also used, thus approximating the RO-SLAM problem to the sensor network localization problem. In that direction, the authors of [TGdDO14] use the concept of multi-hop measurements in the SLAM context. In spite of their proximity, the fields of SLAM and self-localization of sensor networks have historically been treated separately and by different fields of research, robotics and signal processing. Due to that fact, very different approaches have been taken to produce algorithms for each of these topics, from the Bayesian approach of robotics to the convex optimization-based solutions spawned by signal processing. The proliferation of (wireless) sensor network applications, including environmental monitoring, intrusion detection, search and rescue, and several others, has led to a significant research interest on the development of localization techniques (see [MFA07] and references therein for an overview of these techniques). Due to the inherent distributed computation capabilities of a sensor network, several self-localization algorithms have been developed to explore those capabilities, of which [SHCJ10] is a good example as it provides an optimization algorithm that addresses both range-based and connectivity-based localization. Most algorithms rely on nonlinear optimization problems with convex relaxation techniques or on multidimensional scaling [AM10].

Another typical use of sensor networks is the tracking of moving objects. This is mostly known in the scientific community as the simultaneous localization and tracking problem [WTT$^+$07] and is closely related to SLAM. It is usually formulated in a Bayesian framework [TRB$^+$06], which brings it further closer to SLAM.

BO-SLAM The bearing-only case is even more challenging than the range-only one, because an observation of the former corresponds to an unbounded region. This raises serious issues on the initialization of a landmark, the main topic of research in partially-observable SLAM. This research effort yielded initially only delayed solutions, i.e., algorithms that try to obtain a preliminary landmark estimate from readings at different viewpoints before introducing the initial estimate into the filter. This can be done through triangulation or more advanced probabilistic approaches, using, for example, a sum of Gaussians [LLS05] or deferring the initialization until an approximately Gaussian estimate is achieved [Bai03]. There are some notable exceptions in [KD04] and [SMDL05] that recur to multiple hypothesis directly in the filter. More recently, algorithms with undelayed initialization, at least to some extent, have been proposed. The concept of inverse depth parametrization [CDM08] was inspired by computer vision and it brought to bearing-only EKF-SLAM several advantages, as the inverse depth has better linearity and allows for very low parallax features. However, it requires a six-dimensional (6-D) representation of the feature state, instead of the traditional three-dimensional (3-D) representation. In this case, feature depth, or rather, inverse depth, is initialized with a generic prior that accounts for statistically feasible depths. Another technique was proposed in [CPKL12], where the authors transform the measurement in image coordinates to a three-dimensional (3-D) vector in camera coordinates by assuming an arbitrary depth while defining a very large covariance for that measurement. Even though these two approaches brought a fresh perspective to the field and seem to have good results in practice, there are no guarantees of convergence.

Although research in BO-SLAM is not as prolific as it is in range-and-bearing SLAM, the community has provided several approaches depending on the underlying filtering technique and the sensors used. Most algorithms are based on EKFs. However, some methods are inspired on expectation-maximization or particle filters (see [BCK06] for a comparison of these approaches). Another source of diversity in BO-SLAM algorithms is the type of sensor used. Even though bearing-only localization is historically related to the computation of the angle-of-arrival of signals from beacons through the time difference of arrival at different elements of a receiving array, BO-SLAM is mostly associated with monocular vision [JKFB06] or even catadioptric omnidirectional systems [HMK07]. Monocular SLAM and other vision-based algorithms have been the focus of a intensive research effort, yielding several relevant algorithms, of which one of the most important is [DRMS07], as it was the first real-time SLAM algorithm with a single camera as the only data source (even though it was preceded in computer vision by [CFJS00]). This algorithm initializes features as semi-infinite lines lying in the direction of the detected features. Other, more recent, approaches vary in application, the way features are handled, and the underlying filtering engine. For example, [KE13] represents a departure from the usual exploitation of the ubiquity of cameras, as it presents possible new applications. In this case, underwater hull inspection. Other interesting approach is ORB-SLAM [MAMT15, MAMT17] that uses ORB features [RRKB11] which are rotation invariant, as SURF features [BETG08], and have faster extraction. The main algorithm performs feature tracking and on top has external loop closing and re-localization procedures. Demonstrating the possibilities of graph-based solutions, the authors of [DB15] introduce a closed-form pose-chain optimization algorithm that uses sparse graphs as well as appearance-based loop detection.

Another very important approach to monocular SLAM was imported from computer vision, where structure-from-motion (SFM) has been a subject of research for many years [Ull79]. In that field, there have been some notable advances that can be related to BO-SLAM, such as [CFJS00], where the authors propose an EKF to address SFM in real-time and perform a characterization of observability and minimal realization. Also in SFM, bundle adjustment (BA) techniques [TMHF00] have become the standard. In bearing-only SLAM this concept is much more recent, but there have been notable efforts as detailed in [SMD12]. There the authors perform a comparison between established bundle-adjustment and filtering techniques, concluding that in general BA has a better accuracy/computational cost ratio than EKF filtering. However, it is also mentioned that BA procedures have no general robustness to initialization parameters and that filters may be more capable of dealing with the high uncertainty present in that phase.

1.2.2 Associated concepts

Being one of the major sources of inconsistency, data association is an important component of most SLAM algorithms. It allows the association between the landmarks measurements and the state landmarks, and it is necessary to close a loop. Several strategies are widely used, such as the simplistic nearest neighbor (NN) validation gating (which tries to associate one landmark at a time), the joint compatibility branch and bound (JCBB) [NT01], and the combined constrained data association (CCDA) [Bai02] (that treat the association as a batch problem), while other, more evolved, strategies such as those in [BSC13] use sensors that provide unique characteristics of each detected feature. When these characteristics are available, the problem can be simplified, as is the case with landmarks obtained from images of the environment. The association of feature descriptors with high dimensionality can be done with good results using simpler sequential algorithms, such as the sequential compatibility nearest neighbour [Coo05], as was done in [SLL02] for instance. In range-and-bearing SLAM, the strategies described above mostly solve the problem. In BO-SLAM, however, there are extra difficulties in the initialization process. Some approaches try to deal with all hypothesis when initializing [TR06], but vision-based algorithms may use image information if the frame-rate is high enough to disambiguate measurements. Feature extraction provides tools with which to distinguish different features, that can be combined with well tested association methods, such as the sequential compatibility nearest neighbour or JCBB [NT01].

Data association assumes a decisive role in SLAM with the issue of loop closing [JIS+11], i.e., the ability to recognize previously visited terrain when it is once again within the field-of-view (FOV) of the vehicle. It is closely related to the association problem and with the inconsistency that some SLAM approaches suffer from, see [BNG+06]. Before the widespread use of vision in SLAM, loop closing was based on landmark position or map-to-map comparison, or scan matching when dealing with laser range finders. With the information present in images, other possibilities arose, such as image-to-image or image-to-map correspondence. These are discussed at length in [WCN+09].

1.3 Theoretical challenges

There have been several topics of intense research within the scope of simultaneous localization and mapping, such as computational efficiency, use of different sensing devices, diverse filtering and optimization techniques, optimality, consistency, and convergence of the algorithms. These last three are important aspects of any estimation problem, since together they provide guarantees to the user that the algorithms will perform in a certain way, while also giving a measure of the quality of the estimates in real time.

In subsection 1.2.1, several of the open issues of the various SLAM formulations were enumerated. One of the recurring challenges is the initialization of estimates in RO-SLAM and BO-SLAM. In fact, addressing that problem has been the focal point of a variety of strategies in those areas, and it was indicated that one of the current trends in BO-SLAM, BA-based algorithms, suffers from lack of robustness to initialization parameters. The two main aspects usually taken into account when initializing an estimate are twofold:

(i) a probabilistically consistent description of the initial knowledge is chosen, i.e., the probability distribution that describes the estimate contains all the possible estimates, while maintaining a *sensible* description of the actual estimates. For example, after one additively perturbed range measurement to a landmark the correct estimate would be a sphere with thickness depending on that perturbation, which is something quite difficult to model. In the bearing case, one noisy measurement provides an infinite cone, which is also problematic to model in, for instance, Gaussian-based estimators; and

(ii) either the overall problem is posed in such a way that the initial distribution fits the model for the estimate after the transient period, or the initialization is delayed until such a distribution has been achieved.

Looking at these two aspects, it is apparent that the initialization problem can be traced back to consistency and convergence considerations, which further shows the relevance of addressing these issues.

1.3.1 Consistency of SLAM algorithms

In general, the SLAM problem has a nonlinear nature which can be tackled using EKF-based solutions, as well as other similar filters that usually imply the linearization of the dynamics or measurement models, resulting in lack of consistency and convergence guarantees [JU01, HD07]. Consistency is an important aspect of any estimation scheme, since a consistent algorithm avoids optimistic estimates that may mislead the user, be it a human or software. For example, due to the nonlinear nature of the problem, the widely used extended Kalman filter has well-known problems of consistency and optimistic estimates (see [BNG+06] and [JU01] for an analysis of this issue) due to the use of linearized equations evaluated at the estimated states that can be erroneous and mislead the filter. Another major source of inconsistency in SLAM algorithms is the spurious association of the measured and state landmarks, accompanied with faux loop closings. In order to cope with this inherent characteristic of EKF-SLAM, some effort was put into action by the community, from using only the first estimates on the Jacobian [HMR08], to an increased focus on the observability of the problem [ACS04]. Following this line of thought, the authors of [HMR10] investigate the unobservable space of the linearized error-state system which they show is of smaller dimensionality than the actual unobservable space of the underlying nonlinear system, thus yielding erroneous information gains from a direction where none is available. They then argue that selecting the linearization points of the EKF to ensure that both unobservable spaces are of the same dimension improves the consistency of the filter, and propose a constrained optimization problem with a closed-form solution that minimizes the linearization error subject to conditions that ensure an unobservable subspace of appropriate dimensions. More recently, the (right) invariant EKF [BB15] was proposed as a strategy to improve the consistency of the algorithm, and an in-depth study of its convergence and consistency was presented in [ZWS+17]. Inconsistency was also dealt with earlier on in the history of SLAM by the use of a robocentric approach [CMCTN07]. This algorithm can improve the consistency of the regular EKF, yet, as it still considers the estimation of the (unobservable) incremental pose in the filter state, it cannot provide guarantees of convergence.

1.3.2 Convergence of SLAM algorithms

Apart from dealing with consistency problems, there are also some notable results regarding the convergence of EKF-SLAM (such as [HD07] and [BJ09]). These study the evolution of the covariance of the filter and derive conditions in which the covariances of the filter converge, as long as it is consistent. These studies that offer some guarantees of convergence include methods that usually assume that the linearized system matrices are evaluated at the ideal values of the state variables [DNDW+01, HD07], or other that resort to stochastic stability concepts assuming access to both linear and angular velocity measurements [BJ09]. From a different point of view, [JS11] and [KS11] address the observability properties of nonlinear SLAM, as it is a necessary (but not sufficient) condition for the convergence of any filtering algorithm. In [JS11] an algorithm with monocular vision and inertial measurements is proposed, aiming to perform SLAM and calibration of parameters such as the local gravity vector and the 6 degrees-of-freedom IMU-camera rigid transformation, whereas in [KS11] the authors also estimate the biases of the accelerometer and rate-gyro measurements in an algorithm more focused on calibration. They perform a thorough observability analysis based on differential geometry considerations resulting in conditions on the vehicle trajectory for local weak observability of the underlying nonlinear system (see [HK77] for further details). All these represent important advances in the study of convergence and observability of SLAM algorithms.

Despite the great research efforts devoted to the SLAM problem as previously documented, a formal theoretical result on global convergence for a three-dimensional SLAM algorithm is absent from the literature to the best of the author's knowledge. This is a significant omission, since guaranteeing convergence a priori, or, at the very least, knowing in what conditions a SLAM algorithm converges can solve or help solve a number of problems usually present in any SLAM implementation. For example:

(i) the undelayed initialization problem. As explained before, when global convergence is guaranteed, it is not necessary to have complex processes for initialization since the estimates will converge regardless of the initial guess; and

(ii) the kidnapped robot problem. The idea behind this problem is that of a catastrophic failure in the localization systems of a vehicle. Imagine a situation in which the vehicle localization has converged and the vehicle is moved to a new location through a process not captured by its proprioception sensors. Unless the localization algorithm has a way to detect and cope with these occurrences, the estimates will diverge and probably never recover. EKF-based algorithms have a hard time dealing with this problem [TBF05, Chapter 8], and particle filters use an heuristic approach that usually involves an adaptation of the resampling stage. If global convergence is present, then the instant immediately after the failure can be seen as a new (wrong) initial guess that will of course lead the algorithm to erroneous (and inconsistent) estimates during a transient period. However, the convergence properties ensure that during this period the estimates will be converging to their true values, as long as the vehicle traverses previously visited terrain and is able to recognize it.

Another relevant aspect is that convergence guarantees lead to an easier implementation when compared to most algorithms that may require significant empirical fine tuning in practice to ensure that the algorithm will operate as it should. All these point towards the importance of addressing convergence of SLAM algorithms as a research objective.

1.4 The sensor-based approach

The simultaneous localization and mapping problem can be formulated in several different ways, and many methodologies follow the idea of performing the estimation in an Earth-fixed frame. Hence, the usual SLAM approach requires a single filter to maintain estimates of the map and vehicle pose along with the respective covariances and cross-covariances. Nevertheless, the sensor-based approach has precedents in SLAM history, most importantly the robocentric idea in [CMCTN07], designed to tackle the consistency problems while still considering the estimation of the (unobservable) incremental pose in the filter state. Another related work is Linear SLAM [ZHD13], in which the map joining procedure is followed, while the state is transformed and augmented in order to achieve a linear least squares formulation. Furthermore, it is possible to use an alternate formulation that uses a Kalman filter (KF) that explores the linear time-varying (LTV) nature of the sensor-based SLAM system. In the light of this, the work in [GBSO13] proposed a different solution where the 2-D filter is purely sensor-based, suppressing pose representation in the state and therefore avoiding its pitfalls (singularities and nonlinearities) – an approach deeply rooted in the well-established theory of sensor-based control which explores the fact that vehicle-fixed sensors provide measurements that otherwise would need to be transformed to an Earth-fixed frame (see [WSN87], [KC11], and references therein for more on this subject). This means that error variables are derived directly in the sensor space and therefore use only body-fixed quantities. These can be provided directly by sensors, as in [VRB00], or by state observers such as sensor-based SLAM. This strategy countermands consistency problems and, at last, provides theoretical guarantees of global convergence for two-dimensional SLAM, paving the way for the contributions in this dissertation.

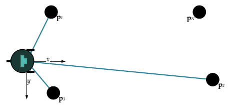
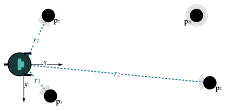
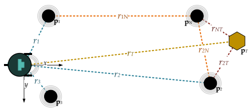
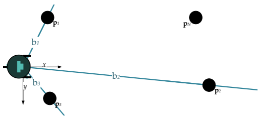
1.5 Problem statement

This dissertation proposes to address the SLAM problem in the formulations explained in Section 1.2.1 and summarized in Table 1.1, while tackling the theoretical challenges enumerated in Section 1.3 with special emphasis on convergence. In particular, the following scenarios are addressed:

(i) range-and-bearing SLAM where angular velocity measurements are the only available ego-motion measurements. The sensor suite considered contains a device able to extract the relative 3-D position of landmarks, such as an RGB-D camera (of which the *Microsoft Kinect* is a well-known example) coupled with a feature extraction algorithm, and an IMU. The linear velocity is estimated from landmark measurements which are naturally detected depending on the feature detection algorithm;

(ii) range-only SLAM with a sensor suite composed of a receiver to communicate with a constellation of beacons, an IMU, and either: (1) an odometer (in the case of ground robots),

Table 1.1: Scenarios used within the thesis. The column on the left indicates the measurements available.

	Range-bearing	Range-only	Sensor-networks	Bearing-only
				
$\mathbf{p}$	natural landmarks	beacons	sensor network / moving agent	natural landmarks
$\mathbf{v}$	×	(visual) odometry	✓	odometry
$\boldsymbol{\omega}$	IMU	IMU	IMU	IMU

or (2) a facing down RGB-D camera or a device for optical flow complemented with an altimeter (for aerial vehicles). The environment is comprised of a constellation of beacons that communicate with the vehicle allowing the determination of the distance between the vehicle and each beacon;

(iii) range-only SLAM where, aside from the communication with the vehicle, the beacons also communicate with each other, and there may be mobile nodes in sensor network formed by the beacons; and

(iv) bearing-only SLAM with a sensor suite containing a monocular camera, an IMU, and one of the linear velocity measuring devices exposed above. As in RB-SLAM, landmarks are naturally extracted from the environment, i.e., detected in the captured RGB images: corners, points-of-interest, and so forth.

1.6 Proposed solution

Following the novel idea introduced in Section 1.4, this dissertation presents a two-part strategy for solving the simultaneous localization and mapping problem in the various flavours detailed in Section 1.2.1 while achieving previously absent convergence guarantees. The resulting algorithms can be used in several scenarios and vehicles, with a direct application to aerial or ground robots.

1.6.1 Part I: Sensor-based SLAM

In the first part, a class of filters rooted in the sensor-based idea is presented along with a systematic approach to the nonlinearities intrinsic to the SLAM problem. The three-dimensional (3-D) sensor-based formulation avoids the representation of the pose of the vehicle in the state, as it becomes deterministic and available by construction, and allows the direct use of odometry-like information that is usually expressed in body-fixed coordinates. This leads to the establishment of global exponentially fast convergence results which facilitate the undelayed initialization of landmarks at any direction and depth.

The sensor-based approach to SLAM, or, as it is commonly known in the SLAM community, the robocentric approach, has been proven more consistent than its inertial, world-centric, counterpart [CMCTN07]. Furthermore, previous observability studies using piece-wise linearizations showed that, for the BO-SLAM case, this approach becomes fully observable in two time steps, in opposition to what happens in the world-centric case [VCBS$^+$07]. In addition, in this family of problems where the measurements are all expressed in local coordinates it makes sense to operate in a sensor-based framework, as it is a way of avoiding the inclusion of the pose of the vehicle in the filter state, one of the main sources of nonlinearity. That is the idea behind the nonlinear systems that underlie the filters designed in this thesis. To these systems is then applied a methodology involving state augmentations and output transformations influenced by the source-localization algorithms proposed in [BSO11b] and [BSO13b]. This leads to the design of linear time-varying (LTV) systems whose observability conditions are given in a constructive analysis with clear physical insight, from which Kalman filters with globally exponentially stable (GES) error dynamics follow naturally.

This novel strategy applied to 3-D SLAM addresses the two main challenges enumerated in Section 1.3 by establishing previously absent global convergence guarantees for this relevant problem, and, through the use of a robocentric framework, by reducing the inconsistency problems present in the typical inertial framework [HMR10]. It should be noted that the paramount idea in this approach is the fact that only the relative coordinates of the map with respect to the vehicle are estimated. In fact, it is possible to achieve similar results in a world-centric strategy, as long as this relative property is kept and external estimates for the attitude are available. An example of this, inspired in the authors sensor-based approach to SLAM [GBSO13, 6, 7] and the theoretical results associated, is given in [JB16].

Although the sensor-based SLAM estimates and uncertainties are sufficient for many applications, the Earth-fixed estimates and uncertainties may also be of interest. One possibility is to use at least two landmarks as anchors with known coordinates, and other is to acquire an external estimate of the attitude of the vehicle (e.g. with an attitude and heading reference system (AHRS) [BSO12]).

1.6.2 Part II: Earth-fixed Trajectory and Map

In the second part of this thesis , a dual algorithm is proposed to tackle this problem, i.e., to estimate the relevant Earth-fixed quantities, the vehicle pose and the landmark map, using only the body-fixed map provided by any of the filters designed in the first part. The Earth-fixed Trajectory and Map (ETM) estimation algorithm is based on an optimization problem to compute the pose of the vehicle from the comparison of the map in the moving and Earth-fixed frames, and on the update of the Earth-fixed map accordingly. Based on the uncertainty provided by the sensor-based filter, a fully characterized uncertainty approximation for this highly nonlinear problem is included.

The optimization problem is shown to be the same as the weighted orthogonal Procrustes problem, the well known problem of finding the rotation matrix and the translation vector that best map one set of points to another, which has a closed-form solution first proposed by [Sch66]. Due to the relevance of this problem to the designed algorithm, the weighted orthogonal Procrustes problem of matching stochastically perturbed point clouds is also studied within the thesis, where a characterization of the uncertainty of the obtained closed-form solution, which admits arbitrary transformations between point clouds and generic (cross) covariance matrices for the points of each cloud, is proposed and validated. Towards this end, perturbation theory is employed, considering arbitrary rotations and translations, individual weights, and individual covariance matrices for the landmarks of the inertial map, improving on previous successful approaches in medical-based applications of the Procrustes problem [Dor05]. These results also hold for n-dimensions, which means that they can be applied directly to existing 2-D and 3-D algorithms. The presented approach builds on the initial idea for the bidimensional RB-SLAM case proposed in [GBSO12] and further developed in [Gue13], while expanding it to a common n-dimensional framework and, coupled with the work relating specifically to the Procrustes problem, providing a more involved characterization with fewer approximations and a more thorough validation. In sum, the complete strategy encompasses (i) a series of sensor-based SLAM filters which estimate the landmark map and other vehicle related quantities expressed in the body-fixed frame; and (ii) an Earth-fixed Trajectory and Map (ETM) estimation algorithm which uses the sensor-based map estimate of the SLAM filter to provide fully uncertainty characterized estimates of the pose of the vehicle and the environment map.

1.6.3 Applications

The SLAM algorithm is extremely important to any mobile robotics problem, as the sheer magnitude of the number of works in the field shows. SLAM, in general, allows the deployment of a robot/vehicle/agent in a new environment without any *a priori* knowledge whatsoever. In that sense, SLAM has an easier overall implementation process than simple localization, which requires some form of prior map. In fact, SLAM algorithms can even be useful in providing a map for future localization usages. For example, in large industrial settings such as a warehouse [Gui08] where autonomous vehicles navigate daily, it may be complex and time-consuming to deploy and produce an exact map of a large number of marks for localization purposes. Using a SLAM algorithm while knowing only the position of a short number of landmarks (as small as two) to serve as anchors for the map is a more expedite way to achieve the same result.

As mentioned before, possible applications range from search and rescue to automatic inspection of critical infrastructures, to consumer-grade applications such as cleaning robots [PRSF00]. However, certain types of sensors (and consequently, formulations) may be more adequate to certain situations than others. For example, both RB-SLAM and BO-SLAM are ready for applications where there is no possibility to prepare the environment prior to the mission, such as search-and-rescue [QSS+16]. On the other hand, in outdoors settings where infrared-based RGB-D cameras such as the *Microsoft Kinect* fail, either RB-SLAM with more demanding stereo or trinocular camera systems or monocular BO-SLAM are the main possibilities. When the use of vision is out of the question, due to its computational cost or due to environmental constraints (nocturnal missions, missions in highly structured and symmetrical environments, missions on environments low on features, and so forth), if prior deployment of a constellation of beacons is possible, then RO-SLAM is the obvious choice. This is further true if there already exists a sensor-network covering the area. RO-SLAM is also interesting since it demands the least computational power of the various formulations, and ensures the most reliable estimation since there is no probabilistic data association connecting measurements and the map. In general, any SLAM algorithm that uses artificial landmarks that can be uniquely identified, be it radio/acoustic beacons or visual marks shares this advantage.

1.7 Summary of contributions and publications

This section briefly presents the main contributions of the thesis, while more detailed outlines are given in the subsequent sections. The results obtained in the scope of this thesis have naturally led to a number of publications, enumerated in the List of Publications that closes this manuscript.

This dissertation presents novel algorithms for simultaneous localization and mapping, reporting the design, analysis, implementation and validation stages. The general objective of the thesis is to find global convergence guarantees for this nonlinear problem. The strategy proposed is to divide the problem in two, noting that, since all the measurements associated with SLAM are relative, removing the unobservable pose from the main filter is beneficial. Following this, a class of systems rooted in sensor-based considerations is designed, and a strategy to convert sensor-based to Earth-fixed maps by computation of the vehicle pose is proposed, its statistical coherence validated and its experimental performance evaluated. In sum, this dissertation aims to provide a comprehensive view on how to achieve global convergence of metric SLAM algorithms in their various formulations.

Aside from the common thread uniting every chapter within their respective part, there is a clear fundamental idea behind the whole thesis: to provide complete SLAM algorithms that can serve as the basis for a wide variety of practical applications. However, the proposed strategy has a two-sided nature that naturally justifies the division of the thesis in two parts: (i) the first is dedicated to the design, analysis and practical evaluation through simulation and experimental results of a series of sensor-based filters; and (ii) the last addresses the design, characterization of the uncertainty, and experimental validation of an algorithm focused on moving from a sensor-based map to an Earth-fixed framework.

In particular, the contributions of the **first part** of the thesis are

(i) a systematic methodology to overcome the nonlinearities present in the design of SLAM dynamic systems, regardless of the type of sensors at hand, inspired in previous sensor-based algorithms such as [GBSO13], [BSO11b], [BSO14], and [BSO13b]. This methodology paves the way for

(ii) the analysis of the observability of the nonlinear problems at hand, which provides physical insight into the requirements on the motion of the vehicle, and that is constructive in the sense that it leads to

(iii) the design of globally exponentially stable sensor-based SLAM Kalman filters for RB-SLAM, RO-SLAM and its derivatives, and BO-SLAM in the scenarios presented in Table 1.1. This is, to the best of the author's knowledge, the first time that global convergence results are achieved for each of these formulations of SLAM in 3-D. The relevance of this cannot be understated, as, for example it intrinsically solves the initialization and kidnapping problems; and

(iv) the implementation in real world experiments of the three main formulations of sensor-based SLAM, providing a proof-of-concept and validation of the proposed approach, as well as allowing the evaluation of its performance.

The research on this topic was published in four conference papers [1, 3, 4, 5], three journal papers [6, 7, 10], and one book chapter [8].

The contributions of the **second part** of this dissertation are

(i) the uncertainty characterization of the closed-form solution for the weighted orthogonal Procrustes problem, i.e., the problem of obtaining the transformation between corresponding n-dimensional point sets. These sets are assumed to be perturbed by anisotropic noise that can be described by the first two moments, and the points are not required to be independent nor identically distributed. Perturbation theory is applied and expressions are found for the mean and (cross-)covariance of all uncertain quantities; and

(ii) an n-dimensional online algorithm to compute the Earth-fixed pose and map and their associated uncertainty given the initial pose and a sensor-based map throughout time, building on the GES SLAM filters and on the weighted orthogonal Procrustes problem.

The research on this topic was published in a conference paper [2] and a journal paper [9] addressing the uncertainty of the Procrustes problem. A journal paper [11] reporting the complete algorithm is also under review.

1.8 Outline

This thesis is organized in 9 chapters and 4 appendices, including this introductory chapter. The remaining are structured as follows.

Chapter 2 is devoted to explaining the sensor-based approach to SLAM that underlies all the work in the first part of the dissertation. It presents all the necessary definitions and results for the establishment of the theoretical work on later chapters, while providing a brief summary of how this theoretical work is laid out and what results from it. Finally, a broad overview of the implementation of a generic sensor-based SLAM filter is given, regardless of the sensor suite chosen, paving the way for the specific details presented in each of the chapters of Part I. An introduction to the design of the algorithm for converting the sensor-based maps provided by each filter to Earth-fixed quantities is also presented, contextualizing the work of Part II.

Part I refers to the subject of sensor-based SLAM in its various formulations, containing Chapters 3 to 6. The range-and-bearing SLAM nonlinear system is the one that requires the least manipulation to transform into an LTV system of the ones studied along the thesis. Furthermore, conceptually, it is the *simplest* problem to tackle since measurements provide all the information necessary for mapping. For these reasons, this was the first formulation of SLAM to be addressed chronologically in the development of this work and is the first to be presented in the dissertation.

Chapter 3 addresses the problem of the design, analysis, implementation and experimental validation of a sensor-based SLAM filter for 3-D environments, rooted in the work presented in Chapter 2. This chapter is focused on the range-and-bearing variant, and includes the estimation of both body-fixed linear velocity and rate gyro measurement bias. The resulting GES filter is implemented and tested experimentally, using an instrumented quadrotor equipped with an RGB-D camera.

In Chapter 4, a 3-D range-only simultaneous localization and mapping GES filter is proposed. While the problem is intrinsically nonlinear, due to the measurement model, a state augmentation is exploited allowing the proposed formulation to be considered as linear time-varying without linearizing the original nonlinear system. A Kalman filter with GES error dynamics is implemented and simulation results are provided. Furthermore, the algorithm is also experimentally validated with an experimental setup containing a constellation of radio/acoustic beacons and an instrumented quadrotor equipped with a radio/acoustic receiver, an IMU, and an RGB-D camera for visual odometry.

This framework is adapted for applications in sensor networks, which exhibit some particularities, in Chapter 5. There the range-only approach of Chapter 4 is expanded, exploiting the communication capabilities of the transceivers associated with range-only navigation systems, and establishing a proper sensor network. This leads to the design of a simultaneous localization and mapping in sensor networks (SN-SLAM) filter. Furthermore, capabilities of moving object tracking are also addressed. A filter for simultaneous localization and mapping in sensor networks with moving object tracking (SN-SLAMMOT) with GES error dynamics is designed using ranges between all the members of the network, and simulation results are presented.

Chapter 6 addresses the remaining formulation of SLAM, using bearing-only sensors. Following the approach detailed in Chapter 2, and using the experience gained with the work of the previous chapters, the nonlinearity present in the measurement model is avoided through an output transformation and a state augmentation. Then, a sensor-based bearing-only simultaneous localization and mapping filter is designed and implemented. Further specifics of the implementation are provided, and simulation results are accompanied by a detailed report on experimental results obtained using a monocular camera in a wheeled ground robot and data from the *Rawseeds* dataset [BBF+06, CFG+09].

Simultaneous localization and mapping is usually associated with autonomous vehicles.

These operate in a large number of environments, and come in a wide variety, from underwater to airborne vehicles with every possibility in between. SLAM is generically useful for all of them, but each specific formulation may have advantages in certain kinds of environments. However, these advantages are related more to the physical characteristics of the kind of sensors typically associated with each formulation than to the actual formulation itself. With that in mind, it should be noted that even though the experiments reported in each chapter of this thesis are centred on a specific kind of vehicle (either a quadrotor or a wheeled ground robot), the algorithms proposed therein are applicable to any type of vehicle or agent, autonomous or not, provided that the sensors available supply the necessary measurements for the operation.

The sensor-based SLAM filters of the previous chapters raise the need for an external tool to estimate inertial/Earth-fixed quantities. Part II addresses this issue. This can be accomplished defining an optimization problem equivalent to the orthogonal Procrustes problem. Chapter 7 proposes a novel uncertainty characterization of this optimization grounded on perturbation theory, and presents its extensive statistical validation. While this work is the foundation of the algorithm proposed in that part of the thesis, it is useful in a wide variety of applications, such as medical imaging [FW01] and machine learning [WM08].

Chapter 8 formulates an optimization problem with a solution that corresponds to an estimate of the transformation between the body-fixed frame $\{B\}$ and the Earth-fixed reference frame $\{E\}$, yielding the algorithm here proposed. An error function is defined and then used to construct a cost function for the optimization problem. The algorithm builds on the derivation in Chapter 7 and the uncertainty characterization proposed therein, and aims at estimating, in real time, both the vehicle trajectory and the Earth-fixed map described in the same frame, with the respective uncertainty characterization. The complete SLAM methodology composed of the cascade of a sensor-based filter and the Earth-fixed trajectory and map (ETM) algorithm is validated by a variety of simulated and experimental trials using the data acquired through the course of Part I.

Finally, some concluding remarks and directions of future work are provided in Chapter 9.

In Appendix A, several definitions and supplementary results necessary for the theoretical work of this dissertation are provided, Appendix B presents the proofs of some minor results that are not shown in the proper chapters in order to lighten the reading, Appendix C details the simulated environment used in several of the chapters, and Appendix D provides an overview of the experimental setup of the experiments presented in Chapters 3, 4, and 8 organized by the author in the Sensor-based Cooperative Robotics Research (SCORE) Laboratory at the Department of Electrical and Computer Engineering of the Faculty of Science and Technology, University of Macau.

1.9 Notation

Throughout this thesis, vectors and matrices are represented in small and capital boldface letters, respectively. Scalar symbols are expressed in italic: constants by capital letters, and scalar variables in small letters. Particularly, the symbol $\mathbf{0}_{n\times m}$ denotes an $n \times m$ matrix of zeros (if only one subscript is present, the matrix is square), $\mathbf{I}_n$ is an identity matrix with dimension $n\times n$, $\mathbf{1}_n$ is a $n\times 1$ column vector of ones, and $\bar{\mathbf{1}}_i$ is a vector whose i-th element is equal to 1 and is the only non-zero one. These last two can be distinguished from context.

A block diagonal matrix is represented by $\operatorname{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_n)$. The determinant of a generic square matrix is denoted by $|\mathbf{A}|$ or $\det \mathbf{A}$ when distinction from the absolute value is necessary, and, for a generic matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$, the Frobenius norm is adopted unless otherwise stated, i.e., $\|\mathbf{A}\| = \sqrt{\operatorname{tr}\left(\mathbf{A}\mathbf{A}^T\right)}$. For two square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \preceq \mathbf{A}$ is equivalent to $\mathbf{A} - \mathbf{B} \succeq 0$, meaning that the left-hand side is a positive semidefinite matrix. This extends to positive definite and negative (semi)definite matrices.

The orthogonal group is denoted by $\mathrm{O}(n) := \left\{\mathbf{X} \in \mathbb{R}^{n\times n} : \mathbf{X}\mathbf{X}^T = \mathbf{X}^T\mathbf{X} = \mathbf{I}\right\}$, and the special orthogonal group is denoted by $\mathrm{SO}(n) := \{\mathbf{X} \in \mathrm{O}(n) : |\mathbf{X}| = 1\}$. The operator $\operatorname{skew} : \mathbb{R}^{n\times n} \to \mathfrak{so}(n)$, where $\mathfrak{so}(n)$ is the special orthogonal Lie algebra, yields the skew-symmetric component of a square matrix, $\operatorname{skew}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^T\right)$, and $\mathbf{S}(\mathbf{a})$ is a special skew-symmetric matrix, henceforth called the cross-product matrix, as $\mathbf{S}(\mathbf{a})\,\mathbf{b} = \mathbf{a} \times \mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

The expectation operator is denoted as $\langle . \rangle$, and the covariance matrix between two generic stochastic vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ is denoted by $\Sigma_{ab} = \langle (\mathbf{a} - \langle \mathbf{a} \rangle)\,(\mathbf{b} - \langle \mathbf{b} \rangle)^T \rangle$ or Σ_a, if $\mathbf{a} = \mathbf{b}$.

The derivative with respect to time of any quantity a is denoted by $\dot{a} := \frac{\mathrm{d}a}{\mathrm{d}t}$. These two possibilities are used interchangeably depending on what is best for legibility.

The superscript ^{E} indicates a vector or matrix expressed in the Earth-fixed frame $\{E\}$. For the sake of clarity, when no superscript is present, the vector is expressed in the body-fixed frame $\{B\}$. Regardless of the frame in which it is expressed, the term position of a landmark and landmark are used interchangeably throughout the dissertation.

2

DESIGNING A SENSOR-BASED SLAM ALGORITHM

THIS chapter presents a general overview of the work developed in the remainder of this dissertation. It introduces a class of systems deeply rooted in a sensor-based approach to the simultaneous localization and mapping (SLAM) problem, encompassing the more usual range-and-bearing SLAM problem, as well as the range-only and bearing-only SLAM problems. The generic steps to deal with the nonlinear nature of these systems are proposed and several necessary definitions and results are presented that lay the groundwork for the design of the SLAM filters detailed in Part I of this dissertation. This chapter aims at serving as a reference for the dynamic models and definitions shared between the different chapters, along with the underlying methodology that permeates the design of the different sensor-based filters throughout the thesis, while trying to avoid the redundancy that otherwise would be apparent. In summary, this chapter presents (i) the consolidation and definition of a class of sensor-based SLAM problems such as 3-D range-and-bearing simultaneous localization and mapping (RB-SLAM), range-only simultaneous localization and mapping (RO-SLAM), and bearing-only simultaneous localization and mapping (BO-SLAM); (ii) the methodology used to move from the nonlinear systems that underlie each formulation to systems that can be regarded as linear time-varying (LTV) for observability purposes; (iii) the filter implementation details that ensure the global exponential stability of the respective error dynamics; and (iv) an introduction to the complementary method detailed in Part II to obtain the Earth-fixed quantities from the results of the sensor-based filters.

Chapter Structure

The remaining of the chapter is organized as follows. Section 2.1 introduces the sensor-based SLAM problems while their observability analysis is explained in Section 2.2, along with several preliminary results and concepts necessary for said analysis. The sensor-based filter implementation details are provided in Section 2.3, preceded by a short analysis of the convergence of the filters. Finally, Section 2.4 introduces the Earth-fixed trajectory and map algorithm.

2.1 The sensor-based framework

Simultaneous localization and mapping is the quintessential issue in operating robotic autonomous vehicles, and, as such, a myriad of different approaches have been spawned since its inception by a flourishing scientific community. These approaches differ in a variety of aspects, from the filtering engine to the perception sensors (and therefore type of measurements). Most of these employ a typical idea in the robotics community that stems from the ages-old concept of human-readable maps, which is to work in an inertial frame. This requires that all the measurements, obtained in a reference frame fixed to the vehicle, are transformed into a common inertial frame that is only known deterministically in the very beginning of the process and whose uncertainty grows until the vehicle visits the same place twice.

A different approach is to take advantage of the fact that all the measurements are expressed in a common frame and therefore perform the estimation process in that same frame - the sensor-based approach. This has itself two varieties: (i) the incremental pose of the vehicle is kept as an unobservable state; or (ii) pose representation is suppressed.

This thesis addresses the purely sensor-based framework, which tries to avoid the pitfalls of pose representation in the filter state, one of the main sources of nonlinearity due to the fact that all measurements are expressed in local coordinates. It must be noted that, even though a filter in that framework provides strictly sensor-based estimates, it is still possible to obtain inertial estimates by having at least two landmarks as anchors with known coordinates, or using an algorithm based on the Procrustes problem [SC70] as will be explained in the last section of this chapter, and further expanded upon in Part II of this dissertation.

2.1.1 Coordinate frames

The SLAM filters proposed in this dissertation are based on linear motion kinematics, and explore the possibilities brought to light by the careful choice of reference frame on which to define the quantities to estimate. This section aims at presenting the necessary definitions that underlie all the mathematical derivations spread throughout this dissertation. For that purpose, take the situation schematized in Figure 2.1. There are presented landmarks in the form of points in 3-D space, along with two coordinate frames: (i) an Earth-fixed frame, denoted by $\{E\}$, which is considered stationary for the purpose of this work, making it an inertial frame; and (ii) the frame $\{B\}$, which denotes the body-fixed frame and is also referred to as the sensor-based frame, as it is the frame in which all the vehicle-mounted sensors provide their measurements. Points in the latter frame are mapped to the former through a transformation that is fully defined by a rotation and a translation, encoded by the pair $\left(\mathbf{R}(t), {}^{E}\mathbf{p}(t)\right) \in \mathrm{SO}(3) \times \mathbb{R}^3$, where $\mathrm{SO}(n)$ is the special orthogonal group. ${}^{E}\mathbf{p}(t) \in \mathbb{R}^3$ represents the position of the vehicle in the Earth-fixed frame, thus also denoting the origin of the body-fixed frame, and $\mathbf{R}(t) \in \mathrm{SO}(3)$ is a rotation matrix that represents the attitude of the vehicle, and is defined as follows.

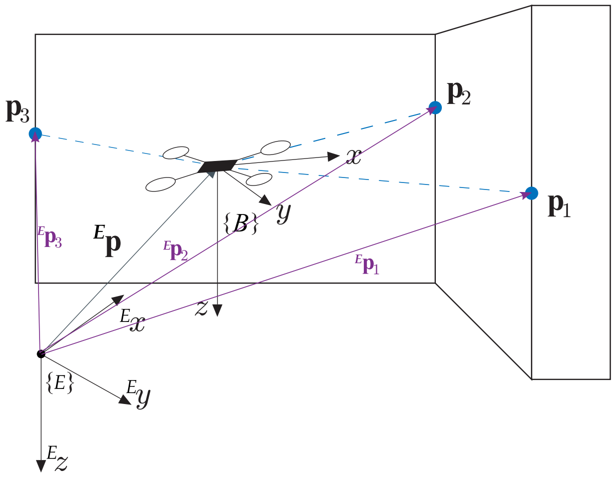


Figure 2.1: Schematic of the reference frames $\{B\}$ and $\{E\}$, and the landmark-based environment.

Definition 2.1 (Rotation matrix, [Cra05, Section 2] and [Gan60, Section XI.14])**.** *A matrix* $\mathbf{R}$ *belonging to the special orthogonal group* $\mathrm{SO}(3)$ *represents a norm-preserving linear transformation and respects the following conditions:*

1. *The determinant of* $\mathbf{R}$ *is unitary,* $|\mathbf{R}| = 1$*;*
2. *The inverse of* $\mathbf{R}$ *is its transpose,* $\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$*;*
3. *The derivative with respect to time is given by*

$$\dot{\mathbf{R}}(t) = \mathbf{R}(t)\mathbf{S}(\boldsymbol{\omega}(t))$$

where $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ *is the angular velocity of the frame* $\{B\}$ *expressed in its own coordinates, and the matrix* $\mathbf{S}(\boldsymbol{\omega}) \in \mathfrak{so}(3)$ *is a skew-symmetric matrix parametrized by the vector, and given by*

$$\mathbf{S}(\boldsymbol{\omega}) = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}.$$

It is henceforth called the cross-product matrix, as $\mathbf{S}(\mathbf{a})\mathbf{b} = \mathbf{a} \times \mathbf{b}$ *for* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

4. *The matrix* $\mathbf{S}(\boldsymbol{\omega})$ *can be mapped to* $\mathrm{SO}(3)$ *through the exponential map, i.e.,*

$$\mathbf{R} = \exp(\mathbf{S}(\boldsymbol{\omega})), \quad \textit{or} \quad \mathbf{S}(\boldsymbol{\omega}) = \log(\mathbf{R}).$$

The pair $\left(\mathbf{R}(t), {}^E\mathbf{p}(t)\right)$ is also a representation of the pose (position and attitude) of the vehicle.

2.1.2 Nonlinear System Design

Following the discussion in the introductory chapter, this section details the design of dynamical systems as part of the sensor-based simultaneous localization and mapping filters using only one source of external environment perception, apart from vehicle motion information.

The environment is characterized by static point landmarks, henceforth denoted merely as landmarks, that may be naturally extracted or artificially placed, and whose relative coordinates, ranges (distances), or bearings (directions) can be perceived by the vehicle that is at the heart of this scenario. Recalling Table 1.1, landmarks can be (i) obtained through feature detection/extraction in images, in which case corners, edges, blobs, and points-of-interest are part of the possibilities; or (ii) artificial beacons/sensors placed in the environment. Even though the figure depicts a rendition of an aerial vehicle, the work in this dissertation does not make assumptions on the type of vehicle used, apart from the ones regarding the on-board sensor-suite for each SLAM formulation that are referred to directly in the text. This is possible given that the systems to be designed rely on exact tridimensional kinematic models.

Let $\mathcal{M} := \{1, \ldots, N\}$ be a set of N landmarks fixed in the environment, to be mapped. Depending on the pose of the vehicle, some of these landmarks may be visible or not, which motivates the definition of two subsets of landmarks, $\mathcal{M}_O = \{1, \ldots, N_O\}$ and $\mathcal{M}_U = \{N_O + 1, \ldots, N\}$. The first contains the N_O observed, or visible, landmarks while the latter contains the N_U unobserved, or non-visible, ones, such that $\mathcal{M} = \mathcal{M}_O \cup \mathcal{M}_U$. Note that, without loss of generality, the landmarks are ordered for simplicity of analysis. Furthermore, suppose that $\mathbf{p}_i(t) \in \mathbb{R}^3$ corresponds to a sensor-based landmark in the set $\mathcal{M}$, i.e., the position of the i-th landmark relative to the vehicle expressed in $\{B\}$, given by

$$\mathbf{p}_i(t) = \mathbf{R}^T(t)\left({}^E\mathbf{p}_i(t) - {}^E\mathbf{p}(t)\right),$$

where ${}^E\mathbf{p}_i(t) \in \mathbb{R}^3$ corresponds to the Earth-fixed position of the landmark. As in the Earth-fixed frame landmarks are assumed static, considering the motion of the landmark in $\{B\}$, it is possible to write

$$\dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t))\,\mathbf{p}_i(t) - \mathbf{v}(t)$$

where $\mathbf{v}(t) = \mathbf{R}^T(t)^E\dot{\mathbf{p}} \in \mathbb{R}^3$ is the linear velocity of the vehicle expressed in $\{B\}$, and $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is its angular velocity. It is generally considered that the vehicle is equipped with a triad of orthogonally mounted rate-gyros, rendering the angular velocity available. This is not necessarily the case for the linear velocity, which may have to be estimated or may be available through (visual) odometry or optical flow measurements.

Landmarks belonging to $\mathcal{M}_O$ will have some kind of system output associated, which leads to the definition of

$$\mathbf{y}_i(t) = \mathbf{h}(\mathbf{p}_i(t)), \quad i \in \mathcal{M}_O$$

where $\mathbf{y}_i(t)$, a component of the system output signal, can be equal to $\mathbf{p}_i(t)$, $r_i(t) := \|\mathbf{p}_i(t)\|$, or $\mathbf{b}_i(t) := \frac{\mathbf{p}_i(t)}{\|\mathbf{p}_i(t)\|}$ according to the version of sensor-based SLAM filter to be designed (see Figure 2.2). Combining this information it is now possible to write the generic nonlinear system

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t))\,\mathbf{p}_i(t) - \mathbf{v}(t) & i \in \mathcal{M} \\ \mathbf{y}_j(t) = \mathbf{h}(\mathbf{p}_j(t)) & j \in \mathcal{M}_O \end{cases}, \tag{2.1}$$

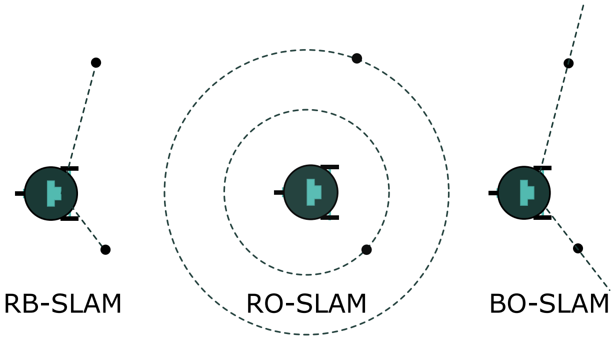


Figure 2.2: For RB-SLAM (Chapter 3), sensors measure the position of a landmark relative to the vehicle, for RO-SLAM (Chapters 4 and 5), the distance to a landmark, and for BO-SLAM (Chapter 6), the relative direction to a landmark.

Table 2.1: Summary of the information of the dynamic systems that underlie each SLAM formulation. The last row summarizes the state augmentation and output transformation details used to obtain LTV-like systems necessary for the following sections.

	RB-SLAM	RO-SLAM	BO-SLAM
States	▷ map ▷ linear velocity ▷ bias estimation	▷ map ▷ ranges	▷ map ▷ ranges
Inputs/ Outputs	▷ biased angular velocity ▷ landmarks	▷ angular velocity ▷ linear velocity ▷ ranges	▷ angular velocity ▷ linear velocity ▷ bearings
Details	▷ $\mathbf{S}(\omega)\,\mathbf{p}_i$ ▷ $\omega = \omega_m - \mathbf{b}_\omega$ ▷ $\mathbf{S}(\mathbf{y}_i)\,\mathbf{b}_\omega$	▷ new state r_i	▷ $\mathbf{b}_i = \mathbf{p}_i / \lVert\mathbf{p}_i\rVert$ ▷ $\mathbf{b}_i \,\lVert\mathbf{p}_i\rVert - \mathbf{p}_i = \mathbf{0}$ ▷ new state r_i

▷ States ▷ Measurements

which is purely based on the rigid-body kinematics.

If $\omega(t)$ is an input and $\mathbf{v}(t)$ is either added as a state with constant dynamics or kept as an input, then the first equation in (2.1) can be considered linear time-varying for observability analysis purposes. The main problem rests with the output equation that may be nonlinear. In that case, further action is necessary to obtain a linear-like system: state augmentation and/or output transformation, leading to a new state $\mathbf{x}(t)$ containing the state of the original nonlinear system and whichever new states the augmentation introduced. With these manipulations, it is possible to obtain dynamical systems for sensor-based SLAM that, when looking only at the visible landmarks, resemble LTV systems that mimic the original nonlinear systems. While the full details of each case are explained in the respective chapter, Table 2.1 summarizes these manipulations depending on the version of SLAM at hand, along with the state and output information. In the details row of this table, it is shown a glimpse of the nonlinear terms at hand and how that problem is tackled. The following section deals with the steps necessary to assess the observability of the resulting systems and their relation with the original nonlinear systems.

2.2 Observability

Despite the linear-like nature of the systems resulting from the system design procedure described in the previous section, the systems are in fact nonlinear: due to the presence of non-visible landmarks in all of them, the state augmentation introduces nonlinear terms that prevent the use of linear tools for analysis. The quantities associated with the non-visible landmarks are, by definition, not observable. Therefore, it is reasonable to discard them from the state when analysing observability, stability, and convergence (see [GBSO13] for a previous successful application of this approach). This yields a reduced system of the form

$$\begin{cases} \dot{\mathbf{z}}(t) = \mathbf{A}(t,\mathbf{y}(t),\mathbf{u}(t))\mathbf{z}(t) + \mathbf{B}(t,\mathbf{y}(t),\mathbf{u}(t))\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}(t,\mathbf{y}(t),\mathbf{u}(t))\mathbf{z}(t) \end{cases}, \tag{2.2}$$

whose dynamics, even though still nonlinear, do not depend on the state $\mathbf{z}(t)$ itself but on measurable quantities, such as inputs $\mathbf{u}(t)$ and outputs $\mathbf{y}(t)$. As these are known functions of time, then, for observability analysis and observer design purposes, the system (2.2) can in fact be considered as linear time-varying.

Having reached this stage in the system design and analysis, it is important to present several concepts necessary for the observability analysis.

2.2.1 Preliminary definitions and results

Consider the definitions of the transition matrix and observability Gramian, paramount to the observability analysis that is the main focus of the theoretical work in the design of globally convergent SLAM filters.

Definition 2.2 (Transition matrix, [Rug96, Chapter 3]). *Consider the generic system* (2.2). *The transition matrix $\boldsymbol{\phi}(t,t_0)$ of a given state space model maps the state from time instant t_0 to time t. The general solution of the system* (2.2) *is given by*

$$\mathbf{z}(t) = \boldsymbol{\phi}(t,t_0)\mathbf{z}(t_0) + \int_{t_0}^{t} \boldsymbol{\phi}(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\mathrm{d}\tau,$$

where the transition matrix $\boldsymbol{\phi}(t,t_0)$ must always satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}(t,t_0) = \mathbf{A}(t)\boldsymbol{\phi}(t,t_0),$$

with the initial condition $\boldsymbol{\phi}(t_0,t_0) = \mathbf{I}$.

Definition 2.3 (Observability Gramian, [Rug96, Chapter 9]). *The observability Gramian in the time interval $\mathcal{T} = [t_0, t_f]$ of system* (2.2) *is defined as*

$$\mathcal{W}(t_0,t_f) := \int_{t_0}^{t_f} \boldsymbol{\phi}^T(\tau,t_0)\mathbf{C}^T(\tau)\mathbf{C}(\tau)\boldsymbol{\phi}(\tau,t_0)d\tau, \tag{2.3}$$

where $\boldsymbol{\phi}(\tau,t_0)$ is the transition matrix associated with (2.2).

Consider now the following lemma that provides the basis for the observability analysis of the class of systems (2.2).

Lemma 2.1 (Observability, [BSO11b, Lemma 1])**.** *Consider the nonlinear system* (2.2)*. If the observability Gramian* $\mathcal{W}(t_0, t_f)$*, associated with the pair* $(\mathbf{A}(t, \mathbf{y}(t), \mathbf{u}(t)), \mathbf{C}(t, \mathbf{y}(t), \mathbf{u}(t)))$ *on* $\mathcal{T}$ *is invertible, then the system* (2.2) *is observable in the sense that, given the system input* $\{\mathbf{u}(t), t \in \mathcal{T}\}$ *and the system output* $\{\mathbf{y}(t), t \in \mathcal{T}\}$*, the initial condition* $\mathbf{z}(t_0)$ *is uniquely defined.*

Proof. The proof can be found in [BSO11b, Section 3]. ☐

Since the system (2.2) can be considered LTV for observability purposes, the dependences on $(t, \mathbf{y}(t), \mathbf{u}(t))$ will be dropped in favour of a lighter notation showing only the dependence on time.

The ultimate goal of the observability analysis is to allow the design of an observer with guaranteed convergence and stability properties. Due to the time-varying nature of the reduced systems at hand, a stronger form of observability than that which is described in the previous lemma is necessary.

Definition 2.4 (Uniform complete observability, [SD82, Definition VI.3])**.** *The pair* $(\mathbf{A}(t), \mathbf{C}(t))$ *from system* (2.2)*, regarded as LTV, is called uniformly completely observable (UCO) if there exist positive constants* δ*,* α_1*, and* α_2 *such that*

$$\alpha_1 \mathbf{I} \leq \mathcal{W}(t + \delta, t) \leq \alpha_2 \mathbf{I},$$

for all $t \geq t_0$*. Alternatively, if the system matrices* $\mathbf{A}(t)$ *and* $\mathbf{C}(t)$ *are norm-bounded (see [SA68, Section 2]), the above definition is equivalent to*

$$\underset{\substack{\delta > 0 \\ \alpha > 0}}{\exists} \quad \underset{t \geq t_0}{\forall} \quad \underset{\substack{\mathbf{c} \in \mathbb{R}^{n_z} \\ \|\mathbf{c}\| = 1}}{\forall} : \quad \mathbf{c}^T \mathcal{W}(t, t + \delta)\mathbf{c} \geq \alpha. \tag{2.4}$$

2.2.2 The analysis

The observability analysis tries to ascertain whether and in what conditions the LTV systems are observable by studying their observability Gramians, as indicated by Lemma 2.1. In order to facilitate the computation of the transition matrices, leading to that of the observability Gramians, Lyapunov transformations (see Definition A.1) aligning the landmark positions with the Earth-fixed frame are proposed for each flavour of SLAM studied. Since Lyapunov transformations maintain observability properties, the transformed systems are used for the analysis. However, it must be noted that, since the class of systems (2.2) was designed with the intention of mimicking the dynamics of the various versions of the original nonlinear system (2.1), when discarding the non-visible landmarks, an observability study of each version of SLAM would not be complete without properly addressing that system. The study of the observability of the original nonlinear system starts with a comparison between the state of the LTV system and that of the nonlinear system. This includes investigating the conditions in which the state augmentation relations become naturally imposed by the dynamics. It is in that case that the states of the two systems become equivalent, which validates the augmentation approach in the sense that a state observer with uniformly globally exponentially stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, and the estimation error converges exponentially fast for all initial conditions. The final step before proceeding to observer design is to study the uniform complete observability (UCO) of the LTV system. This is a stronger form of observability that is necessary to guarantee the global exponential stability of the Kalman filter. For this purpose, uniform bounds on the observability Gramian calculated on a moving time interval are investigated. Due to this uniformity, the resulting conditions are more demanding, and can be regarded as persistent excitation-like conditions. Figure 2.3 summarizes the whole process presented in the previous section and further explained here, establishing the methodology that permeates the design and analysis of the sensor-based filters in Part I.

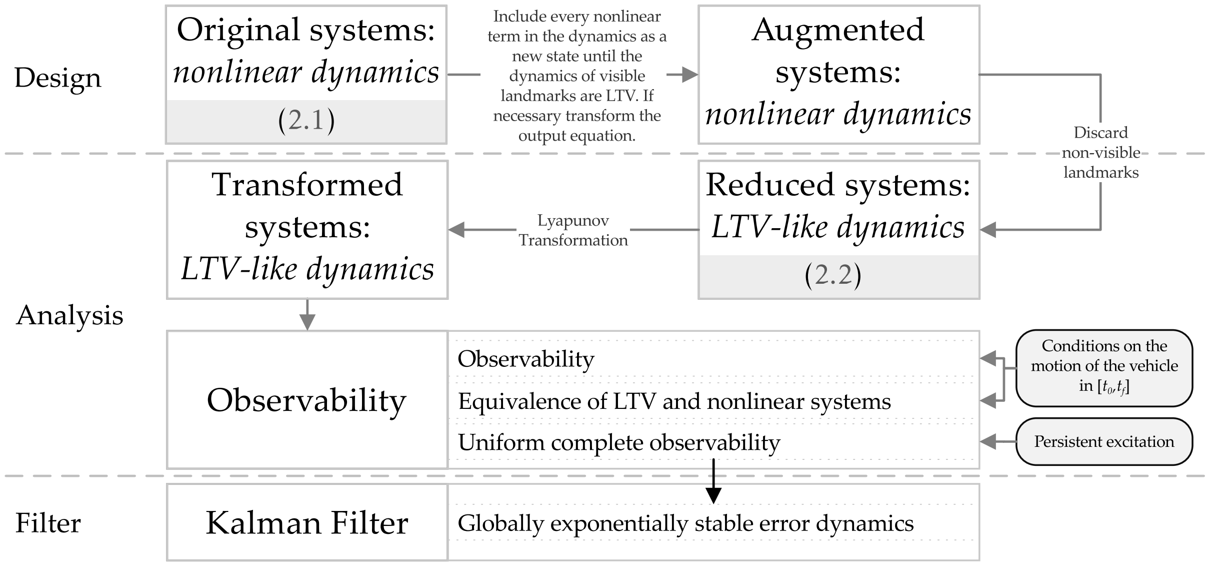

Design
Original systems:
nonlinear dynamics
(2.1)
Include every nonlinear term in the dynamics as a new state until the dynamics of visible landmarks are LTV. If necessary transform the output equation.
Augmented systems:
nonlinear dynamics
Discard non-visible landmarks
Analysis
Transformed systems:
LTV-like dynamics
Lyapunov Transformation
Reduced systems:
LTV-like dynamics
(2.2)
Observability
Observability
Equivalence of LTV and nonlinear systems
Uniform complete observability
Conditions on the motion of the vehicle in $[t_0, t_f]$
Persistent excitation
Filter
Kalman Filter
Globally exponentially stable error dynamics

Figure 2.3: Schematic description of the process of designing a globally convergent sensor-based SLAM filter.

In range-and-bearing SLAM, the conditions for observability depend on the number of landmarks observed, due to the fact that the non-measured quantities (linear velocity and rate-gyro bias) affect all the landmarks. On the other hand, the conditions for the observability of both bearing-only and range-only SLAM do not depend on the number of landmarks, as they are independent of each other. There is another important distinction between the two classes of problems pertaining the quantity of information made available with each measurement. In range-and-bearing SLAM one single measurement provides all the necessary information to estimate the position of a landmark, even though several landmarks are needed for immediate full state recovery, whereas in range-only and bearing-only SLAM measurements from several viewpoints have to be acquired to allow for landmark estimation. This is also the case for range-and-bearing SLAM when the number of available landmarks is not enough to guarantee observability without motion. Since observability is only guaranteed with measurements from different viewpoints, this imposes restrictions on the trajectories of the vehicle that lead to an observable system, and hence to a convergent filter. However, these conditions are intrinsic to each estimation problem and not specific of the algorithms. For example, a stopped vehicle will render RO-SLAM and BO-SLAM unobservable, but may still be able to provide enough information for an RB-SLAM filter estimating both the linear velocity and the rate gyro bias to be convergent. This will be seen in detail in the corresponding chapters.

2.3 Filter design

The theoretical analysis of the previous section was established in a deterministic setting, and thus the presence of measurement noise raises the need for a filtering solution. As detailed in the previous section, the underlying systems that define the several SLAM problems can be regarded as LTV systems, which can be shown to be uniformly completely observable under certain conditions. This leads towards the design of a state observer with globally exponentially stable error dynamics for the system (2.2), such as the Kalman filter for linear time-varying systems. The sequel addresses the conditions in which these convergence and stability properties hold.

2.3.1 Convergence Analysis

Consider the dynamics of the continuous-time observer at hand, the Kalman-Bucy filter [KB61],

$$\begin{cases} \dot{\hat{\mathbf{z}}}(t) = \mathbf{A}(t)\hat{\mathbf{z}}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{P}(t)\mathbf{C}^T(t)\boldsymbol{\Theta}^{-1}\mathbf{C}(t)(\mathbf{z}(t) - \hat{\mathbf{z}}(t)) \\ \dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \boldsymbol{\Xi}(t) - \mathbf{P}(t)\mathbf{C}^T(t)\boldsymbol{\Theta}^{-1}(t)\mathbf{C}(t)\mathbf{P}(t) \end{cases},$$

where $\hat{\mathbf{z}}(t)$ is the state estimate, $\mathbf{P}(t)$ is the filter covariance, $\boldsymbol{\Theta}(t) \in \mathbb{R}^{n_h \times n_h}$ is a positive definite covariance matrix associated with the measurement noise and $\boldsymbol{\Xi}(t) \in \mathbb{R}^{n_z \times n_z}$ is a positive semi-definite covariance matrix associated with the process disturbances. To analyse the convergence properties of the observer, let $\tilde{\mathbf{z}}(t) := \mathbf{z}(t) - \hat{\mathbf{z}}(t)$ denote the estimation error, and consider its derivative with respect to time, given by

$$\dot{\tilde{\mathbf{z}}}(t) = \left(\mathbf{A}(t) - \mathbf{P}(t)\mathbf{C}^T(t)\boldsymbol{\Theta}^{-1}(t)\mathbf{C}(t)\right)\tilde{\mathbf{z}}(t), \tag{2.5}$$

where the error dynamics matrix is now denoted as $\tilde{\mathbf{A}}(t) = \left(\mathbf{A}(t) - \mathbf{P}(t)\mathbf{C}^T(t)\boldsymbol{\Theta}^{-1}(t)\mathbf{C}(t)\right)$. The convergence and stability of the error dynamics (2.5) of the Kalman filter for norm-bounded matrices $\mathbf{A}(t)$, $\mathbf{C}(t)$, $\boldsymbol{\Xi}(t)$, and $\boldsymbol{\Theta}(t)$ are well-known, but are summarized in the following result.

Theorem 2.2 (Globally exponential stability of the observer). *If the system* (2.2) *is uniformly completely observable, then the origin of the observer error dynamics* (2.5) *is a globally exponentially stable equilibrium point.*

Proof. This proof will follow the steps of [Kha02, Example 8.11]. Consider the quadratic function

$$V(t,\tilde{\mathbf{z}}(t)) = \tilde{\mathbf{z}}^T(t)\mathbf{P}^{-1}(t)\tilde{\mathbf{z}}(t). \tag{2.6}$$

From [And71, Lemma 3.2] it is known that if $\mathbf{A}(t)$, $\mathbf{\Xi}(t)$, and $\mathbf{\Theta}^{-1}(t)$ are bounded, and if $(\mathbf{A}(t),\mathbf{C}(t))$ is uniformly completely observable, then $\mathbf{P}(t)$ is bounded, i.e., $\mathbf{P}^{-1}(t)$ is definite positive, for all $t \geq t_0$. In the light of this result, $V(t,\tilde{\mathbf{z}})$ is a valid Lyapunov candidate function in the conditions of the theorem, and, in fact it satisfies

$$\alpha_m\|\tilde{\mathbf{z}}(t)\|^2 \leq V(t,\tilde{\mathbf{z}}(t)) \leq \alpha_M\|\tilde{\mathbf{z}}(t)\|^2, \tag{2.7}$$

for all $\tilde{\mathbf{z}}(t) \in \mathbb{R}^{n_z}$, with $\alpha_m, \alpha_M > 0$. This is the first assumption of [Kha02, Theorem 8.5] for global exponential stability.

The derivative of (2.6) is given by

$$\dot{V}(t,\tilde{\mathbf{z}}(t)) = \tilde{\mathbf{z}}^T(t)\left(\tilde{\mathbf{A}}^T(t)\mathbf{P}^{-1}(t) + \mathbf{P}^{-1}(t)\tilde{\mathbf{A}}(t) + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}^{-1}(t)\right)\tilde{\mathbf{z}}(t), \tag{2.8}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}^{-1}(t) = -\mathbf{P}^{-1}(t)\mathbf{A}(t) - \mathbf{A}^T(t)\mathbf{P}^{-1}(t) - \mathbf{P}^{-1}(t)\mathbf{\Xi}(t)\mathbf{P}^{-1}(t) + \mathbf{C}^T(t)\mathbf{\Theta}^{-1}(t)\mathbf{C}(t).$$

Replacing this in (2.8) and expanding $\tilde{\mathbf{A}}(t)$, leads to

$$\begin{aligned}\dot{V}(t,\tilde{\mathbf{z}}(t)) &= -\tilde{\mathbf{z}}^T(t)\left(\mathbf{P}^{-1}(t)\mathbf{\Xi}(t)\mathbf{P}^{-1}(t) + \mathbf{C}^T(t)\mathbf{\Theta}^{-1}(t)\mathbf{C}(t)\right)\tilde{\mathbf{z}}(t)\\ &\leq -\tilde{\mathbf{z}}^T(t)\mathbf{C}^T(t)\mathbf{\Theta}^{-1}(t)\mathbf{C}(t)\tilde{\mathbf{z}}(t) \qquad\qquad ,\\ &\leq 0\end{aligned}$$

which satisfies the second assumption of [Kha02, Theorem 8.5].

The solution of system (2.5) is given by $\tilde{\mathbf{z}}(\tau) = \tilde{\boldsymbol{\phi}}(\tau,t)\tilde{\mathbf{z}}(t)$, where $\tilde{\boldsymbol{\phi}}(\tau,t)$ is the transition matrix for the error dynamics. Consider now the last condition of [Kha02, Theorem 8.5], for some $\delta > 0$,

$$\begin{aligned}V(t+\delta,\tilde{\mathbf{z}}(t+\delta)) - V(t,\tilde{\mathbf{z}}(t)) &= \int_t^{t+\delta} \dot{V}(\tau,\tilde{\mathbf{z}}(\tau))\mathrm{d}\tau\\ &= -\tilde{\mathbf{z}}^T(t)\int_t^{t+\delta} \tilde{\boldsymbol{\phi}}^T(\tau,t)\mathbf{P}^{-1}(t)\mathbf{\Xi}(t)\mathbf{P}^{-1}(t)\tilde{\boldsymbol{\phi}}(\tau,t)\mathrm{d}\tau\,\tilde{\mathbf{z}}(t)\\ &\quad -\tilde{\mathbf{z}}^T(t)\int_t^{t+\delta} \tilde{\boldsymbol{\phi}}^T(\tau,t)\mathbf{C}^T(t)\mathbf{\Theta}(t)\mathbf{C}(t)\tilde{\boldsymbol{\phi}}(\tau,t)\mathrm{d}\tau\,\tilde{\mathbf{z}}(t)\\ &\leq -\tilde{\mathbf{z}}^T(t)\tilde{\boldsymbol{\mathcal{W}}}(t,t+\delta)\tilde{\mathbf{z}}(t)\end{aligned}$$

where $\tilde{\boldsymbol{\mathcal{W}}}(t,t+\delta)$ is the observability Gramian of the pair $(\tilde{\mathbf{A}}(t),\mathbf{\Theta}^{-\frac{1}{2}}(t)\mathbf{C}(t))$. Consider now that the conditions of the theorem hold, i.e., the pair $(\mathbf{A}(t),\mathbf{C}(t))$ is uniformly completely observable. Since $\mathbf{P}(t)$ and $\mathbf{\Theta}^{-1}(t)$ are positive definite for all $t \geq t_0$, the pair $(\tilde{\mathbf{A}}(t),\mathbf{\Theta}^{-\frac{1}{2}}(t)\mathbf{C}(t))$ can also be shown to be uniformly completely observable (see [And71] and references therein for details). This means that $\tilde{\mathcal{W}}(t, t+\delta)$ is bounded, which, using (2.7) yields

$$\begin{aligned} V(t+\delta, \tilde{\mathbf{z}}(t+\delta)) - V(t, \tilde{\mathbf{z}}(t)) &\leq -\tilde{\alpha}_m \|\tilde{\mathbf{z}}(t)\|^2 \\ &\leq -\frac{\tilde{\alpha}_m}{\alpha_M} V(t, \tilde{\mathbf{z}}(t)) \end{aligned},$$

thus satisfying the last assumption of [Kha02, Theorem 8.5]. Since all the assumptions of that result are satisfied globally, the error dynamics of the Kalman filter for the LTV system (2.2) are globally exponentially stable when the conditions of the theorem hold, i.e., when that system is uniformly completely observable. □

2.3.2 Discrete implementation

Due to the sample/digital-based characteristics of the sensor suites available to each type of SLAM formulation, i.e., their discrete nature, a discrete-time version of system (2.2) is considered and a discrete-time implementation of the Kalman filter [Kal60] is used.

2.3.2.1 Discrete dynamics

Let T_s denote the sampling period of the synchronized array of sensors used in each solution, noting that a multi-rate implementation can be devised. Then, the discrete time steps can be expressed as $t_k = kT_s + t_0$ where $k \in \mathbb{N}_0$ and t_0 is the initial time. The full discretized system is characterized by the state $\mathbf{x}_k := \mathbf{x}(t_k)$, and is obtained using the forward Euler discretization[1]. Considering additive disturbances, the generic discretized system is given by

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{F}_{F_k}\mathbf{x}_k + \mathbf{G}_{F_k}\mathbf{v}_k + \boldsymbol{\xi}_k \\ \mathbf{y}_{k+1} = \mathbf{H}_{F_{k+1}}\mathbf{x}_{k+1} + \boldsymbol{\theta}_{k+1} \end{cases}, \tag{2.9}$$

where the dynamics matrices have the structure

$$\mathbf{F}_{F_k} = \begin{bmatrix} \mathbf{F}_k & \mathbf{0}_{n_z \times n_U} \\ \mathbf{F}_{UO_k} & \mathbf{F}_{U_k} \end{bmatrix}, \mathbf{G}_{F_k} = \begin{bmatrix} T_s\mathbf{B}_k \\ T_s\mathbf{B}_{U_k} \end{bmatrix}, \text{ and } \mathbf{H}_{F_k} = \begin{bmatrix} \mathbf{C}_k & \mathbf{0}_{n_O \times n_U} \end{bmatrix},$$

and where $\mathbf{F}_k$ is the discretized version of matrix $\mathbf{A}(t_k)$ in (2.2), which accounts for the observable part of the system dynamics, whereas matrices with subscripts $(.)_U$ and $(.)_{UO}$ denote the unobservable states and cross terms, respectively. The components of the input and output matrices are defined as follows, $\mathbf{B}_k := \mathbf{B}(t_k)$ and $\mathbf{C}_k := \mathbf{C}(t_k)$. Also, the vectors $\boldsymbol{\xi}_k$ and $\boldsymbol{\theta}_k$ represent the model disturbance and measurement noise, respectively, assumed to be zero-mean discrete white Gaussian noise with covariances $\langle \boldsymbol{\xi}_k \boldsymbol{\xi}_l^T \rangle = \boldsymbol{\Xi}_k \delta_{k-l}$ and $\langle \boldsymbol{\theta}_k \boldsymbol{\theta}_l^T \rangle = \boldsymbol{\Theta}_k \delta_{k-l}$, respectively. This is an approximation of the real noise description, since there are multiplicative terms due to the presence of the angular velocity in $\mathbf{F}_{F_k}$ and of the measured quantities ($\mathbf{p}_i(t)$, $r_i(t)$, or $\mathbf{b}_i(t)$) in the dynamics, input or output matrices.

---

[1] For the formulations that do not consider biased angular rate measurements, special care is afforded to the rotation of a landmark from one instant to the following. See Section 4.4 and Subsection 6.3.3 for more information on this matter.

In what concerns the observability of this new discrete system, the applicability of the observability results derived for the continuous case is not trivial, given the nonlinear nature of the system. Nevertheless, the authors of [BSO13a] have designed a source-localization filter in a discrete-time setting with dynamics very similar to the class of systems addressed in this dissertation, and found observability conditions that are directly related to the ones presented here, thus hinting at the applicability of the continuous time conditions to the discrete system, and to the global exponential stability of the ensuing discrete-time Kalman filter.

2.3.2.2 Prediction step

From these discrete LTV systems, the filter prediction and update steps are computed using the standard equations of the Kalman filter for LTV systems [Gel74], with the detail that the non-visible landmarks must be propagated in open loop. The predict equations are

$$\begin{cases}\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{F}}_{F_{k|k}} \hat{\mathbf{x}}_{F_{k|k}} + \mathbf{G}_{F_k} \mathbf{v}_k \\ \Sigma_{k+1|k} = \hat{\mathbf{F}}_{F_{k|k}} \Sigma_{F_{k+1|k}} \hat{\mathbf{F}}^T_{F_{k|k}} + \boldsymbol{\Xi}_{F_k}\end{cases}$$

where $\Sigma_{k+1|k}$ and $\hat{\mathbf{x}}_{k+1|k}$ are the predicted versions of the filter covariance and state estimate, respectively, and the full discrete dynamics matrix is in fact an estimated version, as the block that relates to the non-visible landmarks is created with the estimated non-visible quantities, $\hat{\mathbf{x}}_{U_{k|k}}$. This prediction step uses the measurements of the rate gyros, propagating the state every time a reading is available. Note that, even though the visible landmarks and the vehicle state have LTV-like dynamics, that is only achieved by using directly the landmark and angular velocity measurements in the dynamics. For this reason, this approach is not guaranteed to be optimal, and its noise characterization is not exact. However, the noise parameters $\boldsymbol{\Xi}_{F_k}$ and $\boldsymbol{\Theta}_{F_k}$ can be calibrated a priori with Monte Carlo analysis and with actual measurements, to better cope with this matter.

2.3.2.3 Update step

The update step is divided in two different stages, landmark association and the actual Kalman filter update. This step occurs every time a batch of measurements is obtained from the chosen exteroception sensor. In SLAM formulations that use artificially placed landmarks, the association process is more or less straightforward, depending solely on measurement processing (e.g., tagged signals for RO-SLAM, or visual markers in RB-SLAM or BO-SLAM). In the cases where mapping of natural untagged landmarks is performed, association is much more complex. More information on the actual processes is detailed in the respective chapters, namely in Sections 3.4, 4.4, and 6.4. The association process redefines the new sets of visible and non-visible landmarks, and also provides the innovation vector $\boldsymbol{\nu}_{k+1}$ and its covariance matrix $\Sigma_{\nu_{k+1}}$ given by

$$\begin{cases}\boldsymbol{\nu}_{k+1} = \mathbf{y}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_{k+1|k} \\ \Sigma_{\nu_{k+1}} = \mathbf{H}_{k+1} \Sigma_{k+1|k} \mathbf{H}^T_{k+1} + \boldsymbol{\Theta}_{k+1}\end{cases}. \tag{2.10}$$

The update equations are standard, and given by

$$\begin{cases} \hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\boldsymbol{\nu}_{k+1} \\ \Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\Sigma_{k+1|k}, \\ \mathbf{K}_{k+1} = \Sigma_{k+1|k}\mathbf{H}_{k+1}^{T}\Sigma_{\nu_{k+1}}^{-1} \end{cases}$$

where $\Sigma_{k+1|k+1}$ and $\hat{\mathbf{x}}_{k+1|k+1}$ are the filter covariance and state estimate, updated according to the latest measurement, and $\mathbf{K}_{k+1}$ is the Kalman gain.

2.4 Earth-fixed Trajectory and Map

Simultaneous localization and mapping is intrinsically a nonlinear problem due to the nonlinear nature of the relation between the external perception, which is relative to the vehicle, and the map and the pose of the vehicle, which are expressed in an inertial frame of reference. As such, SLAM assumes many interpretations, approaches, and solutions. However, a great number of algorithms choose to describe the problem in an Earth-fixed frame. To be statistically consistent, this formulation must include the pose of the vehicle either in its state or in the probability distribution that describes the belief of the algorithm in the case of particle filter implementations. Nevertheless, the on-board sensors provide vehicle-centered measurements that, in these formulations, need to be transformed to the Earth-fixed frame. Working directly on a sensor-based framework enables to leave the Earth-fixed pose representation out of the filtering process and avoids the transformation of the measurements. However, there may still be need for an estimate of the Earth-fixed pose of the vehicle, as well as the map of the environment. To complement the sensor-based filter, and due to the possible interest in obtaining Earth-fixed estimates for the vehicle pose and the environment map, it is important to have a strategy that takes the body-fixed map and the initial position and attitude as inputs and is able to compute for each time instant the current Earth-fixed map and the pose of the vehicle. For that purpose, another problem addressed in this dissertation is that of designing an algo-

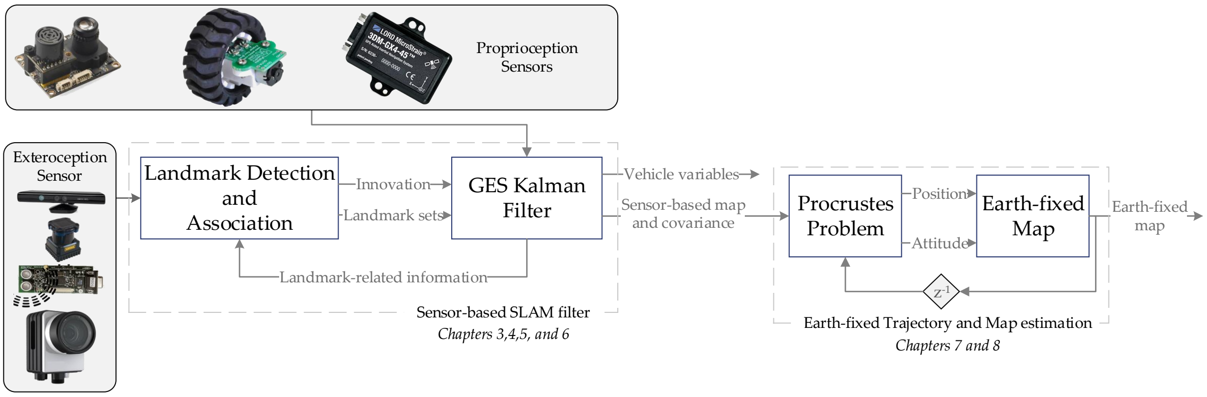

Proprioception Sensors
Exteroception Sensor
Landmark Detection and Association
Innovation
Landmark sets
GES Kalman Filter
Landmark-related information
Sensor-based SLAM filter
Chapters 3,4,5, and 6
Vehicle variables
Sensor-based map and covariance
Procrustes Problem
Position
Attitude
Earth-fixed Map
z^{-1}
Earth-fixed map
Earth-fixed Trajectory and Map estimation
Chapters 7 and 8

Figure 2.4: The generic SLAM algorithm proposed in this dissertation, with references to the relevant chapters, possible exteroception sensors (RGB-D camera, LIDAR, radio/acoustic beacons, and RGB camera), and possible proprioception sensors (optical flow, wheel encoders, and IMU).

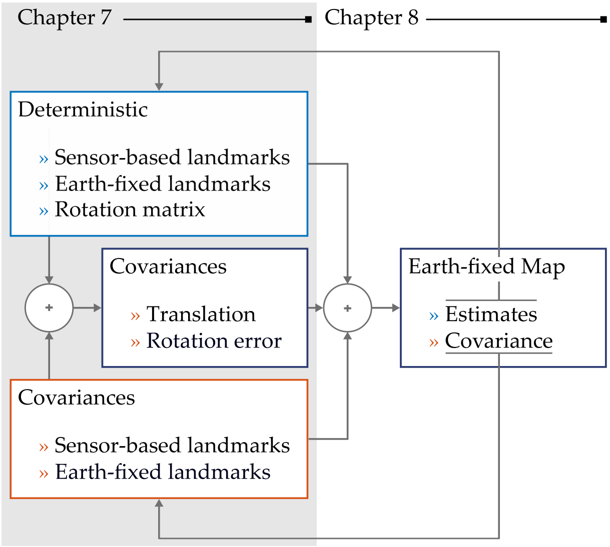


Figure 2.5: Summary of the work performed with the objective of designing the Earth-fixed Trajectory and Map estimation algorithm which is the focus of Part II of this dissertation, showing the focus on the uncertainty of the Procrustes problem in Chapter 7 and the complete algorithm described in Chapter 8.

rithm to compute the Earth-fixed trajectory and map based on the information provided by a sensor-based SLAM filter as detailed so far in this chapter. This is the focus of Part II of this thesis.

Consider the fact that the initial pose of the vehicle in $\{E\}$ is known *a priori* in any SLAM algorithm, as everything is then rooted back to that initial position and orientation. By design, the sensor-based filter provides a landmark map in the body-fixed frame, which means that the Earth-fixed map is readily available in the first instant. Since landmarks are static in $\{E\}$, it is possible to compare the sensor-based map with that initial Earth-fixed estimate through time, which can be updated with new landmarks as they appear. This comparison fits precisely in the definition of the orthogonal Procrustes problem.

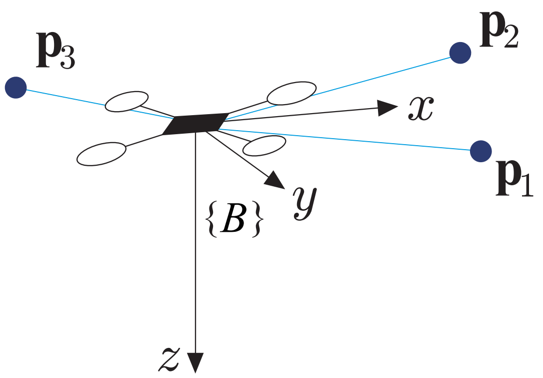
Definition 2.5 (The orthogonal Procrustes Problem)**.** *The problem of finding the translation and rotation that best describe the transformation between two related sets of points is called the Procrustes problem, named after a character in Greek mythology who made his victims fit his bed by either stretching their limbs or cutting them off [GD04].*

A closed-form solution to a particular formulation of this problem was first found in [SC70]. However, the uncertainty of that solution when the two sets are perturbed by noise is still a subject of research. In order to allow the use of this existing possibility in the problem at hand, it is necessary to have a measure of the resulting uncertainty. That is the focus of Chapter 7, where perturbation models consisting of deterministic and perturbed parcels are used to allow the computation of the covariances for the rotation and translation estimation errors as shown in Figure 2.5. Chapter 7 thus lays the foundation for the design of the Earth-fixed algorithm presented in Chapter 8, not only because it provides a method to compute the pose of the vehicle with its uncertainty characterization, but also because the methodology used therein can then be also applied to the step of updating the Earth-fixed map allowing its uncertainty characterization as well.

The final objective is then to build a system that integrates a back-end that performs the mapping in a relative frame (the sensor-based filter), thus guaranteeing convergence and consistency, as well as a front-end that fixes that map to an Earth-fixed, or inertial, frame while providing the pose of the vehicle and the map expressed in the new coordinates. A diagram of the overall structure of the proposed SLAM methodology is presented in Figure 2.4, where the sensor-based SLAM filter can be one of the filters to be detailed in the coming chapters and the Earth-fixed trajectory and map (ETM) algorithm is detailed in Chapter 8.

Part I

SENSOR-BASED SLAM


$\mathbf{p}_3$
$\mathbf{p}_2$
x
$\mathbf{p}_1$
y
$\{B\}$
z

3

SIMULTANEOUS LOCALIZATION AND MAPPING WITH RGB-D VISION

THIS chapter presents the design, analysis, and experimental validation of a globally exponentially stable (GES) filter for range-and-bearing simultaneous localization and mapping (RB-SLAM) with application to unmanned aerial vehicles (UAVs). The main contributions of this chapter are the results of global convergence and stability for RB-SLAM in 3-D environments. The methodology laid out in Chapter 2 is followed, leading to the proposed solution, a Kalman filter with GES error dynamics which includes the estimation of both body-fixed linear velocity and rate gyro measurement biases. Experimental results from several runs, using an instrumented quadrotor equipped with an RGB-D camera, are included in the chapter to illustrate the performance of the algorithm under realistic conditions.

3.1 Introduction

RB-SLAM is a very versatile formulation of simultaneous localization and mapping with the widest variety of possible applications in robotics, autonomous vehicles, among other fields. As long as the necessary computational power to deal with (at least) two sources of images – RGB and depth or stereo cameras – is available, cameras are cheap and provide rich information of the environment which makes them an obvious choice for systems from micro-aerial vehicles in an industrial setting to cars in a large city.

The algorithm proposed in this chapter is rooted in extended Kalman filter (EKF)-SLAM procedures such as [CMCTN07], in which the filtering process is centered on the vehicle. The EKF approach uses a single filter to maintain estimates of the map and vehicle pose, as well as the cross-covariances. This chapter proposes an alternate formulation that uses a Kalman filter and achieves global convergence results by exploring the linear time-varying nature of the sensor-based SLAM system and analysing its observability. Differing from the observability-heavy works in [JS11] and [KS11] that tackle monocular vision SLAM with angular rate and accelerometer measurements mentioned previously, this sensor-based system assumes that 3-D landmark positions and the angular rates are measured, and only includes the local relative map, the linear velocity and the rate-gyro bias. It does not include accelerometer readings, thus avoiding the need to estimate its bias and the gravity vector and rendering the problem less complex. However, and most importantly, the observability results here presented are a sufficient condition for the global exponential stability of the sensor-based RB-SLAM filter, the main contribution of this chapter. This relieves the filter design process of the careful consideration usually devoted to filter initialization. The proposed globally exponentially stable sensor-based RB-SLAM filter: (i) resorts to the linear and angular motion kinematics, which are exact; (ii) uses the low-cost Kinect, in opposition to the two-dimensional (2-D) landmark approach, which demands the use of considerably more expensive laser range finders; (iii) builds on the well-established Kalman filter for linear time-varying systems; and (iv) explicitly estimates the rate gyro measurement bias.

Chapter Structure

The chapter is organized as follows. Section 3.2 presents a short description of the problem, with the definition of the system dynamics. The observability analysis is performed in Section 3.3. The filter design is described in Section 3.4 including landmark detection, data association, and loop closing procedures. Comprehensive experimental results using an instrumented quadrotor are detailed in Section 3.5 to illustrate the performance of the algorithm in real-world conditions, whereas in Section 3.6 the main concluding remarks are provided.

3.2 Problem formulation

Building on the sensor-based idea explained in the previous chapter, this chapter addresses the problem of designing a navigation system in a sensor-based framework for an aerial vehicle capable of sensing landmarks in a previously unknown environment. This is done resorting to a novel SLAM algorithm, where no linearization or approximation is used whatsoever. The only available sensors are a triaxial rate gyro and an RGB-D camera, such as the *Microsoft Kinect*, which provides angular rate measurements and RGB-D images, from where 3-D landmarks may be extracted based on feature detection (corners, edges, blobs, points-of-interest).

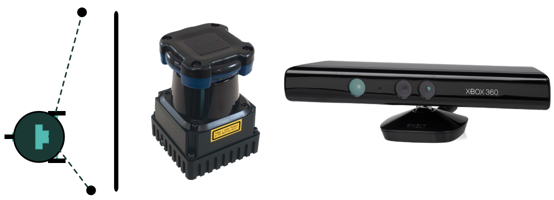


Figure 3.1: In range-and-bearing SLAM, the vehicle is equipped with sensors that are capable of measuring the position of a landmark relative to the vehicle. For example, a LIDAR in 2-D or an RGB-D camera for 3-D. In any of these cases, a feature detection/extraction tool has to be ran on the 2-D scans or the RGB images to obtain usable landmarks.

3.2.1 Nonlinear system dynamics

Recall the generic system (2.1), and consider that the vehicle is equipped with a triad of orthogonally mounted rate gyros, rendering the angular velocity of the vehicle available through the biased rate gyros measurements $\boldsymbol{\omega}_m(t)$ given by

$$\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_\omega(t),$$

where the bias $\mathbf{b}_\omega(t) \in \mathbb{R}^3$ is assumed constant. Taking this into account, and using the cross product property $\mathbf{S}(\mathbf{a})\,\mathbf{b} = -\mathbf{S}(\mathbf{b})\,\mathbf{a}$, it is possible to rewrite the dynamics in (2.1) as

$$\dot{\mathbf{p}}_i(t) = -\mathbf{v}(t) - \mathbf{S}(\mathbf{p}_i(t))\;\mathbf{b}_\omega(t) - \mathbf{S}(\boldsymbol{\omega}_m(t))\;\mathbf{p}_i(t). \tag{3.1}$$

In RB-SLAM, the landmark measurements provide the relative position of each landmark with respect to the vehicle (see Figure 3.1). This is the most information possible in terms of point-based maps, and it may be exploited in order to further estimate other quantities. Following that line of reasoning, assume that the linear velocity is not directly measured, and, as such, needs to be estimated.

The variables presented so far can be divided in the vehicle-related or landmark-related categories. The former, i.e., the linear velocity and the angular measurement bias will constitute the vehicle state, denoted by $\mathbf{x}_V(t) := \begin{bmatrix}\mathbf{v}^T(t) & \mathbf{b}_\omega^T(t)\end{bmatrix}^T \in \mathbb{R}^{n_V}$, with simple dynamics given by

$$\dot{\mathbf{x}}_V(t) = 0,$$

which means that both are assumed, in a deterministic setting, as constant. While this may seem unrealistic in the case of the linear velocity (biases of this kind are usually constant), in the filtering framework, the inclusion of state disturbances allows to consider them as slowly time-varying. The landmarks in each of the sets defined in Section 2.1.2 can be agglomerated in state vectors denoted as $\mathbf{x}_O(t) = \{\mathbf{p}_i(t) \in \mathcal{M}_O\} \in \mathbb{R}^{n_O}$ and $\mathbf{x}_U(t) = \{\mathbf{p}_i(t) \in \mathcal{M}_U\} \in \mathbb{R}^{n_U}$, which together form the map state vector $\mathbf{x}_M(t) = \begin{bmatrix}\mathbf{x}_O^T(t) & \mathbf{x}_U^T(t)\end{bmatrix}^T \in \mathbb{R}^{n_M}$.

With the previous definitions, it is now possible to derive the full state dynamics. For that purpose consider the position landmark dynamics (3.1), which may now be expressed as a function of the state vector, yielding

$$\dot{\mathbf{x}}_{M_i}(t) = \mathbf{A}_{MV_i}(\mathbf{p}_i(t))\,\mathbf{x}_V(t) - \mathbf{S}(\boldsymbol{\omega}_m(t))\mathbf{x}_{M_i}(t),$$

where $\mathbf{A}_{MV_i}(\mathbf{p}_i(t)) = \begin{bmatrix}-\mathbf{I}_3 & -\mathbf{S}(\mathbf{p}_i(t))\end{bmatrix}$.

The two state vectors here defined constitute the full state vector $\mathbf{x}(t) = \begin{bmatrix}\mathbf{x}_V^T(t) & \mathbf{x}_M^T(t)\end{bmatrix}^T$, with the full system dynamics reading as

$$\begin{cases}\dot{\mathbf{x}}(t) = \mathbf{A}_F(t, \mathbf{y}(t), \mathbf{x}_U(t))\,\mathbf{x}(t)\\ \mathbf{y}(t) = \mathbf{x}_O(t)\end{cases}, \tag{3.2}$$

with

$$\mathbf{A}_F(t,\mathbf{y}(t),\mathbf{x}_U(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_M} \\ \mathbf{A}_{MV}(\mathbf{y}(t),\mathbf{x}_U(t)) & \mathbf{A}_M(t) \end{bmatrix},$$

where

$$\mathbf{A}_{MV}(\mathbf{x}_M(t)) = \begin{bmatrix} \mathbf{A}^T_{MV_1}(\mathbf{y}_1(t)) & \cdots & \mathbf{A}^T_{MV_{N_O}}(\mathbf{y}_{N_O}(t)) & \mathbf{A}^T_{MV_{N_O+1}}(\mathbf{x}_{M_{N_O+1}}(t)) & \cdots & \mathbf{A}^T_{MV_N}(\mathbf{x}_{M_N}(t)) \end{bmatrix}^T$$

and $\mathbf{A}_M(t) = \operatorname{diag}(-\mathbf{S}(\boldsymbol{\omega}_m(t)), \cdots, -\mathbf{S}(\boldsymbol{\omega}_m(t)))$. It must be noted that, as $\mathbf{y}(t) = \mathbf{x}_O(t)$, the dynamics matrix can be written as a function of the system output and the non-visible landmarks. From (3.2) it follows that the system may be expressed in a way similar to the usual linear system form. However, it can be seen that the system above is nonlinear, as the dynamics matrix depends on the landmarks that constitute the state and the output.

3.2.2 Problem statement

The problem addressed in this chapter is the design of a range-and-bearing SLAM filter in the space of the sensors by means of a filter for system (3.2), providing a sensor-based map and the velocity of the vehicle. Besides the map, which is defined by tridimensional landmarks, the filter shall have as by-product the estimation of the angular measurement bias. The pose of the vehicle is deterministic as its position corresponds to the origin of the body-fixed frame, which also defines its attitude, i.e, the vehicle is always at the origin of the sensor-based map.

3.3 Observability analysis

Observability is of the utmost importance in any filtering problem and it is the main focus of the theoretical work presented in this chapter. This section aims at analysing the dynamical system previously defined, by establishing several observability results, with the ultimate purpose of designing an RB-SLAM filter with stability and convergence guarantees, and follows the steps described in Section 2.2. It is important to notice that, although system (3.2) is inherently nonlinear, discarding the non-visible landmarks $\mathbf{x}_U(t)$ makes it possible to regard the resulting system as linear time-varying. This is done because the non-visible landmarks are clearly unobservable, and it is shown here that they do not influence the observability of the visible landmarks and of the vehicle state.

Consider the new state vector $\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}^T_V(t) & \mathbf{x}^T_O(t) \end{bmatrix}^T$, which does not include the non-visible landmarks, for which the resulting system dynamics can be written as

$$\begin{cases} \dot{\mathbf{z}}(t) = \mathbf{A}(t,\mathbf{y}(t))\mathbf{z}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{z}(t) \end{cases}, \tag{3.3}$$

where

$$\mathbf{A}(t,\mathbf{y}(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \mathbf{A}_{MV_O}(\mathbf{y}(t)) & \mathbf{A}_{M_O}(t) \end{bmatrix}$$

and

$$\mathbf{C} = \begin{bmatrix} \mathbf{0}_{n_O \times n_V} & \mathbf{I}_{n_O} \end{bmatrix}.$$

The submatrices with subscript $(.)_O$ are equal to their subscript-less counterparts but use only the visible landmarks and have the appropriate dimensions. Furthermore, as indicated in Section 3.2, the visible part of the landmark state coincides with the output, and can therefore be substituted, eliminating the dependency of the dynamics matrix on the system state. However, the matrix $\mathbf{A}(t, \mathbf{y}(t))$ depends not only on time but also on the system output, even though the system can be seen as a linear time-varying system for observability analysis purposes as the dependency on the system state is now absent and the system output is known. This property is established in Lemma 2.1. This result will be used throughout this section. However, before proceeding with this analysis the following assumption is introduced.

Assumption 3.1. *Any two detected landmarks are assumed to be different and nonzero, i.e.,* $\mathbf{y}_i(t) \neq \mathbf{y}_j(t)$ *and* $\mathbf{y}_i(t) \neq \mathbf{0}$ *for all* $t \geq t_0$ *and* $i, j \in \mathcal{M}_O$*, where* $\mathbf{y}_i(t) = \mathbf{p}_i(t)$ *is the* i*-th vector component of the output,* $i \in \{1, \ldots, N_O\}$.

It is important to notice that it is physically impossible to have two collinear landmarks, let alone equal, visible at the same time, as the angle of view of the camera is always smaller than 180°, as shown in Figure 3.3(b). Furthermore, a landmark cannot be coincident with the vehicle. All this makes it a very mild assumption.

The analysis starts by investigating the observability of the LTV system (3.3) and then by showing that the original nonlinear system (3.2) is equivalent to the LTV. In short, the results of this section explore the observability and uniform complete observability of the LTV system and extend the observability result to the original nonlinear system. It is found that the geometric properties of the set of landmark observations in time drive the observability of both systems, in the sense that the complete set of observations of all time must contain sensor-based landmarks that define a plane, either naturally or through the motion of the vehicle. The following theorem states the sufficient and necessary conditions for the observability of system (3.3).

Theorem 3.1 (Observability of the RB-SLAM LTV system)**.** *Consider system* (3.3) *and let* $\mathcal{T} := [t_0, t_f]$*. The system is observable on* $\mathcal{T}$ *in the sense that, given the system output, the initial condition is uniquely defined, if and only if there exist* $\{t_1, t_2, t_3\} \in \mathcal{T}$ *such that at least one of these conditions holds:*

(i) there are, at least, three visible landmarks at the same time t_1 *that define a plane;*

(ii) there exist two visible landmarks in the interval $[t_1, t_2]$ *such that at least one of the landmark sets* $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_2(t_2)\}$ *and* $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_1(t_2)\}$ *defines a plane; or*

(iii) there is a visible time-varying landmark whose coordinates, $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$*, define a plane.*

Proof. The proof follows by transforming the system in analysis by means of a Lyapunov transformation (see Definition A.1 in Appendix A), and then proving that the observability Gramian of the transformed system is non-singular in the conditions of the theorem, which, as Lemma 2.1 states, implies the observability of the transformed system. A Lyapunov transformation preserves the observability properties of a system, hence it suffices to prove that the new, transformed system is observable. This approach has been used successfully in the past, see for example [BSO11b] and [BSO10].

Let $\mathbf{T}(t)$ be a Lyapunov transformation such that

$$\chi(t) = \mathbf{T}(t)\,\mathbf{z}(t), \tag{3.4}$$

where $\mathbf{T}(t) = \operatorname{diag}\left(\mathbf{I}_{n_V}, \mathbf{R}_m(t), \cdots, \mathbf{R}_m(t)\right)$ and the rotation matrix $\mathbf{R}_m(t) \in \mathrm{SO}(3)$ satisfies the differential equation $\dot{\mathbf{R}}_m(t) = \mathbf{R}_m(t)\mathbf{S}(\omega_m(t))$. Please note that $\mathbf{R}_m(t) \neq \mathbf{R}(t)$ in general. The computation of the new system dynamics and output is simple, yielding

$$\begin{cases} \dot{\chi}(t) = \boldsymbol{\mathcal{A}}(t, \mathbf{y}(t))\chi(t) \\ \mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\,\chi(t) \end{cases}. \tag{3.5}$$

The dynamics matrix is given by

$$\boldsymbol{\mathcal{A}}(t, \mathbf{y}(t)) = \begin{bmatrix} \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \boldsymbol{\mathcal{A}}_{MV}(t, \mathbf{y}(t)) & \mathbf{0}_{n_O} \end{bmatrix},$$

where

$$\boldsymbol{\mathcal{A}}_{MV}(t, \mathbf{y}(t)) = \begin{bmatrix} -\mathbf{R}_m(t) & -\mathbf{R}_m(t)\mathbf{S}(\mathbf{y}_1(t)) \\ \vdots & \vdots \\ -\mathbf{R}_m(t) & -\mathbf{R}_m(t)\mathbf{S}\left(\mathbf{y}_{N_O}(t)\right) \end{bmatrix}.$$

The output matrix is simply

$$\boldsymbol{\mathcal{C}}(t) = \begin{bmatrix} \mathbf{0}_{n_O \times n_V} & \operatorname{diag}\left(\mathbf{R}_m^T(t), \cdots, \mathbf{R}_m^T(t)\right) \end{bmatrix}.$$

Before proceeding to compute the observability Gramian associated with the transformed system (3.5), it is necessary to know its transition matrix. Computing $\chi(t)$ as a function of $\chi(t_0)$ by solving $\boldsymbol{\phi}(t, t_0)\chi(t_0) = \chi(t_0) + \int_{t_0}^t \boldsymbol{\mathcal{A}}(\tau, y(\tau))\chi(\tau)\mathrm{d}\tau$, or using the Peano-Baker series, yields

$$\boldsymbol{\phi}(t, t_0) = \begin{bmatrix} \mathbf{I}_{n_V} & \mathbf{0}_{n_V \times n_O} \\ \boldsymbol{\phi}_{MV}(t, t_0) & \mathbf{I}_{n_O} \end{bmatrix},$$

where

$$\boldsymbol{\phi}_{MV}(t, t_0) = \begin{bmatrix} -\int_{t_0}^t \mathbf{R}_m(\sigma)\mathrm{d}\sigma & -\int_{t_0}^t \mathbf{R}_m(\sigma)\mathbf{S}(\mathbf{p}_1(\sigma))\,\mathrm{d}\sigma \\ \vdots & \vdots \\ -\int_{t_0}^t \mathbf{R}_m(\sigma)\mathrm{d}\sigma & -\int_{t_0}^t \mathbf{R}_m(\sigma)\mathbf{S}\left(\mathbf{p}_{N_O}(\sigma)\right)\mathrm{d}\sigma \end{bmatrix}.$$

It is noted that $\mathbf{y}_i(t)$ was replaced by $\mathbf{p}_i(t)$ as it is the same for all $i \in \mathcal{M}_O$.

Recall now the observability Gramian, defined in (2.3). If $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is invertible, then the system (3.5) is observable, in the sense that given the system input and output, the initial condition $\chi(t_0)$ is uniquely defined. The next step is to prove, by contraposition, that this is the case, i.e, by assuming that the system is unobservable, which implies that the observability

Gramian on $\mathcal{T}$ is singular, and showing that the conditions of the theorem do not hold under that hypothesis. Suppose then that $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is not invertible, which means that there exists a unit vector

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_v^T & \mathbf{c}_b^T & \mathbf{c}_1^T & \cdots & \mathbf{c}_{N_O}^T \end{bmatrix}^T \in \mathbb{R}^{(n_V+n_O)},$$

such that,

$$\mathbf{c}^T \boldsymbol{\mathcal{W}}(t_0, t_f)\mathbf{c} = 0. \tag{3.6}$$

Expanding (3.6) gives

$$\mathbf{c}^T \boldsymbol{\mathcal{W}}(t_0, t_f)\mathbf{c} = \int_{t_0}^{t_f} \|\mathbf{f}(\tau, t_0)\|^2 \,\mathrm{d}\tau, \tag{3.7}$$

where $\mathbf{f}(\tau, t_0) := \operatorname{diag}(\mathbf{R}_m(\tau), \ldots, \mathbf{R}_m(\tau))\boldsymbol{\mathcal{C}}(\tau)\boldsymbol{\phi}(\tau, t_0)\mathbf{c}$. The evaluation of

$$\mathbf{f}(\tau, t_0) = \begin{bmatrix} \boldsymbol{\phi}_{MV}(\tau, t_0) & \mathbf{I}_{n_O} \end{bmatrix}\mathbf{c}$$

and its derivative yields

$$\mathbf{f}(\tau, t_0) = \begin{bmatrix} \mathbf{c}_1 - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{c}_v \mathrm{d}\sigma - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{S}(\mathbf{p}_1(\sigma))\,\mathbf{c}_b \mathrm{d}\sigma \\ \vdots \\ \mathbf{c}_{N_O} - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{c}_v \mathrm{d}\sigma - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{S}\big(\mathbf{p}_{N_O}(\sigma)\big)\,\mathbf{c}_b \mathrm{d}\sigma \end{bmatrix} \tag{3.8}$$

and

$$\frac{\partial \mathbf{f}(\tau, t_0)}{\partial \tau} = \begin{bmatrix} -\mathbf{R}_m(\tau)\mathbf{c}_v - \mathbf{R}_m(\tau)\mathbf{S}(\mathbf{p}_1(\tau))\,\mathbf{c}_b \\ \vdots \\ -\mathbf{R}_m(\tau)\mathbf{c}_v - \mathbf{R}_m(\tau)\mathbf{S}\big(\mathbf{p}_{N_O}(\tau)\big)\,\mathbf{c}_b \end{bmatrix}, \tag{3.9}$$

respectively. In order for (3.6) to be true, both quantities must be zero for all $\tau \in \mathcal{T}$. For the first case, this restriction immediately leads to $\mathbf{c}_i = 0$ for all $i \in \mathcal{M}_O$, when $\mathbf{f}(\tau, t_0) = 0$ is evaluated at $\tau = t_0$. Equating the first time derivative in (3.9) to zero implies that

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}(\mathbf{p}_1(\tau)) \\ \vdots & \vdots \\ \mathbf{I}_3 & \mathbf{S}\big(\mathbf{p}_{N_O}(\tau)\big) \end{bmatrix} \begin{bmatrix} \mathbf{c}_v \\ \mathbf{c}_b \end{bmatrix} = \mathbf{0}, \quad \forall \tau \in \mathcal{T}. \tag{3.10}$$

Thus, to prove the sufficiency of the conditions of the theorem it remains to show that those conditions do not hold. For that purpose, consider the situation where there are three visible landmarks $\mathbf{p}_i(t_1)$, $i \in \{1, 2, 3\}$. In this case (3.10) can be rewritten as

$$\begin{bmatrix} \mathbf{I}_3 & \mathbf{S}(\mathbf{p}_1(t_1)) \\ \mathbf{0}_3 & \mathbf{S}(\mathbf{p}_2(t_1) - \mathbf{p}_1(t_1)) \\ \mathbf{0}_3 & \mathbf{S}(\mathbf{p}_3(t_1) - \mathbf{p}_1(t_1)) \end{bmatrix} \begin{bmatrix} \mathbf{c}_v \\ \mathbf{c}_b \end{bmatrix} = \mathbf{0}. \tag{3.11}$$

From this, it is simple to find either that $\mathbf{c}_b = 0$ or that all three landmarks form a line. The first case implies that $\mathbf{c}_v$ is also zero, contradicting the hypothesis that $\mathbf{c}$ is a unit vector. Thus, all the landmarks must form a line and condition 3.1.(i) cannot hold.

In the case where any of the remaining conditions applies, an equation similar to (3.11) may be constructed, this time with the sets $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_2(t_2)\}$ or $\{\mathbf{p}_1(t_1), \mathbf{p}_2(t_1), \mathbf{p}_1(t_2)\}$, for condition 3.1.(ii) and $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$ for condition 3.1.(iii). Hence, if the observability Gramian is singular, none of the hypothesis of the theorem can hold, which means that, if at least one of the conditions of Theorem 3.1 holds, then $\mathcal{W}(t_0, t_f)$ is invertible on $\mathcal{T}$, and, using Lemma 2.1, it follows that (3.5) is observable. Moreover, as the Lyapunov transformation (3.4) preserves observability, the system (3.3) is also observable, thus concluding the proof of sufficiency.

Proving the necessity of the conditions of the theorem for the observability follows naturally by contraposition, i.e., by assuming that neither 3.1.(i), 3.1.(ii), or 3.1.(iii) hold and showing that system (3.5) is not observable. In this situation, there is no possible combination of landmarks in the time interval $\mathcal{T}$ that defines a plane, which is equivalent to say that either there are no visible landmarks or all the visible ones are static or move in the same line. Hence, all the N_O observed landmarks must respect

$$\mathbf{p}_i(t) = p_{d_i}(t)\mathbf{d} + \mathbf{p}_i(t_0), \tag{3.12}$$

where $\mathbf{d} \in \mathbb{R}^3$ defines the line, $p_{d_i}(t) \in \mathbb{R}$, and $i \in \mathcal{M}_O$. The static case is trivially provided by a constant $p_{d_i}(t)$ for all t. The proof starts by computing the explicit time evolution of the output of the transformed system (3.5), which is given by $\mathbf{y}(t) = \mathcal{C}(t)\boldsymbol{\phi}(t, t_0)\boldsymbol{\chi}(t_0)$, whose i-th element is

$$\mathbf{y}_i(t) = \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_i(t_0) + \mathbf{R}_m^T(t)\int_{t_0}^{t} \mathbf{R}_m(\sigma)\Big(\mathbf{v}(t_0) - \mathbf{S}(\mathbf{b}(t_0))\,\mathbf{p}_i(\sigma)\Big)\mathrm{d}\sigma. \tag{3.13}$$

Substituting the negation of the theorem conditions presented in (3.12) in this expression yields

$$\mathbf{R}_m(t)\mathbf{y}_i(t) = \mathbf{R}_m(t_0)\mathbf{p}_1(t_0) - \int_{t_0}^{t} \mathbf{R}_m(\sigma)\Big[\mathbf{v}(t_0) - \mathbf{S}(\mathbf{b}(t_0))\big(p_{d_i}(\sigma)\mathbf{d} + \mathbf{p}_1(t_0)\big)\Big]\mathrm{d}\sigma,$$

for all $i \in \mathcal{M}_O$. If, for the same input, there exist two different initial conditions that lead to the same output at all times, then the system is not observable. With that purpose in mind, let

$$\bar{\mathbf{z}}_0 := \begin{bmatrix}\bar{\mathbf{v}}_0^T & \bar{\mathbf{b}}_0^T & \mathbf{p}_1^T(t_0) & \cdots & \mathbf{p}_{N_O}^T(t_0)\end{bmatrix}^T, \tag{3.14}$$

and

$$\tilde{\mathbf{z}}_0 := \begin{bmatrix}\tilde{\mathbf{v}}_0^T & \tilde{\mathbf{b}}_0^T & \mathbf{p}_1^T(t_0) & \cdots & \mathbf{p}_{N_O}^T(t_0)\end{bmatrix}^T \tag{3.15}$$

be two different initial conditions for system (3.5), such that $\bar{\mathbf{b}}_0 = a\mathbf{d}$, $\bar{\mathbf{v}}_0 = \mathbf{S}\big(\bar{\mathbf{b}}_0\big)\mathbf{p}_1(t_0)$, $\tilde{\mathbf{b}}_0 = b\mathbf{d}$, and $\tilde{\mathbf{v}}_0 = \mathbf{S}\big(\tilde{\mathbf{b}}_0\big)\mathbf{p}_1(t_0)$, with $a \neq b$ and $a, b \in \mathbb{R}\backslash\{0\}$. It is possible to see that given either initial states (3.14) and (3.15) the output of the system is the same and given by

$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_1(t_0) \\ \vdots \\ \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{p}_{N_O}(t_0)\end{bmatrix}.$$

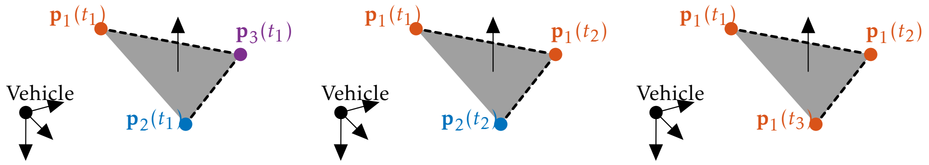


Figure 3.2: Geometrical interpretation of the observability conditions for range-and-bearing SLAM in 3-D. Three landmarks form a plane in one observation (left), two observations of two landmarks form a plane (center) or three observations of one landmark form a plane (right).

Since the two initial states are different and lead to the same output for all $t \in \mathcal{T}$, the system is rendered unobservable if the conditions of the theorem are not satisfied. Hence, if the system (3.5) is observable, at least one of the conditions must hold, which, combined with the fact that the Lyapunov transformation preserves the observability properties, thus concludes the proof of necessity. ☐

Given the sufficient and necessary conditions for observability (see Figure 3.2), this section aims at designing a filter for the nonlinear system (3.2), with globally exponentially stable error dynamics. However, further results are necessary to achieve this goal. The following theorem addresses the equivalence between the state of the nonlinear system (3.3), regarded as LTV, and that of the nominal nonlinear system (3.2), when the non-visible landmarks are not considered.

Theorem 3.2 (Equivalence of the RB-SLAM LTV and nonlinear systems)**.** *Consider that the conditions of Theorem 3.1 hold. Then,*

(i) the initial state of the nonlinear system (3.2), discarding the non-visible landmarks, is uniquely determined, and is the same of the nonlinear system (3.3), regarded as LTV;

(ii) a state observer with uniformly globally exponentially stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, with uniformly globally exponentially stable error dynamics.

Proof. Consider the transformed system (3.5), whose state, and therefore initial condition, is related with the state of the nonlinear system (3.3), regarded as LTV. The proof follows with the transformed system for simplicity of analysis.

Let the initial condition for this system be given by

$$\bar{\chi}(t_0) = \begin{bmatrix} \bar{\mathbf{v}}^T(t_0) & \bar{\mathbf{b}}_\omega^T(t_0) & \bar{\chi}_1^T(t_0) & \cdots & \bar{\chi}_{N_O}^T(t_0) \end{bmatrix}^T,$$

where $\bar{\chi}_i(t) = \mathbf{R}_m(t)\mathbf{y}_i(t)$ for all $i \in \mathcal{M}_O$, resulting from $\mathbf{y}(t) = \mathcal{C}(t)\chi(t)$. It is noted that $\chi(t) = \phi(t,t_0)\bar{\chi}(t_0)$, which then yields $\mathbf{y}(t) = \mathcal{C}(t)\phi(t,t_0)\bar{\chi}(t_0)$. Recall (3.13) and that the i-th component of the output of the system is given by

$$\mathbf{y}_i(t) = \mathbf{R}_m^T(t)\mathbf{R}_m(t_0)\mathbf{y}_i(t_0) - \mathbf{R}_m^T(t)\int_{t_0}^{t} \mathbf{R}_m(\sigma)\Big(\bar{\mathbf{v}}(t_0) + \mathbf{S}(\mathbf{y}_i(\sigma))\bar{\mathbf{b}}_\omega(t_0)\Big)d\sigma,$$

where the relation between $\mathbf{y}_i(t)$ and $\bar{\chi}_i(t_0)$ was used. The next steps include left multiplying this expression by $\mathbf{R}_m(t)$, differentiating, and further simplifying by left multiplying both sides by $\mathbf{R}_m(t)^T$, which yields the first derivative of $\mathbf{y}_i(t)$,

$$\dot{\mathbf{y}}_i(t) = -\bar{\mathbf{v}}(t_0) - \mathbf{S}(\mathbf{y}_i(t))\big(\bar{\mathbf{b}}_\omega(t_0) - \boldsymbol{\omega}_m(t)\big). \tag{3.16}$$

Now consider the nonlinear system (3.2). Let the initial condition of this system be given by

$$\mathbf{x}(t_0) = \begin{bmatrix} \mathbf{v}^T(t_0) & \mathbf{b}_\omega^T(t_0) & \mathbf{p}_1^T(t_0) & \cdots & \mathbf{p}_N^T(t_0) \end{bmatrix}^T,$$

where $\mathbf{p}_i(t_0) = \mathbf{y}_i(t_0)$ for all $\mathbf{p}_i(t_0) \in \mathcal{M}_O$. The output of the system is related to the state by

$$\mathbf{y}(t) = \mathbf{C}\int_{t_0}^{t} \mathbf{A}(\mathbf{x}(\sigma),\sigma)\mathbf{x}(\sigma)\mathrm{d}\sigma + \mathbf{C}\mathbf{x}(t_0),$$

which yields

$$\mathbf{y}_i(t) = \mathbf{y}_i(t_0) - \int_{t_0}^{t} \Big(\mathbf{v}(t_0) + \mathbf{S}\big(\mathbf{p}_i(\sigma)\big)(\mathbf{b}_\omega(t_0) - \boldsymbol{\omega}_m(\sigma))\Big)\mathrm{d}\sigma, \tag{3.17}$$

with first time derivative given by

$$\dot{\mathbf{y}}_i(t) = -\mathbf{v}(t_0) - \mathbf{S}\big(\mathbf{p}_i(t)\big)(\mathbf{b}_\omega(t_0) - \boldsymbol{\omega}_m(t)). \tag{3.18}$$

Comparison of (3.16) with (3.18) after substituting $\mathbf{p}_i(t)$ by $\mathbf{y}_i(t)$ yields

$$\mathbf{0} = (\bar{\mathbf{v}}(t_0) - \mathbf{v}(t_0)) - \mathbf{S}(\mathbf{y}_i(t))\big(\bar{\mathbf{b}}_\omega(t_0) - \mathbf{b}_\omega(t_0)\big)$$

for all t in $\mathcal{T}$ and $i \in \mathcal{M}_O$. When the conditions of Theorem 3.1 hold, this system yields $\bar{\mathbf{v}}(t_0) = \mathbf{v}(t_0)$ and $\bar{\mathbf{b}}_\omega(t_0) = \mathbf{b}_\omega(t_0)$ by a similar reasoning to that which was used to prove the sufficiency of those conditions to ensure the observability of the system. As mentioned before, the initial condition $\bar{\chi}(t_0)$ is related to that of the nonlinear system (3.3) by the Lyapunov transformation $\mathbf{T}(t)$. Hence, under the conditions of Theorem 3.1, the initial state of the nonlinear system (3.3), regarded as LTV, and the initial state of the nonlinear system (3.2), discarding the non-visible landmarks, are the same and uniquely defined.

The first part of the theorem, now proven, gives insight for the proof of the second part. An observer designed for a LTV system with globally exponentially stable error dynamics has an estimation error convergent to zero, implying that the estimates exponentially tend to the true state. Therefore, as the true state of the nonlinear system and the state of the LTV system are shown to be the same under the conditions of the theorem, the estimation error of the same observer applied to the nonlinear system converges to zero exponentially fast. □

Theorem 3.2 established sufficient conditions for the observability of the nominal nonlinear system by relating it to the LTV system (3.3). This can be complemented with the following theorem to ensure a stronger result, by showing that those conditions are not only sufficient but in fact necessary.

Theorem 3.3 (Observability of the RB-SLAM nonlinear system)**.** *The nonlinear system* (3.3)*, discarding the non-visible landmarks, is observable if and only if the conditions of Theorem* 3.1 *hold.*

Proof. Theorems 3.1 and 3.2 establish the ground for this proof, as the sufficiency part comes directly from those results. For the necessity part, consider the output of the nonlinear system expressed in terms of time, as presented in (3.17). The proof is made by substituting the negation of the conditions of the theorem in the output equation and showing that two different initial conditions lead to the same output, following very similar terms to that of the necessity part of Theorem 3.1. Recall that, if the conditions of the theorem do not hold, then all the observed landmarks must respect (3.12). Replacing this in the output of the nonlinear system as defined in (3.17) yields

$$\mathbf{y}_i(t) = \mathbf{p}_i(t_0) - \int_{t_0}^{t} \Big(\mathbf{v}(t_0) + \mathbf{S}\big(p_{d_i}(\sigma)\mathbf{d} + \mathbf{p}_i(t_0)\big)(\mathbf{b}_\omega(t_0) - \boldsymbol{\omega}_m(\sigma))\Big)\mathrm{d}\sigma.$$

Choosing two initial states as in (3.14) and (3.15), yields, for every output and both initial states

$$\mathbf{y}_i(t) = \mathbf{p}_i(t_0) + \int_{t_0}^{t} \mathbf{S}\big(\mathbf{p}_i(t_0)\big)\boldsymbol{\omega}_m(\sigma)\mathrm{d}\sigma,$$

which means that, when the conditions of the theorem do not hold, there are initial states that lead to undistinguishable outputs. Hence, the conditions of the theorem are necessary for the observability of the nonlinear system at hands. □

Given that a GES observer for system (3.3) is an observer for the nominal nonlinear system, the design of such an observer for the LTV system follows. This step requires that the pair $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$ is uniformly completely observable as shown in Theorem 2.2.

The concept of uniform complete observability implies uniform bounds on the observability Gramian in time intervals of length δ. It is a stronger form of observability for linear time-varying systems needed in the steps to design a GES observer. The following theorem states the conditions for this property to be verified, and requires the introduction of a new assumption.

Assumption 3.2. *The position of any detected landmark is upper and lower bounded in norm, i.e.*

$$\underset{t\geq t_0}{\forall}\quad \underset{i\in\mathcal{M}_O}{\forall}\quad \underset{P_m,P_M>0}{\exists}\ :\quad P_m \leq \|\mathbf{y}_i(t)\| \leq P_M.$$

This imposes upper and lower bounds on the norm of the position of each landmark, which may be seen as a demanding assumption. However, given that the vehicle cannot be arbitrarily distant to any landmark and that the range of the RGB-D camera has also a lower and upper limit, these bounds are perfectly justified. Furthermore, they are not used for filter design. With the introduction of this assumption, the analysis can now proceed with the following result.

Theorem 3.4 (Uniform complete observability of the RB-SLAM LTV system). *Consider system* (3.3) *regarded as LTV. The pair* $(\mathbf{A}(t,\mathbf{y}(t)),\mathbf{C})$ *is uniformly completely observable if and only if there exist* $\delta > 0$ *and* $\alpha_* > 0$ *such that, for all* $t \geq t_0$*, it is possible to choose* $\{t_1,t_2,t_3\} \in \mathcal{T}_\delta$, $\mathcal{T}_\delta = [t, t+\delta]$*, such that at least one of the following conditions hold:*

(i) there are at least three visible landmarks $\mathbf{p}_1(t)$, $\mathbf{p}_2(t)$ *and* $\mathbf{p}_3(t)$ *such that*

$$\|(\mathbf{p}_1(t_1)-\mathbf{p}_2(t_1))\times(\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1))\| \geq \alpha_*;$$

(ii) there exist two visible landmarks at times t_1, t_2 *such that*

$$\|(\mathbf{p}_1(t_1)-\mathbf{p}_2(t_1))\times\left(\mathbf{p}_j(t_1)-\mathbf{p}_j(t_2)\right)\| \geq \alpha_*$$

for some $j \in \{1,2\}$*; or*

(iii) there exists one visible time-varying landmark such that its coordinates, at three different instants of time $\{t_1,t_2,t_3\}$*, satisfy*

$$\|(\mathbf{p}_1(t_1)-\mathbf{p}_1(t_2))\times(\mathbf{p}_1(t_1)-\mathbf{p}_1(t_3))\| \geq \alpha_*.$$

Proof. Consider again system (3.5) and recall that it is related by a Lyapunov transformation to system (3.3). Thus, the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t,\mathbf{y}(t)),\boldsymbol{\mathcal{C}}(t))$ implies the uniform complete observability of the pair $(\mathbf{A}(t,\mathbf{y}(t)),\mathbf{C})$. Hence, the proof will continue with the transformed system.

Recall the definition of uniform complete observability in (2.4), reproduced here for convenience,

$$\underset{\substack{\delta>0\\ \alpha>0}}{\exists} \quad \underset{t\geq t_0}{\forall} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_x}\\ \|\mathbf{c}\|=1}}{\forall} \quad : \quad \mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} \geq \alpha,$$

meaning that, in contrast with the observability definition used in Theorem 3.1, the Gramian must have uniform bounds at all times. The proof follows by exhaustion, by analysing the quadratic form in the previous expression for all the possible cases of unit vectors $\mathbf{c}$ for all time. It is recalled from (3.7) and (3.8) that $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} = \int_t^{t+\delta}\|\mathbf{f}(\tau,t)\|^2\,\mathrm{d}\tau$, with

$$\|\mathbf{f}(\tau,t)\|^2 = \sum_{i=1}^{N_O}\left\|\mathbf{c}_i - \mathbf{f}_{v_i}(\tau,t)\right\|^2, \tag{3.19}$$

where

$$\mathbf{f}_{v_i}(\tau,t) := \int_t^\tau \mathbf{R}_m(\sigma_i)(\mathbf{c}_v + \mathbf{S}(\mathbf{p}_i(\sigma_i))\mathbf{c}_b))\,\mathrm{d}\sigma_i.$$

In order to be able to proceed, Proposition A.1 on Appendix A is needed. It is possible to see that it applies to $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$, and therefore it states conditions upon which the quadratic form in analysis is lower bounded if (3.19) is lower bounded, uniformly in time. Then, it suffices to prove that $\|\mathbf{f}(\tau, t)\| \geq \alpha_1$ with $\tau \in \mathcal{T}_\delta$ and $\alpha_1 > 0$ for every possible $\mathbf{c}$, provided that the conditions of Proposition A.1 are satisfied uniformly in time.

Consider the case where there exists at least one $i \in \mathcal{M}_O$ such that $\|\mathbf{c}_i\| \geq \alpha_m$ for some $0 < \alpha_m < 1$ and no restriction is imposed on $\mathbf{c}_v$ and $\mathbf{c}_b$. Then,

$$\|\mathbf{f}(t,t)\|^2 = \sum_{j=1}^{N_O} \left\|\mathbf{c}_j\right\|^2 \geq \|\mathbf{c}_i\|^2 \geq \alpha_m^2$$

for all t. Consider now that $\|\mathbf{c}_i\| < \alpha_m$, for all $i \in \{1, \ldots, N_O\}$ and for some $0 < \alpha_m < 1$. For that purpose, let $\mathbf{c}_p$ and $\mathbf{f}_v(\tau, t)$ be the stacking of all the $\mathbf{c}_i$ and $\mathbf{f}_{v_i}(\tau, t)$, respectively, and note that $\mathbf{f}(\tau, t) = \mathbf{c}_p - \mathbf{f}_v(\tau, t)$. Then, it is possible to write

$$\begin{aligned}\|\mathbf{f}(\tau,t)\|^2 &\geq \|\mathbf{f}_v(\tau,t)\|\Big(\|\mathbf{f}_v(\tau,t)\| - 2\|\mathbf{c}_p\|\Big)\\ &\geq \|\mathbf{f}_v(\tau,t)\|\Big(\|\mathbf{f}_v(\tau,t)\| - 2\sqrt{N_O}\alpha_m\Big).\end{aligned}$$

Assuming that $\|\mathbf{f}_v(\tau, t)\|$ is lower bounded at $\tau = t^*$ in $\mathcal{T}_\delta$ by some $\alpha_2 > 0$ so that α_m can be chosen to satisfy $\alpha_m \leq \frac{\alpha_2}{4\sqrt{N_O}}$, it follows that

$$\|\mathbf{f}(t^*,t)\|^2 \geq \frac{1}{2}\alpha_2^2.$$

It is necessary then to show that $\|\mathbf{f}_v(t^*, t)\| \geq \alpha_2 \geq 4\sqrt{N_O}\alpha_m$ under the conditions of the theorem. In this situation $\mathbf{f}_v(\tau, t)$ is zero for $\tau = t$, and therefore Proposition A.1 applies once again. Thus, it is enough to show that the norm of the derivative of $\mathbf{f}_v(\tau, t)$, given by

$$\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|^2 = \sum_{i=1}^{N_O}\left\|\mathbf{c}_v + \mathbf{S}\big(\mathbf{p}_i(\tau)\big)\mathbf{c}_b\right\|^2, \tag{3.20}$$

is lower bounded by some $\alpha_3 > 0$ for a $\tau \in \mathcal{T}_\delta$. Then, there exists an α_2 that bounds $\|\mathbf{f}_v(\tau, t)\|$ below for some $t^* \in \mathcal{T}_\delta$ and α_m can be chosen accordingly. Under the restriction $\|\mathbf{c}_i\| < \alpha_m$ for all $i \in \mathcal{M}_O$, there exist three possibilities for $\mathbf{c}$, depending on $\mathbf{c}_v$ and $\mathbf{c}_b$. The first case is set by $\|\mathbf{c}_v\| \geq \alpha_v$ and $\|\mathbf{c}_b\| < \alpha_b$ for some α_v and α_b in the interval $(0, 1)$. It is possible to write

$$\begin{aligned}\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|^2 &\geq \sum_{i=1}^{N_O}\|\mathbf{c}_v\|\Big(\|\mathbf{c}_v\| - 2\|\mathbf{p}_i(\tau)\|\|\mathbf{c}_b\|\Big)\\ &\geq N_O\alpha_v\left(\alpha_v - 2P_M\alpha_b\right)\\ &\geq \frac{N_O}{2}\alpha_v^2,\end{aligned}$$

where Assumption 3.2 was employed, and α_v was chosen so that it satisfies $\alpha_b \leq \frac{\alpha_v}{4P_M}$. The second case, where $\|\mathbf{c}_v\| < \alpha_v$ for some $0 < \alpha_v < 1$ and $\|\mathbf{c}_b\| \geq \alpha_b$ for some $0 < \alpha_b < 1$, and the third case, where $\|\mathbf{c}_v\| \geq \alpha_v$ and $\|\mathbf{c}_b\| \geq \alpha_b$, can be analysed together. Consider then that $\|\mathbf{c}_b\| \geq \alpha_b$ and that $0 \leq \|\mathbf{c}_v\| \leq 1$, in which case the conditions of the theorem must be addressed separately. The first condition of the theorem states that there are at least three landmarks ($N_O \geq 3$) that are sufficiently away from collinearity, i.e., the plane defined by the vectors that unite each pair of landmarks is well-defined uniformly in time and sufficiently away from degenerating into a line. This can be illustrated by taking the cross product between each vector that defines the plane and $\mathbf{c}_b$, summing the norms of the results, and noting that the lowest possible value for this sum occurs when $\mathbf{c}_b$ is collinear with the largest of these vectors, say, $\mathbf{c}_b = \pm\frac{\|\mathbf{c}_b\|}{\|\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1)\|}(\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1))$. This worst case can be substituted in the conditions of the theorem to yield

$$\|\mathbf{S}\big(\mathbf{p}_1(t_1)-\mathbf{p}_2(t_1)\big)\mathbf{c}_b\| \geq \frac{\alpha_b\alpha_*}{2P_M}$$

where Assumption 3.2 was used to show that $\|\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1)\| \leq 2P_M$. This means that

$$\|\mathbf{S}\big(\mathbf{p}_1(t_1)-\mathbf{p}_2(t_1)\big)\mathbf{c}_b\| + \|\mathbf{S}\big(\mathbf{p}_1(t_1)-\mathbf{p}_3(t_1)\big)\mathbf{c}_b\| \geq \frac{\alpha_b\alpha_*}{2P_M}$$

for any $\mathbf{c}_b$ with norm greater than $0<\alpha_b<1$. It is a matter of algebraic manipulation to obtain

$$\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v-(\mathbf{S}\big(\mathbf{p}_2(t_1)\big)\mathbf{c}_b+\mathbf{c}_v)\| + \|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v-(\mathbf{S}\big(\mathbf{p}_3(t_1)\big)\mathbf{c}_b+\mathbf{c}_v)\| \geq \frac{\alpha_b\alpha_*}{2P_M}$$

from the previous expression, which can be further manipulated using the triangle inequality to yield

$$2\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_2(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_3(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| \geq \frac{\alpha_b\alpha_*}{2P_M}. \tag{3.21}$$

There are two possible conclusions to draw from this inequality depending on the values of $\|\mathbf{S}\big(\mathbf{p}_2(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\|$ and $\|\mathbf{S}\big(\mathbf{p}_3(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\|$. If either is greater than or equal to some $\alpha_{p_1}>0$, (3.20) is lower bounded by α_{p_1} at $\tau=t_1$, as intended. On the other hand, if both are smaller than $\alpha_{p_1}>0$, then (3.21) leads to

$$\begin{aligned}\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|\Bigg|_{\tau=t_1} &\geq \|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| \\ &\geq \frac{\alpha_b\alpha_*}{4P_M}-\alpha_{p_1} \geq \frac{\alpha_b\alpha_*}{8P_M}\end{aligned}$$

if α_{p_1} is chosen so that $\alpha_{p_1} \leq \frac{\alpha_b\alpha_*}{8P_M}$. The second condition can be treated in similar terms. Consider, without loss of generality that the landmark set that is considered is $\{\mathbf{p}_1(t_1),\mathbf{p}_2(t_1),\mathbf{p}_1(t_2)\}$. Using the same reasoning used for the analysis of the first conditions, it is possible to write

$$\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v-(\mathbf{S}\big(\mathbf{p}_2(t_1)\big)\mathbf{c}_b+\mathbf{c}_v)\| + \|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v-(\mathbf{S}\big(\mathbf{p}_1(t_2)\big)\mathbf{c}_b+\mathbf{c}_v)\| \geq \frac{\alpha_b\alpha_*}{2P_M},$$

which can also be manipulated to obtain a lower bound to the sum of the norms, yielding

$$2\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_2(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_1(t_2)\big)\mathbf{c}_b+\mathbf{c}_v\| \geq \frac{\alpha_b\alpha_*}{2P_M}. \tag{3.22}$$

Once again, the analysis depends on the behaviour of $\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\|$ and $\|\mathbf{S}\big(\mathbf{p}_2(\tau)\big)\mathbf{c}_b+\mathbf{c}_v\|$. If either is greater than or equal to some $\alpha_{p_1}>0$, then, for $\tau=t_1$, $\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\| \geq \alpha_{p_1}$. Conversely, if both are smaller than α_{p_1}, then from (3.22) it is possible to write

$$\begin{aligned}\left\|\frac{\partial}{\partial\tau}\mathbf{f}_v(\tau,t)\right\|\Bigg|_{\tau=t_1} &\geq \|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b+\mathbf{c}_v\| \\ &\geq \frac{\alpha_b\alpha_*}{4P_M}-\alpha_{p_1} \geq \frac{\alpha_b\alpha_*}{8P_M},\end{aligned}$$

for the previous choice of α_{p_1}. Finally, for the third condition, the set of landmarks considered is $\{\mathbf{p}_1(t_1), \mathbf{p}_1(t_2), \mathbf{p}_1(t_3)\}$. In the same line of thought followed previously, one knows that

$$\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}\big(\mathbf{p}_1(t_2)\big)\mathbf{c}_b + \mathbf{c}_v)\| + \|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b + \mathbf{c}_v - (\mathbf{S}\big(\mathbf{p}_1(t_3)\big)\mathbf{c}_b + \mathbf{c}_v)\| \geq \frac{\alpha_b \alpha_*}{2P_M}.$$

Applying the triangle inequality to the previous expression as was done when manipulating the similar expressions for the first two conditions yields

$$2\|\mathbf{S}\big(\mathbf{p}_1(t_1)\big)\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_1(t_2)\big)\mathbf{c}_b + \mathbf{c}_v\| + \|\mathbf{S}\big(\mathbf{p}_1(t_3)\big)\mathbf{c}_b + \mathbf{c}_v\| \geq \frac{\alpha_b \alpha_*}{2P_M}. \tag{3.23}$$

If $\|\mathbf{S}\big(\mathbf{p}_1(\tau)\big)\mathbf{c}_b + \mathbf{c}_v\| \geq \alpha_{p_1}$ for $\tau = t_2$, t_3 or both, then (3.20) is bounded by α_{p_1} in one or both of these instants. However, if $\|\mathbf{S}\big(\mathbf{p}_1(\tau)\big)\mathbf{c}_b + \mathbf{c}_v\| < \alpha_{p_1}$ for all $\tau \in \mathcal{T}_\delta$, $\tau \neq t_1$ and nothing is imposed for $\tau = t_1$, (3.23) leads to

$$\left.\left\|\frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t)\right\|^2\right|_{\tau = t_1} \geq \frac{\alpha_b \alpha_*}{4P_M} - \alpha_{p_1} \geq \frac{\alpha_b \alpha_*}{8P_M}$$

where the fact that $\left\|\frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t)\right\| = \|\mathbf{S}\big(\mathbf{p}_1(\tau)\big)\mathbf{c}_b + \mathbf{c}_v\|$ was used, and α_{p_1} was chosen so that $\alpha_{p_1} \leq \frac{\alpha_b \alpha_*}{8P_M}$. In these three conditions, positive lower bounds were found for $\left\|\frac{\partial}{\partial \tau}\mathbf{f}_v(\tau, t)\right\|$, which, from Proposition A.1 implies that $\mathbf{f}_v(\tau, t)$ is also lower bounded and, therefore, $\mathbf{f}(\tau, t)$ has a lower bound too within the possibilities studied. All possible cases are now enumerated, hence the sufficiency of the conditions of Theorem 3.4 for the pair $(\boldsymbol{\mathcal{A}}(t, \mathbf{y}(t)), \boldsymbol{\mathcal{C}}(t))$ to be uniformly completely observable is proved. Because (3.4) is a Lyapunov transformation, the uniform complete observability of the transformed pair implies uniform complete observability of the $(\mathbf{A}(t, \mathbf{y}(t)), \mathbf{C})$, and thus the proof of sufficiency of the conditions of the theorem is concluded.

The proof of the necessity of the conditions of the theorem follows by contraposition. The hypothesis that the conditions of the theorem do not hold is considered, and it is shown that this implies that the pair $(\boldsymbol{\mathcal{A}}(t), \boldsymbol{\mathcal{C}}(t))$ is not uniformly completely observable.

Assume then that the conditions of the theorem do not hold. The negation of these conditions means that all the landmarks in the set $\mathcal{M}_O$ are not sufficiently away from collinearity. Consider the alternative definition for a landmark, given by

$$\mathbf{p}_i(t) = p_{d_i}(t)\mathbf{d} + p_{d_i^{\perp 1}}(t)\mathbf{d}^{\perp 1} + p_{d_i^{\perp 2}}(t)\mathbf{d}^{\perp 2} + \mathbf{p}_1(t_0),$$

where $\{\mathbf{d}, \mathbf{d}^{\perp 1}, \mathbf{d}^{\perp 2}\} \in \mathbb{R}^3$ are orthonormal vectors that span the planes defined by all the combinations of three landmarks in $\mathcal{M}_O$, and $p_{d_i}(t), p_{d_i^{\perp 1}}(t), p_{d_i^{\perp 2}}(t) \in \mathbb{R}^3$. The negated conditions imply that, either there are no landmarks which immediately renders the pair unobservable or, considering, without loss of generality, that $\mathbf{d}$ is the predominant direction of landmarks, there are at least two directions $\mathbf{d}^{\perp 1}$ and $\mathbf{d}^{\perp 2}$ along which the vector $\mathbf{p}_i(t) - \mathbf{p}_1(t_0)$ is bounded and as small as wanted for some $t \geq t_0$, i.e.,

$$\underset{\beta > 0}{\forall} \quad \underset{i \in \mathcal{M}_O}{\forall} \quad \underset{t \geq t_0}{\exists} : \quad |p_{d_i^{\perp 1}}(t)| < \beta \wedge |p_{d_i^{\perp 2}}(t)| < \beta. \tag{3.24}$$

The proof follows by showing that this implies that the pair in analysis is not uniformly completely observable, which can be stated as

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_x}\\ \|\mathbf{c}\|=1}}{\exists} \; : \quad \mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}<\epsilon.$$

For that purpose, the expansion of $\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}$ is considered for some unit $\mathbf{c}$, yielding the following expression that can be obtained by substituting (3.19) into (3.7),

$$\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}=\int_t^{t+\delta}\sum_{i=1}^{N_O}\left\|\mathbf{c}_i-\int_t^{\tau}\mathbf{R}_m(\sigma_i)\Big(\mathbf{c}_v+\mathbf{S}\big(\mathbf{p}_i(\sigma_i)\big)\mathbf{c}_b\Big)\mathrm{d}\sigma_i\right\|^2\mathrm{d}\tau.$$

In the following steps, a particular $\mathbf{c}$ is chosen and the negated condition (3.24) is used. Considering that $\mathbf{c}_i=\mathbf{0}$ for all $i\in\mathcal{M}_O$, and using the Cauchy-Schwartz inequality in the previous expression leads to

$$\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}\leq\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\left\|\mathbf{c}_v-\mathbf{S}(\mathbf{c}_b)\,\mathbf{p}_i(\sigma)\right\|^2\mathrm{d}\sigma\mathrm{d}\tau.$$

Recall that $\mathbf{R}_m(t)$ maintains the norm, which is why it is omitted in the previous expression. Let a be a real constant, and choose $\mathbf{c}_b=a\mathbf{d}$ and $\mathbf{c}_v=a\mathbf{S}(\mathbf{d})\,\mathbf{p}_1(t_0)$ such that $\|\mathbf{c}\|=1$. This yields

$$\begin{aligned}\mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}&\leq a^2\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\left\|\mathbf{S}(\mathbf{d})\Big(p_{d_i^{\perp1}}(\sigma)\mathbf{d}^{\perp1}+p_{d_i^{\perp2}}(\sigma)\mathbf{d}^{\perp2}\Big)\right\|^2\mathrm{d}\sigma\mathrm{d}\tau\\&\leq a^2\sum_{i=1}^{N_O}\int_t^{t+\delta}\int_t^{\tau}\Big(|p_{d_i^{\perp1}}(\sigma)|+|p_{d_i^{\perp2}}(\sigma)|\Big)^2\mathrm{d}\sigma\mathrm{d}\tau\\&\leq a^2N_O\int_t^{t+\delta}\int_t^{\tau}(2\beta)^2\mathrm{d}\sigma\mathrm{d}\tau,\end{aligned}$$

which can be rewritten as

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_x}\\ \|\mathbf{c}\|=1}}{\exists} \; : \quad \mathbf{c}^T\mathcal{W}(t,t+\delta)\mathbf{c}<\epsilon.$$

as $\beta:=\frac{1}{a\delta}\sqrt{\frac{\epsilon}{2N_O}}$. With this step it was shown that the transformed system cannot be uniformly completely observable if the conditions of the theorem do not apply, which, along with the fact that the Lyapunov transformation preserves the observability properties of the system, leads to the necessity of the conditions for the uniform complete observability of the nonlinear system (3.3), regarded as LTV, and thus concluding the proof. ☐

Remark 3.1. *The conditions of Theorem 3.4 together with Assumption 3.2 imply that the vectors defined by the set of observed landmarks must be sufficiently away from collinearity so that the plane defined by them does not degenerate in time. Moreover, conditions 3.4.(ii) and 3.4.(iii) can be seen as persistent excitation conditions.*

The results shown in this section lead towards the design of a state observer, such as the Kalman filter for linear time-varying systems, with globally exponentially stable error dynamics for the nominal nonlinear system (3.2) according to Theorem 2.2. The discrete-time implementation of such a filter is discussed in the following section.

3.4 SLAM filter design

This section addresses the design of the sensor-based 3-D RB-SLAM filter. Following the discussion in subsection 2.3.2, a discrete Kalman filter is designed, considering the sample-based/digital characteristics of both sensors needed for this work: an IMU (or more precisely a triad of rate gyros) and a depth camera (or other tridimensional relative position sensor). Hence, it is important to obtain the discrete-time version of the dynamic system under analysis. Recall the discrete-time system (2.9). In the implementation proposed in this chapter, the dynamics matrix is discretized with a simple forward Euler discretization, yielding, $\mathbf{F}_k := \mathbf{I}_{n_x} + T_s \mathbf{A}_{F_k}(\mathbf{y}_k)$.

The generic description of subsection 2.3.2 serves as the basis for the work detailed in this section, where additional information is presented regarding the update step of the filter as well as state maintenance procedures.

3.4.1 Update step

Every time an RGB-D (colour and depth) image is available from the *Kinect*, an update is triggered in the filter. The RGB two-dimensional frame is processed by an implementation of SURF [BETG08], which detects and provides a 64-dimensional descriptor to points-of-interest on the 2-D pictures of the environment. This algorithm uses a blob detector, which means that the points-of-interest are regions that differ from the vicinity in properties such as brightness or colour, while maintaining some consistency within the region. The resulting features are then matched to a pointcloud built with the depth image. This matching returns a set of observed tridimensional landmarks in cartesian coordinates. One of the main issues in performing SLAM with natural untagged landmarks, or in mapping in general, is the problem of data association: as the landmarks are collected from the environment, there is no way to know *a priori* with certainty if a landmark from the set of measurements is in the current map or if it is being detected for the first time. At this point, some form of landmark association must take place to match the measured data with the existing landmarks in the filter state.

Landmark association: Wrong associations may have a very negative effect on the estimation, so this is a field where a lot of research effort has been put, yielding a multitude of algorithms from the community. There are two ways to ascertain the correspondence between measured and observed landmarks, either by comparison of their positions, usually through computation of the distances between landmarks in different sets, or by comparison of distinctive characteristics. The use of image processing procedures permits the description of the landmarks by labels with a great deal of information, and which are invariant and robust to several transformations. This fact enhances greatly the data association, as in landmark-filled environments spurious associations when using only the coordinates are very common. For this reason, the filter proposed in this chapter employs a greedy search algorithm with mutual exclusion which looks for the landmarks with closest descriptor. Landmark detection provides tens of landmarks that cannot possibly be all incorporated in the filter, and as such there must be a trimming down process. First the scene is divided in quadrants, and the landmarks in each quadrant are organized by the value of the determinant of their Hessian (see [BETG08] for more information on these quantities). For the desired number N_O of landmarks being chosen, it is only a matter of picking from each quadrant in equal parts so that an even distribution of features is achieved. These landmarks constitute the measured set and are then associated with the state landmarks whose predicted coordinates are within the field of view of the camera (see Figure 3.3(b)). Besides the 64-D descriptor that differentiates landmarks and that provides the main information upon which the association decision is based, each SURF feature has a boolean Laplacian, which is used to prune the search tree, by dividing it in two. The association process can then proceed regularly for each search tree (depending on the Laplacian). The candidate pairs of landmarks to associate are chosen according to the lowest absolute difference of descriptors, and the decision is supported upon the Mahalanobis distance of the landmarks themselves, defined as follows.

Definition 3.1 (*Mahalanobis* distance, [NT01])**.** *The Mahalanobis distance, also denoted as normalized innovation square (NIS), is given by*

$$D^2_{ij_k} := \left(\mathbf{y}_{j_k} - \mathbf{H}_{i_k}\hat{\mathbf{x}}_{k+1|k}\right)^T \left(\mathbf{H}_{i_k}\Sigma_k\mathbf{H}^T_{i_k} + \mathbf{\Theta}_{j_k}\right)^{-1} \left(\mathbf{y}_{j_k} - \mathbf{H}_{i_k}\hat{\mathbf{x}}_{k+1|k}\right),$$

where $\mathbf{y}_{i_k}$ represents the i-th observed landmark at time k, $\mathbf{H}_{i_k}$ is the i-th row of the observation matrix at the same instant, i.e., the row associated with the landmark i. Recall that $\mathbf{\Theta}_k$ and Σ_k are the measurement noise and filter covariances, respectively.

The Mahalanobis distance is known to follow a χ^2 distribution with 3 degrees of freedom, as the vectors are tridimensional.

A statistical test based on the 95% percentile is performed to decide whether an association is valid of not. This statistical test takes the form of

$$D^2_{ij_k} < F^{-1}_{\chi^2_3}(0.95)$$

where $F^{-1}_{\chi^2_3}(\alpha)$ is the inverse distribution function of the χ^2 distribution, i.e., $P(\Gamma < F^{-1}_{\chi^2_3}(\alpha)) = \alpha$ for $\Gamma \sim \chi^2_3$. Finally, when a landmark is associated, the measured descriptor is combined with the descriptor of the state landmark with a mean filter. The algorithm described here is summarized in Algorithm 1. It must be remarked that both the landmark detection and association algorithms may be substituted by others, as they are independent from the filtering technique described in this chapter.

Algorithm 1 Landmark detection and association for RB-SLAM

```
Input: set of state landmarks within the field-of-view (FOV) (position, SURF descriptor, Laplacian), set of detected features (position, SURF descriptor, Laplacian, Hessian), state and observation covariance matrices $\Sigma_k$ and $\boldsymbol{\Theta}_k$, number of landmarks to observe $N_O$
Output: set of visible landmarks, innovation vector and covariance $\Sigma_{\nu_k}$
  Divide scene in quadrants.
  Pick the $N_O/4$ features with highest Hessian from each quadrant.
  Separate features and state landmarks by the sign of the Laplacian.
  Search for associations within each of the sets using Algorithm 2.
  Compute innovation vector $\boldsymbol{\nu}_k$ and covariance $\Sigma_{\nu_k}$ according to the association hypothesis and (2.10).
```

Algorithm 2 Search for associations for RB-SLAM

```
Input: subset of state landmarks (position, SURF descriptor), subset of measured features (position, SURF descriptor), state and observation covariance matrices $\Sigma_k$ and $\boldsymbol{\Theta}_k$
Output: association hypothesis
  Compute the matrix of descriptor distances from all state landmarks to all measured features.
  for k = 1 : $N_O$ do
    Find the pair $(i, j)$ with the lowest descriptor distance.
    Compute the Mahalanobis distance $D_{ij}$ of the pair.
    if $D_{ij} < F^{-1}_{\chi^2_3}(0.95)$ then
      Add pair to the association hypothesis.
    else
      Observation $j$ is considered a new landmark.
    end if
    Remove landmark $i$ and observation $j$ from the pool of association.
  end for
```

3.4.2 State maintenance

The number of landmarks and therefore the size of the state, the size of the state covariance, and the size of the search trees for data association can grow unbounded with the mission time. Due to the computational complexity of SLAM algorithms, this limits the operation time and the size of the environment. For this reason, it is usual to introduce state maintenance procedures to clean the state. In this algorithm, each time a measured landmark is not associated with any already existing landmark, it is automatically added to the state which is the mechanism of incorporation of new landmarks to the state. However, this can also lead to the inclusion of spurious measurements that do not repeat themselves in time an are therefore superfluous to the filter. To tackle this issue, and prevent the state from growing indefinitely in size, there is a counter for each time a landmark is observed, as well as a timer. This way it is possible to keep track of how long has passed since the last observation of a specific landmark was observed and how many times it appeared. A landmark is removed from the state if two conditions are met simultaneously: it must not have been observed for more than some time $T_{\max}$ and it must have been observed less than $N_{\min}$ times.

3.5 Experimental results

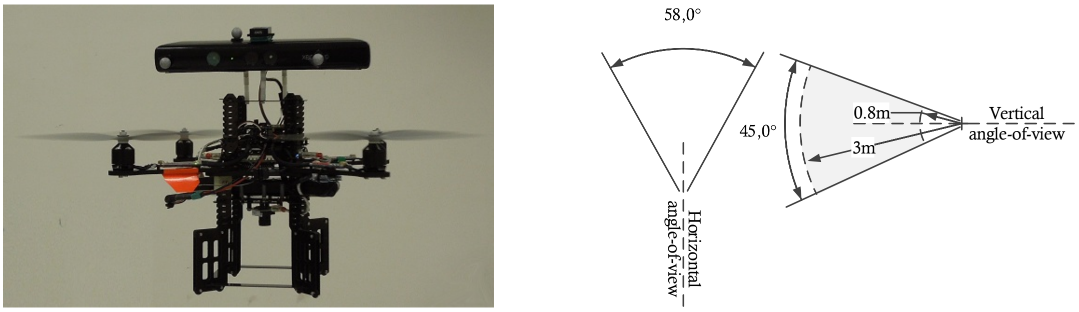
The theoretical results of the previous sections were consolidated by a series of experiments in the Sensor-based Cooperative Robotics Research (SCORE) Lab of the Faculty of Science and Technology of the University of Macau. The experimental setup consists of an *Asctec Pelican* quadrotor (see Figure 3.3(a)), into which was added a *Microstrain 3DM-GX3-25* inertial measurement unit working at 50 Hz and a *Microsoft Kinect* camera, at 10 Hz (see Appendix D for further details on the experimental setup.). The experiments consisted in moving the quadrotor inside a 8m×6m room (usable area of 16 m^2) equipped with a *VICON* motion capture system, which provides accurate estimates of the position, attitude, linear and angular velocities of any vehicle placed inside the working area with the correct markers. These estimates are used in this chapter as ground truth for validation of the proposed RB-SLAM algorithm.

This section provides experimental results for the performance and consistency evaluation of the sensor-based RB-SLAM filter in connection with the Earth-fixed trajectory and map (ETM) algorithm. The results in the Earth-fixed frame are obviously dependent on the individual performance of each algorithm. Conversely, the sensor-based filter does not depend on the ETM algorithm. Furthermore, the results of the ETM algorithm are naturally affected by the nonlinearity intrinsic to the problem of translating and rotating a map arbitrarily between coordinate frames, also found in EKF-based SLAM algorithms. By proposing this algorithm, it is argued that it may be possible to obtain a less uncertain Earth-fixed trajectory and landmark map. With this approach, all the landmark association, loop-closing, control, and decision procedures can be made in the sensor-based frame, minimizing the effects of nonlinearities in the consistency of the filter, whereas the Earth-fixed trajectory and map estimation is used only for completeness of results. As there is no ground truth for the sensor-based map (which is where the SLAM problem is posed and solved), the results of the ETM are used for comparison with the ground truth.

In this section, the results of two different experiments are presented: (i) a short path inside the lab, as shown in Figure 3.4, with ground truth available at all times, and (ii) a larger path starting inside the lab but moving outside where no ground truth is available. The parameters of the algorithm used in these experiments are presented in Table 3.1, that includes

Table 3.1: RB-SLAM filter parameters.

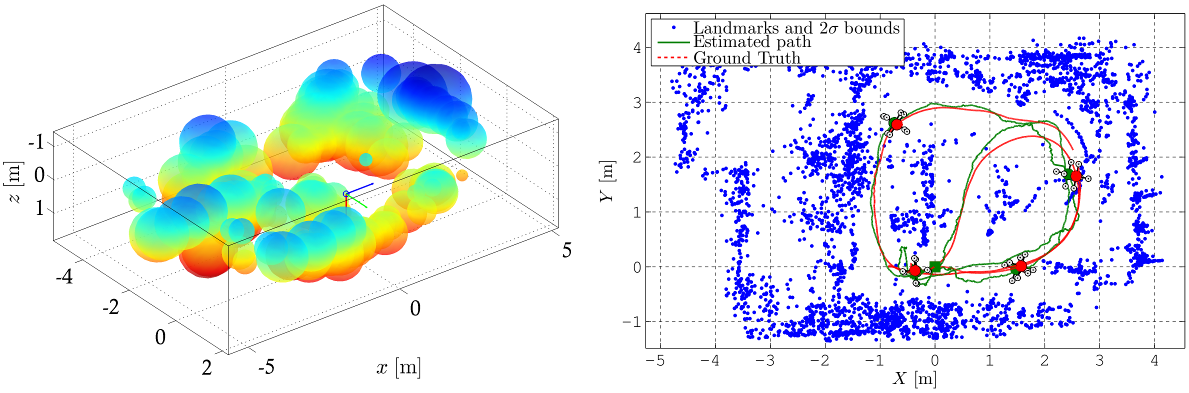
$\mathbf{\Theta}$ $[m^2]$	Ξ_M $[m^2]$	Ξ_v $[(m/s)^2]$	Ξ_{b_ω} $[(^\circ/s)^2]$	desired N_O
$(0.05)^2\mathbf{I}_{n_O}$	$(0.05)^2\mathbf{I}_{n_M}$	$(0.05)^2\mathbf{I}_3$	$(5\times10^{-6})^2\mathbf{I}_3$	15



(a) The AscTec Pelican equipped with a *Microsoft Kinect*, a *Microstrain 3DM-GX3-25* and the *VICON* markers.

(b) The field of view of the *Kinect*.

Figure 3.3: The experimental setup.



(a) Sensor-based map with uncertainty ellipsoids.

(b) Top view of the Earth-fixed map with vehicle trajectory and ground truth.

Figure 3.4: Run #1 – Landmark map of the environment with the estimated and real trajectory.

the observation noise covariance, the process noise covariance, and the desired N_O, which is an indicative parameter that is used when choosing which and how many features in the RGB image to input as measured landmarks in the filter.

3.5.1 Run #1

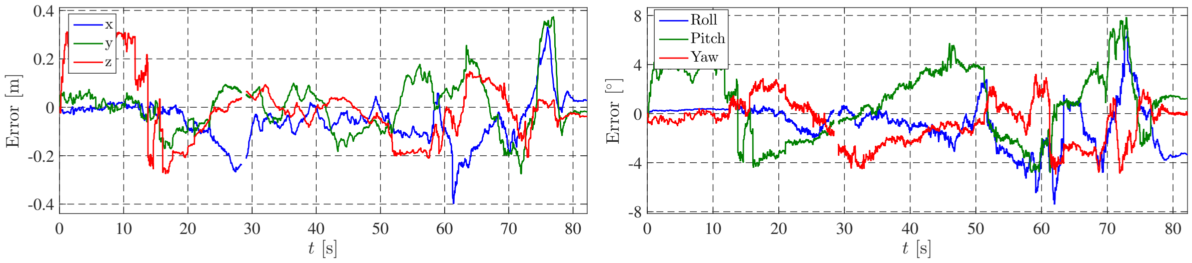
The first run is simply a circular-like lap inside the working area of the lab. The vehicle is driven by hand at constant height and average speed of 0.3 m/s in a 25 m path, which can be seen in Figure 3.4. In Figure 3.4(a) the map in the sensor frame is presented with the respective 95% (or 2σ) uncertainty ellipsoids. One important aspect to retain from this figure is that the uncertainty of the non-visible landmarks increases with time, noting that the most recent landmarks, i.e., those close to the vehicle, have much smaller uncertainty ellipsoids.

In Figure 3.4(b), a top view of the Earth-fixed map is shown along with the estimated trajectory (solid line) and the ground truth trajectory (dashed line) obtained from the *VICON*. The coloured squares, that coincide by construction, and triangles indicate the start and end of the run, respectively. The ellipses are the 2-D projection of the 2σ uncertainty ellipsoids. The small quadrotors represent the pose of the vehicle in several instants, both with ground truth (dashed red) and SLAM estimates (solid green).

These results can be better explained by looking at Figure 3.5, where the estimation error of the Earth-fixed quantities are plotted against time. First, Figure 3.5(a) depicts the evolution of the position estimation error. Note that the vertical estimates are worse than the horizontal ones, which are quite accurate (standard deviation of 0.08 m against 0.14 m). Given that there was no motion in the vertical direction, and that the horizontal angle-of-view of the *Kinect* is greater than the vertical one, these results are not unexpected (see Figure 3.3(b)). This is confirmed in Figure 3.5(b), where the attitude of the vehicle is presented in the form of the Euler angles, estimated with small error, except for the pitch angle (θ) which is off up to 8° (standard deviations of 1.7° for roll and yaw, and 2.8° for pitch). It must be noted that the calibration of the rigid transformation between the camera and the rate-gyros frames was not performed automatically, as in [KS11], but manually. This may explain the errors in the attitude estimation, together with the lack of vertical motion by the vehicle, and the reduced vertical angle-of-view of the *Kinect* that limits the vertical separation of landmarks, and consequently the information extractable therein.

The body-fixed velocity estimation errors and the angular rate measurement bias estimates are shown in Figure 3.6. The velocity estimation error is depicted alongside the 95% uncertainty bounds, and, even though the velocity is modelled to be constant, it follows the velocity accurately (standard deviation of 0.02 m/s in the vertical axis and 0.05 m/s in the horizontal ones). Furthermore, its uncertainty converges while generally maintaining the consistency throughout the run. The measurement bias on the right is obviously presented without ground truth, but its uncertainty can be seen to converge, confirming the results of Section 3.3.

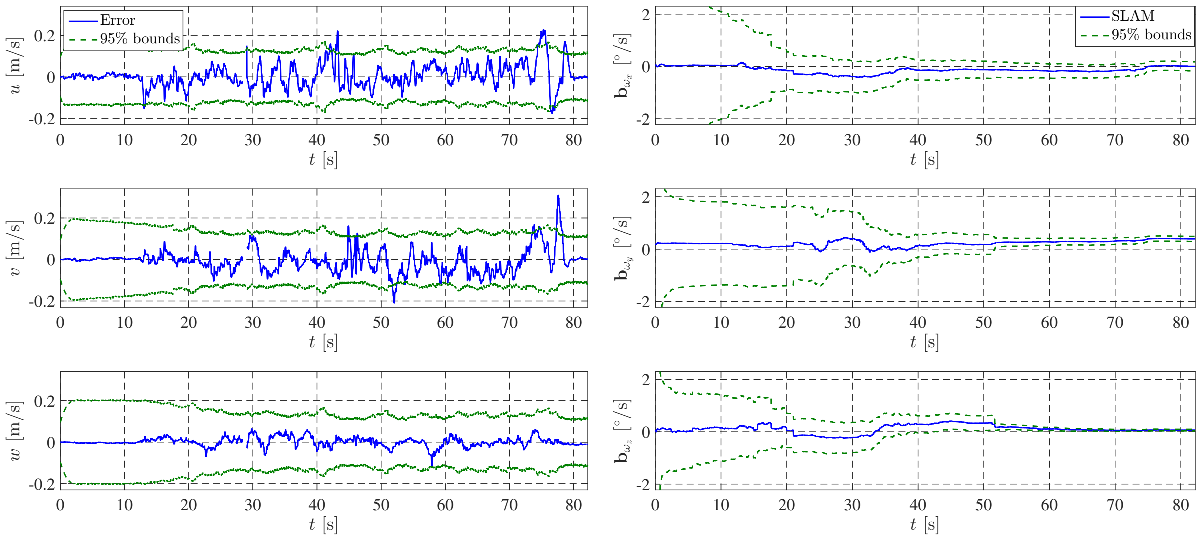
Finally, Figure 3.7 presents information about the sensor-based landmark estimates: the number of landmarks in Figure 3.7(a), the association quality index in Figure 3.7(b), and the standard deviation representing the uncertainty in Figure 3.7(c). The landmark uncertainty plot is of interest because it allows to identify the moments where loop closures occur, at $t \approx 50$ s, as well as the growth rate of the uncertainty of the non-visible landmarks that is smaller than 1 m after a displacement of 12 m. Loop closures occur when the vehicle is exposed to



(a) Position.

(b) Attitude.

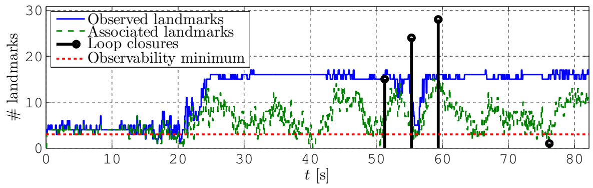
Figure 3.5: Run #1 – Time evolution of the estimation error of the Earth-fixed estimates.



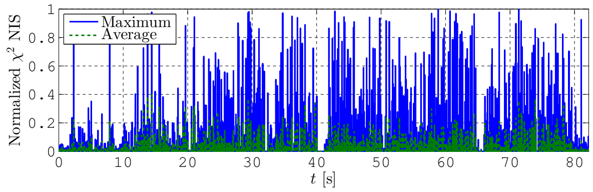
(a) Velocity estimation error.

(b) Angular measurement bias.

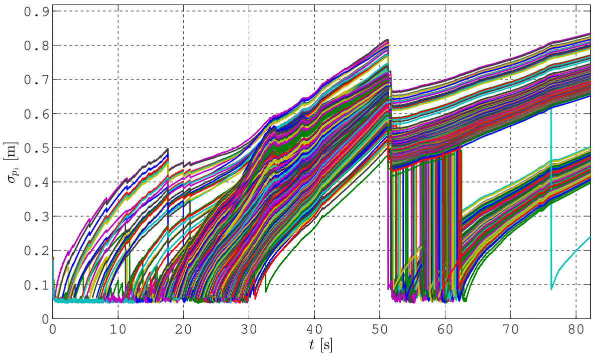
Figure 3.6: Run #1 – Time evolution of the sensor-based estimates with 2σ bounds.



(a) Number of landmarks used in the Kalman filter.



(b) normalized innovation square (NIS) association.



(c) Landmark standard deviation.

Figure 3.7: Run #1 – Time evolution of landmark-related variables.

previously visited terrain, and the old landmarks reappear in its estimated field-of-view being associated with freshly obtained measurements. The solid blue line in Figure 3.7(a) represents the number of measured landmarks in each instant. From the association process results a smaller number of actual associations, presented in dashed green. It can be noticed that the number of landmarks is almost always greater than three, which means that at least one of the observability conditions derived in Section 3.3 is fulfilled. At last, Figure 3.7(b) presents the average and maximum normalized innovation square (NIS) value for each set of associations, given by

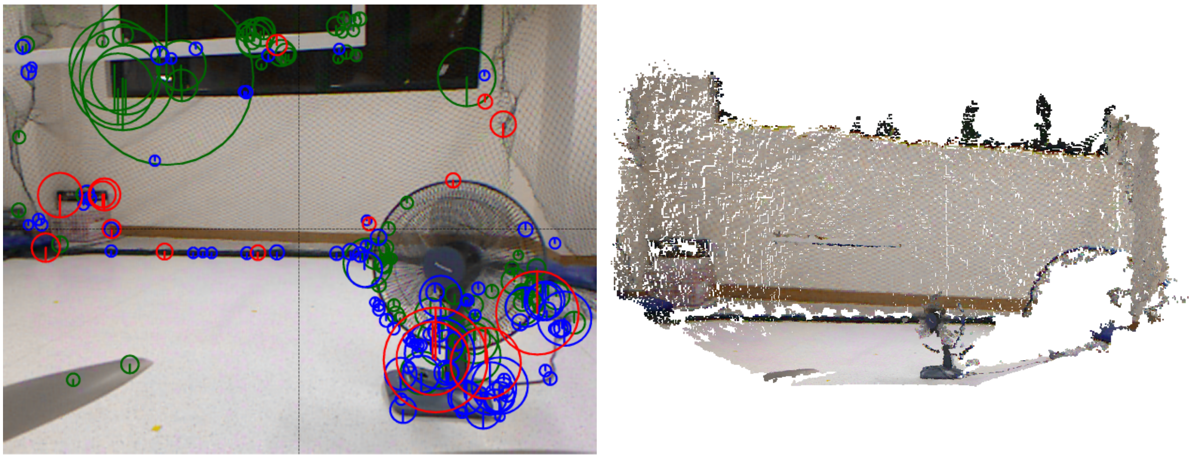
$$\mathrm{NIS}_{i_k} = \boldsymbol{\nu}_{i_k}^T \Sigma_{\nu_{i_k}} \boldsymbol{\nu}_{i_k}.$$

This value is employed as an association index, and is known to follow a chi-squared distribution with 3 degrees of freedom for 3-D measurements (see Definition 3.1). It is compared to the 95% percentile threshold to determine if an association is valid or not. In the figure, the average and maximum NIS values are also normalized with the value corresponding to the 95% percentile. It is observed that although the maximum values approach the threshold, the average association is well below it, indicating that the landmark associations and the overall sensor-based algorithm are consistent.

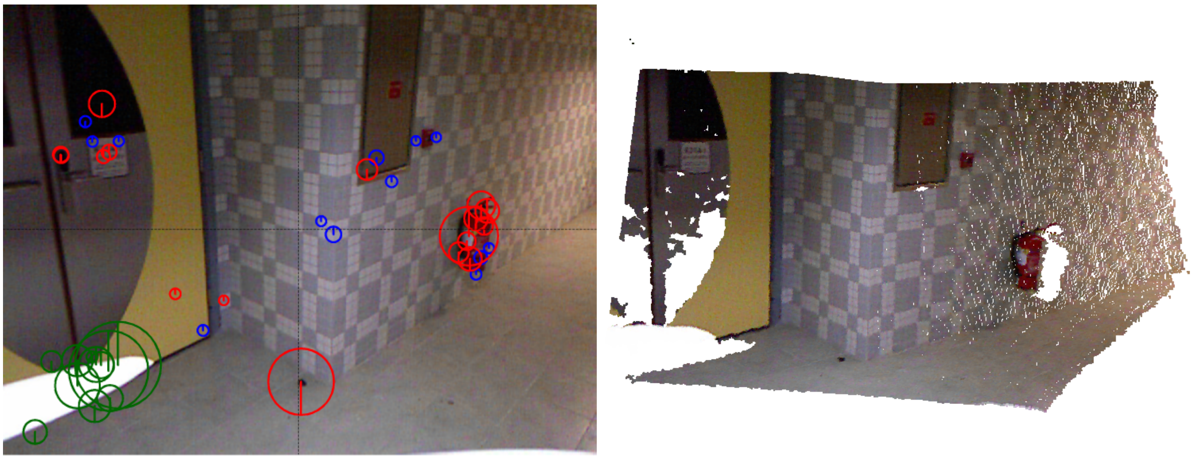
3.5.2 Run #2

In the second run, the vehicle starts inside the usable area of the lab and, after a small lap inside the room, it is hand-driven outside the room into the corridor shown in Figure 3.9 travelling a total of 70 meters at an average speed of 0.4 m/s. The estimated Earth-fixed map and trajectory are presented in Figure 3.9. From the top view depicted in that figure it is possible to recognize the floor plan that is in the background. Please note that the *VICON* system is only available inside the working area of the lab, and that is why ground truth disappears after around 45 s into the run. In this figure (as in Figure 3.4(b)), it is noticeable that a number of landmarks appears inside the lab due to the presence of obstacles not indicated in the blueprint, especially the protection net whose knots are identified as landmarks several times. This is shown in Figure 3.8(a) which depicts the output of the RGB-D camera along with the extracted SURF features in the beginning of the experiment. The green circles represent extracted features that are not matched to the camera pointcloud as they are out of range, the blue circles are the detected features, and the red are used as measurements in the filter. If on one hand the lab has plenty of features, the corridor outside it presents a completely different environment. The most recognizable areas are the ones close to doors, as shown in Figure 3.8(b). This can be confirmed in Figure 3.9 where it can be seen that the greater agglomerations of landmarks occur near doors. In fact, outside those areas, there are situations where the number of detected features is very low (Figure 3.8(c)), which, combined with the limited field-of-view of the camera impairs the performance of the algorithm. On the other hand, the tiled walls of the corridor present an excellent field for feature extraction. However, the environment can become cluttered with features similar to one another (see Figure 3.8(d)), and, although only a reduced set of landmarks from the detected set are actually used as measurements, there can be some difficulty in associating them.

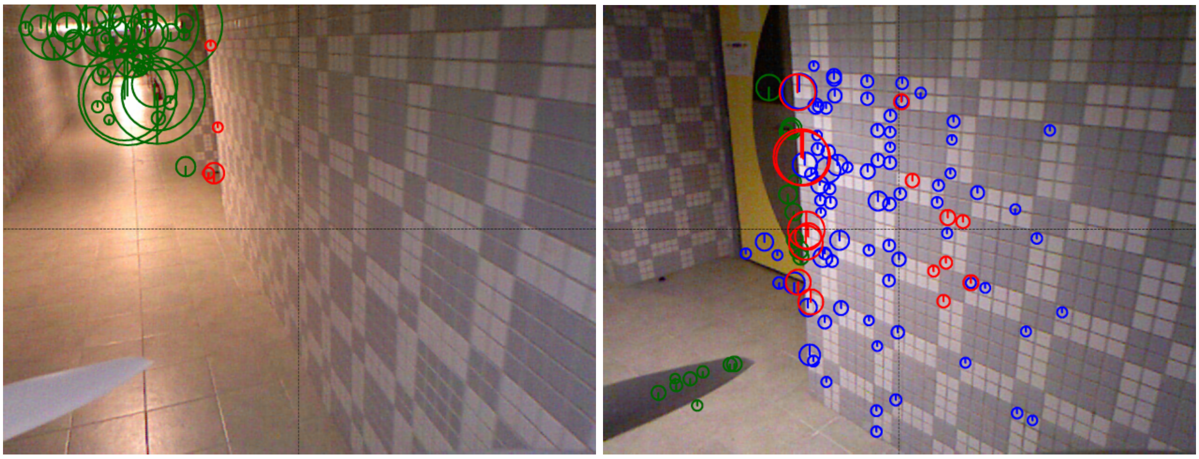
Once again, the horizontal estimates are more accurate than the vertical ones, as can be seen in Figure 3.10(b), where the Earth-fixed estimates are presented in the same fashion as those of Figure 3.5. When compared with the first run, one concludes that the Euler angles are better estimated in this run, although pitch and roll are still erroneous momentarily. The yaw angle is quite accurately estimated once more. Looking at the products of the sensor-based filter, it is easily checked that the linear velocity estimates presented in Figure 3.10(a) are consistent with the ground truth.



(a) RGB-D output at the beginning of the experiment.



(b) RGB-D output near a corridor door.



(c) A zone with few relevant landmarks.

(d) A zone cluttered with similar features.

Figure 3.8: Run #2 – RGB pictures (left) with detected features and coloured pointclouds (right) obtained by the *Kinect* at the beginning of the experiment (top) and while in the corridor (bottom). The size of the circles indicates the magnitude of the Hessian of each feature and the direction of the vertical radius represents the Laplacian (positive-up, negative-down). Green circles represent features extracted but not matched in the point-cloud, blue circles are detected landmarks, and red circles are used as measurements.

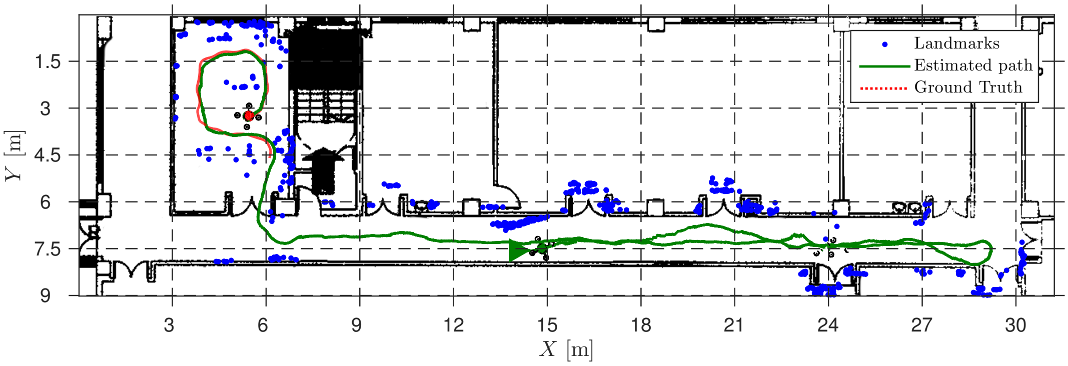
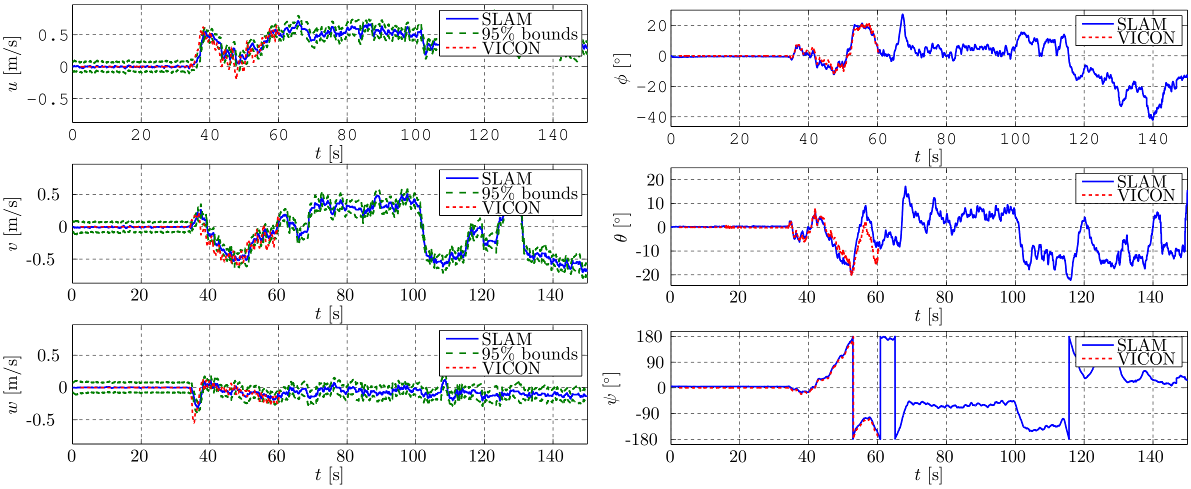


Figure 3.9: Run #2 – Top view of the Earth-fixed map with vehicle trajectory and ground truth against the floor blueprint.



(a) Velocity with 2σ bounds.

(b) Attitude.

Figure 3.10: Run #2 – Time evolution of the attitude and body-fixed linear velocity estimates against the ground truth.

In summary, these experiments were intended to allow the practical validation of the algorithm proposed in this chapter, assess its performance on real conditions and evaluate two of the main issues on any SLAM algorithm: consistency and convergence of the solution. The comparison with the ground truth enabled the performance evaluation, showing that the sensor-based algorithm performs well. The consistency of the sensor-based SLAM filter is also evaluated by the ability to close a loop after revisiting terrain. Finally, the convergence is demonstrated by the decreasing in the uncertainty of all variables, except when landmarks are not visible, confirming the theoretical results of Section 3.3.

3.6 Conclusions

This chapter presented an algorithm for simultaneous localization and mapping, reporting the design, analysis, and validation stages. The main contribution is the novel sensor-based filter with global stability and convergence guarantees. These stem from the provided formal proofs for the observability of the underlying nonlinear system, through the establishment of both necessary and sufficient conditions. This was accomplished by designing the filtering framework in the space of the sensors, which enabled the construction of a nonlinear system that can be regarded as linear time-varying, thus allowing the use of the well-established Kalman filter theory. The sensor-based approach allows the establishment of purely kinematic systems without the need for any pose representation, as well as the direct use of the measurements in the state, especially as the sensors used in the vehicle directly provide measurements in vehicle-centered coordinates.

The theoretical results were verified in real world experiments in an indoor environment, validating the convergence of the error dynamics of the Kalman filter for LTV systems, as predicted by the observability analysis presented in Section 3.3. The performance and consistency evaluation showed a decreasing uncertainty in every variable except the non-visible landmarks, as well as the production of a consistent map. This was complemented by the use of an early implementation of the ETM algorithm working in conjunction with the sensor-based SLAM filter, as the Earth-fixed estimation is also shown to be performing accurately in comparison with the ground truth data.

4

RANGE-ONLY SIMULTANEOUS LOCALIZATION AND MAPPING

THIS chapter proposes the design, analysis, and validation of a globally exponentially stable filter for three-dimensional (3-D) range-only simultaneous localization and mapping. Following the methodology proposed in Chapter 2 and building on the idea in [BSO11b], constructive observability results can be established, leading naturally to the design of a Kalman filter with globally exponentially stable (GES) error dynamics. These results also provide valuable insight on the motion planning of the vehicle. Experimental results using the setup overviewed in Appendix D demonstrate the good performance of the algorithm and help validate the theoretical results presented. For completeness and to illustrate the necessity of a proper trajectory, simulation data are included as well.

4.1 Introduction

One of the possible applications of range-only simultaneous localization and mapping (RO-SLAM) algorithms are scenarios where a number of sensor nodes were deployed at an unknown location. For example, static sensors with unregistered locations monitoring an industrial facility or the first response deployment of real-time monitoring equipment at an accident site [TGdDO17].

The main contributions of this chapter are the design, analysis, and experimental validation of a 3-D RO-SLAM algorithm that (i) has globally exponentially stable (GES) error dynamics; (ii) resorts to the exact linear and angular motion kinematics; (iii) uses as odometry-like measurements the linear and angular velocities; (iv) solves a nonlinear problem with no linearizations whatsoever; and (v) builds on the well-established linear time-varying Kalman filtering theory.

The constructive observability and convergence results achieved provide physical insight on what kind of trajectories the vehicle must take in order for the RO-SLAM algorithm to be able to perform accurately. These results were validated in real conditions, using a Cricket

[PCB00] sensor network as landmarks and a visual odometry procedure to determine the linear velocity. Furthermore, simulation results are also presented to illustrate the good vertical performance when the trajectory is sufficiently rich, which was not possible to perform in the experiments carried out.

Chapter Structure

The chapter is organized as follows: in Section 4.2, the problem addressed in this chapter is stated and the dynamics of the system to be filtered are presented; the observability analysis of the system is performed in Section 4.3 and filter implementation issues are detailed in Section 4.4. The results of simulation and real experiments are presented in Sections 4.5 and 4.6, respectively, and, finally, Section 4.7 addresses some concluding remarks.

4.2 Problem statement and System dynamics

Consider a vehicle moving in a static world where acoustic and/or electromagnetic beacons are installed at unknown locations. The vehicle is equipped with a sensor suite capable of measuring the linear and angular velocities as well as radio and acoustic signals from the static beacons (see Figure 4.1). The distances to the emmitting beacons can then be computed from the time differences of arrival. This section details the design of a dynamical system as part of a simultaneous localization and mapping filter using only, apart from vehicle motion information, the distance to the beacons placed in the environment.

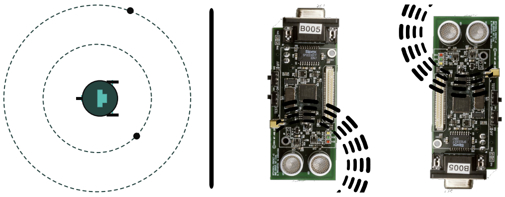


Figure 4.1: In range-only SLAM, the vehicle is equipped with sensors that are capable of measuring the distance to a landmark. For example, acoustic beacons such as the Cricket system. [PCB00]

4.2.1 Problem statement

Recall the generic nonlinear system (2.1), and consider that both the linear and angular velocities are available through sensor measurements. The distances to landmarks, henceforth denominated ranges, are given by $r_i(t) = \|{}^E\mathbf{p}_i(t) - {}^E\mathbf{p}(t)\| = \|\mathbf{p}_i(t)\|$ and are measured as well. Given that the landmarks positions are unknown a priori, the ranges are not sufficient to obtain estimates of the coordinates of the landmarks, raising the need for the knowledge of the motion of the vehicle. That is why the linear velocity is measured.

This information motivates the design of a system whose states are the sensor-based landmarks and the linear velocity, and with outputs that are the ranges to the landmarks and the linear velocity. Thus the generic system (2.1) becomes

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t))\,\mathbf{p}_i(t) - \mathbf{v}(t) \\ \dot{\mathbf{v}}(t) = \mathbf{0} \\ r_j(t) = \|\mathbf{p}_j(t)\| \\ \mathbf{y}_v(t) = \mathbf{v}(t) \end{cases} \tag{4.1}$$

with $i \in \mathcal{M}$ and $j \in \mathcal{M}_O$, and where the first two quantities are the system state and the last two its output. Although the linear velocity in $\{B\}$ is measured, it is included as a state with zero derivative for filtering purposes. Albeit being assumed constant in this deterministic setting, in the filtering framework it can be seen as slowly time-varying, depending on the process and measurement noise covariances.

The problem addressed in this chapter is that of designing a navigation system for a vehicle operating in the environment previously described, by means of a filter for the dynamics in (4.1), assuming noisy measurements. The algorithm consists of a RO-SLAM filter in the space of sensors and, therefore, the pose of the vehicle is deterministic as it simply corresponds to the position and attitude of the body-fixed frame expressed in the same frame.

4.2.2 Augmented system dynamics

The system derived in the previous subsection is clearly nonlinear, as there is a nonlinear relation between the output and the system state. The strategy proposed to avoid this nonlinearity is to augment the system state in order to obtain a linear relation between the system state and output, as it was successfully done in [BSO11b]. The resulting augmented state is

$$\mathbf{x}(t) := \begin{bmatrix} \mathbf{x}_M^T(t) & \mathbf{x}_V^T(t) & \mathbf{x}_R^T(t) \end{bmatrix}^T,$$

where $\mathbf{x}_M(t) \in \mathbb{R}^{n_M}$ is the vector built by stacking all the landmarks present in the landmark set $\mathcal{M}$, both the visible ones, $\mathbf{x}_{M_O}(t) := \{\mathbf{x}_{M_i}(t),\ i \in \mathcal{M}_O\}$, and the invisible ones, $\mathbf{x}_{M_U}(t) := \{\mathbf{x}_{M_i}(t),\ i \in \mathcal{M}_U\}$, $\mathbf{x}_V(t) \in \mathbb{R}^{n_V}$ represents the vehicle state, i.e., the linear velocity of the vehicle in the body-fixed frame, and the vector $\mathbf{x}_R(t) := \begin{bmatrix} \mathbf{x}_{R_O}^T(t) & \mathbf{x}_{R_U}^T(t) \end{bmatrix}^T \in \mathbb{R}^{n_R}$ contains the ranges to all the landmarks in the visible and invisible sets. The following expression summarizes the relations that define the augmented state,

$$\begin{cases} \mathbf{x}_{M_i}(t) := \mathbf{p}_i(t) \\ \mathbf{x}_V(t) := \mathbf{v}(t) \\ x_{R_i}(t) := \|\mathbf{x}_{M_i}(t)\| \end{cases}, \tag{4.2}$$

where $\mathbf{x}_{M_i}(t) \in \mathbb{R}^3$ and $x_{R_i}(t) \in \mathbb{R}$ are part of the full landmark and full range states, respectively, for all $i \in \mathcal{M}$. It is important to notice that, both in the landmark and range states, the first N_O quantities are the visible ones, while the $i \in \{N_O + 1, \ldots, N\}$ are the remaining. The state is chosen this way to simplify the forthcoming analysis, without loss of generality.

The dynamics of the landmark and vehicle states have already been defined in (4.1), hence, the derivative of the range, given by

$$\dot{r}_i(t) = -\frac{1}{r_i(t)}\mathbf{v}^T(t)\mathbf{p}_i(t), \tag{4.3}$$

is needed to derive the full dynamics. Note that, although the system output is now linear, the introduction of the ranges as states has created another nonlinearity in the dynamics. On the other hand, the velocity is directly available as a measurement, as is the distance $r_i(t)$ if the corresponding landmark is visible, i.e., if $i \in \mathcal{M}_O$. Therefore, it is possible to replace the dependence on the state for one on the system output. Observing that the output is given by

$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{y}_v^T(t) & y_{R_1}(t) & \cdots & y_{R_{N_O}}(t)\end{bmatrix}^T,$$

where $\mathbf{y}_v(t) := \mathbf{v}(t)$ and $y_{R_i}(t) = r_i(t)$ for every $i \in \mathcal{M}_O$, it is possible to derive the dynamics of the augmented system, which are given by

$$\begin{cases}\dot{\mathbf{x}}(t) = \mathbf{A}_F(\mathbf{x}_{R_U}(t), \mathbf{y}(t), t)\mathbf{x}(t) \\ \mathbf{y}(t) = \mathbf{C}_F\mathbf{x}(t)\end{cases} \tag{4.4}$$

where

$$\mathbf{A}_F(\mathbf{x}_{R_U}(t), \mathbf{y}(t), t) = \begin{bmatrix}\mathbf{A}_M(t) & \mathbf{A}_{MV} & \mathbf{0}_{n_M\times n_R} \\ \mathbf{0}_{n_V\times n_M} & \mathbf{0}_{n_V} & \mathbf{0}_{n_V\times n_R} \\ \mathbf{A}_{RM}(\mathbf{x}_{R_U}(t), \mathbf{y}(t), t) & \mathbf{0}_{n_R\times n_V} & \mathbf{0}_{n_R}\end{bmatrix},$$

and

$$\mathbf{C}_F = \begin{bmatrix}\mathbf{0}_{3\times n_M} & \mathbf{I}_3 & \mathbf{0}_{3\times N_O} & \mathbf{0}_{3\times N_U} \\ \mathbf{0}_{N_O\times n_M} & \mathbf{0}_{N_O\times 3} & \mathbf{I}_{N_O} & \mathbf{0}_{N_O\times N_U}\end{bmatrix}.$$

The dynamics matrix is composed by three block matrices, namely one that relates the landmark state to itself, given by

$$\mathbf{A}_M(t) := \operatorname{diag}(-\mathbf{S}(\boldsymbol{\omega}(t)), \ldots, -\mathbf{S}(\boldsymbol{\omega}(t))),$$

another that relates the landmark and vehicle states,

$$\mathbf{A}_{MV} := \begin{bmatrix}-\mathbf{I}_3 & \cdots & -\mathbf{I}_3\end{bmatrix}^T,$$

and finally the relation between landmark and range states,

$$\mathbf{A}_{RM}(\mathbf{x}_R(t), \mathbf{y}(t), t) := -\operatorname{diag}\left(\frac{\mathbf{y}_v^T(t)}{y_{R_1}(t)}, \ldots, \frac{\mathbf{y}_v^T(t)}{y_{R_{N_O}}(t)}, \frac{\mathbf{y}_v^T(t)}{x_{R_{N_O+1}}(t)}, \ldots, \frac{\mathbf{y}_v^T(t)}{x_{R_N}(t)}\right).$$

It is clear that the system now derived is still nonlinear, due to the dependence of the dynamics on the velocity and on the ranges. However, as both the first N_O ranges and the linear velocity are measured, the dependence of $\mathbf{A}_{RM}(\mathbf{x}_R(t), \mathbf{y}(t), t)$ in the full range state can be substituted by a dependence on the system output. This is done so that the observability analysis can be performed in a linear fashion, as it will be seen in the next section. Notice also that there are several singularities in the dynamics matrix, if one or more of the ranges becomes null. To prevent that, the following assumption is needed.

Assumption 4.1. *The motion of the vehicle is such that*

$$\underset{i\in\mathcal{M}}{\forall}\quad \underset{t\geq t_0}{\forall}\quad \underset{R_m,R_M>0}{\exists}:\quad R_m < r_i(t) < R_M.$$

Even though this assumption is needed for the dynamics matrix to be well-defined, it is straightforward to see that it is a very mild assumption, as, in practice, the vehicle is never coincident with a landmark nor can it be arbitrarily distant. Moreover, the values of R_m and R_M are not required for the filter design.

One final aspect important to retain is the fact that there is nothing in the system dynamics imposing the state relations expressed by (4.2), particularly the nonlinear relation $x_{R_i}(t) = \|\mathbf{x}_{M_i}(t)\|$. Although they could be established through augmented nonlinear outputs, that would invalidate the option to expand the system state in order to obtain a linear time-varying system. Furthermore, the next section presents a result that demonstrates that, in certain conditions, the dynamics of the system directly imposes these constraints.

4.3 Observability analysis

Before the RO-SLAM filter design can proceed properly, it is necessary to ensure observability of the derived system, and to validate theoretically the procedure of the state augmentation. This section addresses the observability analysis of the nonlinear system derived in the previous section, both in its original and augmented forms. Necessary and sufficient conditions for the observability of the system, with a physical insight on the motion of the vehicle, are obtained, and global convergence results are established.

Although the introduction of the augmented system (4.4) has removed the output nonlinearity existent in the original nonlinear system (4.1), the presence of the ranges to invisible landmarks in the dynamics matrix introduces another kind of problem. Furthermore, given that the only available information with which to obtain $\mathbf{x}_{M_i}(t)$ and $x_{R_i}(t)$ is the corresponding range, it follows that the invisible landmarks and their ranges cannot be observable. For this reason, and according to the approach described in Section 2.2, the ranges to invisible landmarks and the landmarks themselves are removed from the system state, resulting in a reduced state, in which the dynamics matrix does not depend on the state as before, but solely on the system output. Thus, as the output is known, this new system may be regarded as linear time-varying for observability purposes.

For the sake of simplicity, and without loss of generality, in this section it is assumed that there is only one visible landmark, i.e., $\mathcal{M}_O := \{1\}$. This is possible due to the multi-single-range nature of the problem. Furthermore, as described, the invisible landmarks are discarded and left out of the new system state. Let $\mathbf{z}(t) = \begin{bmatrix}\mathbf{z}_{M_1}^T(t) & \mathbf{z}_V^T(t) & z_{R_1}(t)\end{bmatrix}^T$ be the new reduced state, and the corresponding system be

$$\begin{cases}\dot{\mathbf{z}}(t) = \mathbf{A}(t)\mathbf{z}(t)\\ \mathbf{y}(t) = \mathbf{C}\mathbf{z}(t)\end{cases}\tag{4.5}$$

where the dynamics matrix is given by

$$\mathbf{A}(t) := \begin{bmatrix} \mathbf{A}_{M_O}(t) & \mathbf{A}_{MV_O} & \mathbf{0}_{n_O \times N_O} \\ \mathbf{0}_{n_V \times n_O} & \mathbf{0}_{n_V} & \mathbf{0}_{n_V \times N_O} \\ \mathbf{A}_{RM_O}(t) & \mathbf{0}_{N_O \times n_V} & \mathbf{0}_{N_O} \end{bmatrix},$$

and the output matrix is simply

$$\mathbf{C} = \begin{bmatrix} \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} & 1 \end{bmatrix}.$$

Note that the blocks that constitute the dynamics matrix are the ones defined in the previous section, while including only the visible landmark. Also, the dependence on the system output and input can be seen as merely a dependence on time, as the two signals are known. This enables to consider (4.5) as a linear time-varying system (LTV), as shown in Lemma 2.1. This will be used throughout this section.

The following result addresses the observability analysis of the LTV system, but before proceeding with the analysis it is convenient to define ${}^E\mathbf{v}(t) = \mathbf{R}(t)\mathbf{v}(t)$ as the linear velocity of the vehicle in the Earth-fixed frame $\{E\}$.

Theorem 4.1 (Observability of the RO-SLAM LTV system)**.** *Consider the LTV system given by (4.5) and let $\mathcal{T} := [t_0, t_f]$. If and only if there exist three instants $\{t_1, t_2, t_3\} \in \mathcal{T}$ such that the linear velocity of the vehicle expressed in the Earth-fixed frame is linearly independent in those instants, i.e.,* $\det\left(\begin{bmatrix} {}^E\mathbf{v}(t_1) & {}^E\mathbf{v}(t_2) & {}^E\mathbf{v}(t_3) \end{bmatrix}\right) \neq 0$, *then the system is observable in the sense that, given the system output $\{\mathbf{y}(t), t \in \mathcal{T}\}$, the initial condition $\mathbf{z}(t_0)$ is uniquely defined.*

Proof. The proof starts by transforming the LTV system through a Lyapunov transformation, to simplify the analysis. A Lyapunov transformation (see Definition A.1 in Appendix A for details), preserves the observability properties of the original system. Consider then the tranformation

$$\mathbf{T}(t) = \operatorname{diag}(\mathbf{R}(t), \mathbf{I}_3, 1),$$

and the transformed system state given by

$$\chi(t) = \mathbf{T}(t)\mathbf{z}(t). \tag{4.6}$$

If it is considered that the angular velocity $\omega(t)$ is bounded, which is a physically sound assumption, the transformation $\mathbf{T}(t)$ has a continuous and bounded time derivative, while also having a bounded determinant itself, and therefore it is indeed a Lyapunov transformation. This means that it suffices to prove that the transformed system is observable.

Before proceeding with the proof it is necessary to derive the new system dynamics, by simply taking the first time derivative of (4.6) and using the inverse transformation $\mathbf{z}(t) = \mathbf{T}^{-1}(t)\chi(t)$, which results in

$$\begin{cases} \dot{\chi}(t) = \boldsymbol{\mathcal{A}}(t)\chi(t) \\ \mathbf{y}(t) = \mathbf{C}\chi(t) \end{cases}. \tag{4.7}$$

It is a simple matter of computation to see that the dynamics matrix of the transformed system is given by

$$\boldsymbol{\mathcal{A}}(t) = \begin{bmatrix} \mathbf{0}_3 & -\mathbf{R}(t) & \mathbf{0}_{3\times 1} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{3\times 1} \\ -\frac{\mathbf{y}_v^T(t)}{y_{R_1}(t)}\mathbf{R}^T(t) & \mathbf{0}_{1\times 3} & 0 \end{bmatrix},$$

while the output matrix of the transformed system is the same as in (4.5).

The proof follows by computing the transition matrix of the transformed system and subsequently the observability Gramian that will help determine whether the system is observable or not. The computation of the transition matrix can be made either using the Peano-Baker series or, in this case, by simply solving

$$\boldsymbol{\phi}(t,t_0)\boldsymbol{\chi}(t_0) = \boldsymbol{\chi}(t_0) + \int_{t_0}^{t} \boldsymbol{\mathcal{A}}(\tau)\boldsymbol{\chi}(\tau)\mathrm{d}\tau.$$

The computed transition is given by

$$\boldsymbol{\phi}(t,t_0) = \begin{bmatrix} \mathbf{I}_3 & -\mathbf{R}^{[1]}(t,t_0) & \mathbf{0}_{3\times 1} \\ \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_{3\times 1} \\ -\mathbf{v}^{[0]}(t,t_0) & \mathbf{v}^{[1]}(t,t_0) & 1 \end{bmatrix}, \tag{4.8}$$

where the following auxiliary quantities are introduced to simplify this expression,

$$\begin{cases} \mathbf{R}^{[1]}(t,t_0) = \int_{t_0}^{t} \mathbf{R}(\tau)\mathrm{d}\tau \\ \mathbf{v}^{[0]}(t,t_0) = \int_{t_0}^{t} \frac{\mathbf{y}_v^T(\tau)}{y_{R_1}(\tau)}\mathbf{R}^T(\tau)\mathrm{d}\tau \\ \mathbf{v}^{[1]}(t,t_0) = \int_{t_0}^{t} \frac{\mathbf{y}_v^T(\tau)}{y_{R_1}(\tau)}\mathbf{R}^T(\tau)\mathbf{R}^{[1]}(\tau,t_0)\mathrm{d}\tau \end{cases}.$$

The proof follows by contraposition, i.e., by establishing the hypothesis that the system (4.7) is not observable and then showing that, if the conditions of the theorem apply, the hypothesis cannot hold. Suppose then that the system is unobservable. This, from Lemma 2.1, implies that the observability Gramian $\boldsymbol{\mathcal{W}}(t_0,t_f)$, given by (2.3), is singular, which means that there exists a unit vector $\mathbf{c} = \begin{bmatrix}\mathbf{c}_m^T & \mathbf{c}_v^T & c_r\end{bmatrix}^T \in \mathbb{R}^{n_x}$ such that

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0,t_f)\mathbf{c} = 0. \tag{4.9}$$

Then, the substitution of (2.3) in (4.9) yields

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0,t_f)\mathbf{c} = \int_{t_0}^{t_f} \|\mathbf{C}\boldsymbol{\phi}(\tau,t_0)\mathbf{c}\|^2\mathrm{d}\tau, \tag{4.10}$$

where it is possible to substitute the argument of the norm in the integral by a vector function $\mathbf{f}(\tau,t_0)$ such that $\mathbf{f}(\tau,t_0) = \mathrm{diag}\left(\mathbf{R}^T(\tau),1\right)\mathbf{C}\boldsymbol{\phi}(\tau,t_0)\mathbf{c}$. Note that $\|\mathbf{f}(\tau,t_0)\| = \|\mathbf{C}\boldsymbol{\phi}(\tau,t_0)\mathbf{c}\|$.

In order for (4.9) to be true, both $\mathbf{f}(\tau, t_0)$ and its derivative must be zero for all $\tau \in \mathcal{T}$. Consider now the expressions for these two quantities, given by

$$\mathbf{f}(\tau, t_0) = \begin{bmatrix} \mathbf{c}_v \\ f_r(\tau, t_0) \end{bmatrix} \tag{4.11}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau, t_0) = \begin{bmatrix} \mathbf{0}_{3\times 1} \\ \frac{\mathrm{d}}{\mathrm{d}\tau} f_r(\tau, t_0) \end{bmatrix} \tag{4.12}$$

respectively. The component of $\mathbf{f}(\tau, t_0)$ associated with the range output is given by

$$f_r(\tau, t_0) = -\mathbf{v}^{[0]}(\tau, t_0)\mathbf{c}_m + \mathbf{v}^{[1]}(\tau, t_0)\mathbf{c}_v + c_r,$$

while its derivative is

$$\frac{\mathrm{d}}{\mathrm{d}\tau} f_r(\tau, t_0) = -\frac{\mathbf{y}_v^T(\tau)\mathbf{R}^T(\tau)}{y_{R_1}(\tau)}\mathbf{c}_m + \frac{\mathbf{y}_v^T(\tau)\mathbf{R}^T(\tau)}{y_{R_1}(\tau)}\mathbf{R}^{[1]}(\tau, t_0)\mathbf{c}_v.$$

Evaluating (4.11) at $\tau = t_0$ while equating the result to zero, automatically wields $\mathbf{c}_v = \mathbf{0}$ and $c_r = 0$. Furthermore, for (4.9) to be true, $\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau, t_0)$ must be zero in $\mathcal{T}$. This, recalling Assumption 4.1, leads to the final condition, expressed by

$${}^E\mathbf{v}^T(\tau)\,\mathbf{c}_m = 0,\ \forall \tau \in \mathcal{T}.$$

This condition can only be satisfied if $\mathbf{c}_m = \mathbf{0}$, which contradicts the hypothesis of the proof ($\|\mathbf{c}\| = 1$), or if the linear velocity of the vehicle in the Earth-fixed frame for any t_1, t_2, and t_3 in $\mathcal{T}$ is such that

$$\det\left(\begin{bmatrix} {}^E\mathbf{v}^T(t_1) \\ {}^E\mathbf{v}^T(t_2) \\ {}^E\mathbf{v}^T(t_3) \end{bmatrix}\right) = 0,$$

i.e., there do not exist three time instants such that the linear velocity of the vehicle in $\{E\}$ on these instants spans $\mathbb{R}^3$, which contradicts the conditions of the theorem. Therefore, by contraposition, if the conditions of the theorem hold, then $\mathcal{W}(t_0, t_f)$ is invertible on $\mathcal{T}$, and, using Lemma 2.1, it follows that (4.7) is observable. Furthermore, the LTV system (4.5) is also observable, as it is related with the system (4.7) through a Lyapunov transformation, thus concluding the sufficiency part of the proof.

The proof that the observability of the system implies that the conditions of the theorem hold, i.e., the necessity of those conditions, is also done by contraposition. It is assumed that the conditions of the theorem do not hold, and it is shown that, in this case, the system is not observable. Consider then that the conditions of the theorem do not hold. Therefore, the vehicle can only move in a plane or a line, which means that the Earth-fixed velocity vector is always in the same plane, i.e.,

$${}^E\mathbf{v}(t) = v_d(t)\mathbf{d} + v_{d^{\perp_1}}(t)\mathbf{d}^{\perp_1}, \quad \forall t \in \mathcal{T} \tag{4.13}$$

where $v_d(t)$ and $v_{d^{\perp_1}}(t) \in \mathbb{R}$, and $\mathbf{d}$ and $\mathbf{d}^{\perp_1}$ are two orthonormal vectors in $\mathbb{R}^3$ that define the plane or line (case defined by $v_{d^{\perp_1}}(t) = 0$) in which the vehicle moves. The computation of the explicit evolution of the output of the system (4.7) as a function of the initial state $\chi(t_0)$ is needed to proceed with the proof. Recall the definition of the transition matrix in (4.8) and the output in (4.7), which combined yield $\mathbf{y}(t) = \mathbf{C}\boldsymbol{\phi}(t,t_0)\chi(t_0)$ or

$$\mathbf{y}(t) = \begin{bmatrix} \chi_V(t_0) \\ -\mathbf{v}^{[0]}(t,t_0)\chi_M(t_0) + \mathbf{v}^{[1]}(t,t_0)\chi_V(t_0) + \chi_R(t_0) \end{bmatrix}.$$

The scalar part of the output can be further expanded to read

$$y_R(t) = \chi_R(t_0) + \mathbf{v}^{[1]}(t,t_0)\chi_V(t_0) - \int_{t_0}^{t} \frac{\mathbf{y}_v^T(\tau)\mathbf{R}^T(\tau)}{y_R(\tau)} \chi_M(t_0)\mathrm{d}\tau, \tag{4.14}$$

and substituting (4.13) in (4.14) leads to

$$y_R(t) = \chi_R(t_0) + \mathbf{v}^{[1]}(t,t_0)\chi_V(t_0) - \int_{t_0}^{t} \frac{v_d(\tau)}{y_R(\tau)}\mathrm{d}\tau\, \mathbf{d}\cdot\chi_M(t_0) - \int_{t_0}^{t} \frac{v_{d^{\perp_1}}(\tau)}{y_R(\tau)}\mathrm{d}\tau\, \mathbf{d}^{\perp_1}\cdot\chi_M(t_0). \tag{4.15}$$

The observability of the system implies that any two different initial states are distinguishable using only the output, i.e., if two different initial states produce the same output, the system cannot be observable. The proof follows by choosing two particular examples of initial states that produce the same output for all time. For that purpose, consider the two initial conditions

$$\tilde{\chi}(t_0) = \begin{bmatrix} (\mathbf{d}^{\perp_1}\times\mathbf{d})^T & \chi_V^T(t_0) & \chi_R(t_0) \end{bmatrix}^T$$

and

$$\bar{\chi}(t_0) = \begin{bmatrix} (\mathbf{d}\times\mathbf{d}^{\perp_1})^T & \chi_V^T(t_0) & \chi_R(t_0) \end{bmatrix}^T.$$

Observe that the range state $\chi_R(t_0)$ is the same for both initial states, which makes sense, as $\|\mathbf{d}^{\perp_1}\times\mathbf{d}\| = \|\mathbf{d}\times\mathbf{d}^{\perp_1}\|$. Noting that $\mathbf{d}^{\perp_1}\cdot\left(\mathbf{d}^{\perp_1}\times\mathbf{d}\right) = 0$ and $\mathbf{d}\cdot\left(\mathbf{d}^{\perp_1}\times\mathbf{d}\right) = 0$, it is straightforward to see that the two integrals in (4.15) vanish for these two initial states, rendering the output equal in both cases, i.e., $\tilde{\chi}(t_0)$ and $\bar{\chi}(t_0)$ are indistinguishable, as they lead to the same output for all $t \in \mathcal{T}$. Hence, if the conditions of the theorem do not hold, it is possible to find at least two indistinguishable initial states, which means that the transformed system (4.7) is unobservable in $\mathcal{T}$, or, conversely, if the transformed system is observable the theorem conditions must hold. As the two LTV systems are related by the Lyapunov transformation (4.6) which preserves observability properties, the proof of the necessity of those conditions is concluded. □

Remark 4.1. *This theorem establishes requirements for the motion of the vehicle, as it allows to conclude that if there are at least three velocity vectors that span $\mathbb{R}^3$, or, equivalently, if the trajectory of the vehicle is not restricted to a line or a plane, the system is observable. Note that, in a trilateration technique, ranges to four non-coplanar vehicle positions are needed. In this system, the motion of the vehicle provides those positions as Figure 4.2 shows for the bidimensional case (chosen for better visualization).*

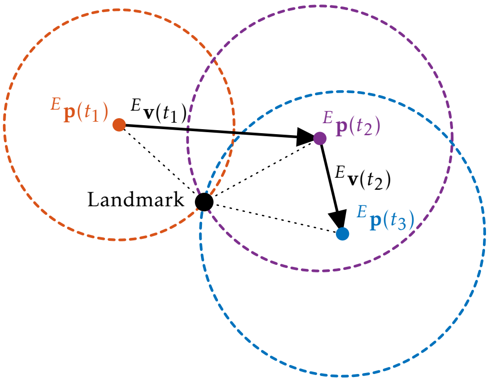


Figure 4.2: Trilateration for positioning a landmark in 2-D.

The LTV system (4.5) is a reduction of the augmented nonlinear system (4.4) in the sense that it does not include the invisible landmarks, and assumes the existence of only a single visible landmark. Due to the independence of the landmarks, whichever their number is, the two systems are completely equivalent in terms of observability, discarding the invisible landmarks. Note that the observability result of Theorem 4.1 when applied to the multi landmark case requires each landmark to be visible in the time instants when the velocity spans $\mathbb{R}^3$. This does not mean that all landmarks must be visible at the same time, but that in the time intervals in which each landmark is visible the conditions of Theorem 4.1 must hold. Furthermore, care must be taken before extending the observability results of this section to the original nonlinear system (4.1). As there is nothing imposing the state relations (4.2), it is not possible to apply the results to the original nonlinear system without some reflection. The following result addresses this issue: firstly, the equivalence of the state of the nonlinear and LTV systems is performed in a similar fashion to what was done in [BSO11b]. Secondly, it is shown that the nonlinear relation $\mathbf{z}_{R_1}(t) = \|\mathbf{z}_{M_1}\|$ is automatically imposed by the system dynamics when the LTV system is observable.

Theorem 4.2 (Equivalence of the RO-SLAM LTV and nonlinear systems)**.** *Consider the LTV system* (4.5) *and the original nonlinear system* (4.1)*. If the conditions of Theorem 4.1 hold, then*

(i) the state of the original nonlinear system and that of the LTV system are the same and uniquely defined, provided that the invisible landmarks are discarded. Furthermore the constraints expressed by (4.2) *become naturally imposed by the dynamics; and*

(ii) a state observer with uniformly globally exponentially stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, and the estimation error converges exponentially fast for all initial conditions.

Proof. The proof of the first part of the theorem follows by considering the system output and its relation to the states of the two systems in analysis, leading to a series of equations which, in the conditions of the theorem, result in the correspondence between the states, while imposing the algebraic constraints.

Consider the output of the LTV system (4.5), given by $\mathbf{y}(t) = \mathbf{C}\mathbf{z}(t)$, and that of the original nonlinear system, $\mathbf{y}(t) = \begin{bmatrix}\mathbf{v}^T(t) & \|\mathbf{p}_1(t)\|\end{bmatrix}^T$. Recall that, in this section, it is assumed that the visible landmark set has only one landmark. The comparison of the two outputs shows that

$$\begin{cases} \mathbf{z}_V(t) = \mathbf{v}(t) \\ z_{R_1}(t) = \|\mathbf{p}_1(t)\| \end{cases}.$$

For the purpose of finding the correspondence of the remaining states, $\mathbf{z}_{M_1}(t)$ and $\mathbf{p}_1(t)$, consider the time evolution of the rotation of each of these states, given by

$$\mathbf{R}(t)\mathbf{z}_{M_1}(t) = \mathbf{R}(t_0)\mathbf{z}_{M_1}(t_0) - \mathbf{R}^{[1]}(t,t_0)\mathbf{z}_V(t_0), \tag{4.16}$$

and

$$\mathbf{R}(t)\mathbf{p}_1(t) = \mathbf{R}(t_0)\mathbf{p}_1(t_0) - \mathbf{R}^{[1]}(t,t_0)\mathbf{v}(t_0), \tag{4.17}$$

respectively. These can be derived following the same reasoning used to compute the transition matrix (4.8). The expressions in (4.16) and (4.17) can be used to compute the derivative of the outputs of the two systems in analysis related to the ranges. Recall that the range output of the LTV system is the same as the range state of that system. Then, using the dynamics of that state given by

$$\dot{z}_{R_1}(t) = -\frac{\mathbf{y}_v^T(t)}{y_{R_1}(t)}\mathbf{z}_{M_1}(t),$$

replacing $\mathbf{z}_{M_1}(t)$ by $\mathbf{R}^T(t)\mathbf{R}(t)\mathbf{z}_{M_1}(t)$, and using (4.16) yields

$$\dot{y}_{R_1}(t) = \frac{\mathbf{y}_v^T(t)\mathbf{R}^T(t)}{y_{R_1}(t)}\left(\mathbf{R}(t_0)\mathbf{z}_{M_1}(t_0) - \mathbf{R}^{[1]}(t,t_0)\mathbf{z}_V(t_0)\right). \tag{4.18}$$

For the original system, recall the derivative of the range output given by (4.3), and substitute $\mathbf{p}_1(t)$ by $\mathbf{R}^T(t)\mathbf{R}(t)\mathbf{p}_1(t)$ in order to be able to use (4.17) and reach

$$\dot{r}_1(t) = \frac{\mathbf{v}^T(t)\mathbf{R}^T(t)}{r_1(t)}\left(\mathbf{R}(t_0)\mathbf{p}_1(t_0) - \mathbf{R}^{[1]}(t,t_0)\mathbf{v}(t_0)\right). \tag{4.19}$$

It is possible to compare these two expressions for the derivatives of the outputs of the two systems, if it is noted that $\mathbf{z}_V(t) = \mathbf{v}(t)$, $r_1(t) = y_{R_1}(t)$, and ${}^E\mathbf{v}(t) = \mathbf{R}(t)\mathbf{y}_v(t) = \mathbf{R}(t)\mathbf{v}(t)$. Then, the comparison of (4.18) with (4.19) yields

$$\frac{{}^E\mathbf{v}^T(t)}{r_1(t)}\mathbf{R}(t_0)\left(\mathbf{z}_{M_1}(t_0) - \mathbf{p}_1(t_0)\right) = 0$$

for all t in $\mathcal{T}$. When the conditions of Theorem 4.1 hold, this expression implies $\mathbf{z}_{M_1}(t_0) = \mathbf{p}_1(t_0)$ by the same reasoning used to prove the sufficiency of these conditions to the observability of the LTV system. Since the dynamics of these states are the same by construction, the equivalence of the initial conditions implies that the system states are equivalent. Hence, if the conditions for the observability of the LTV system apply, then the state of the system (4.5) corresponds directly to that of the original nonlinear system, disregarding the invisible landmarks. Furthermore, noting that it is imposed by the output that $z_{R_1}(t)$ is the norm of $\mathbf{p}_1(t)$ and, if the conditions of the theorem hold, that $\mathbf{p}_1(t) = \mathbf{z}_{M_1}(t)$, then the constraint $z_{R_1}(t) = \|\mathbf{z}_{M_1}(t)\|$ becomes naturally imposed by the dynamics. Thus, the proof of the first part of the theorem is concluded.

The second part of the theorem follows naturally from the first part. An observer for (4.5) with globally exponentially stable error dynamics provides estimates that converge exponentially fast to the true state. Therefore, if the state of the LTV system and that of the original nonlinear system, when the invisible landmarks are discarded, are one and the same, the estimates of the observer will also tend exponentially fast to the true state of system (4.1) with the same error dynamics. □

Theorems 4.1 and 4.2 provided sufficient conditions for the observability of the original nonlinear system while establishing a direct relation between the augmented system 4.5 and the nonlinear one. However, a stronger result can be accomplished by means of the following theorem, which states that these conditions are indeed necessary and sufficient.

Theorem 4.3 (Observability of the RO-SLAM nonlinear system)**.** *The nonlinear system* (4.1) *is observable, when discarding the invisible landmarks, if and only if the conditions of Theorem 4.1 hold.*

Proof. The sufficiency part of the proof is readily provided by Theorems 4.1 and 4.2. The former establishes conditions for the observability of the LTV system (4.5) and the latter relates the state of that system to that of the nonlinear system in analysis. The proof of the necessity part of the theorem is made in very similar terms to that of the second part of Theorem 4.1 in the sense that the system output, written as a function of the initial state using (4.17)

$$r_1(t) = \|\mathbf{R}(t)\mathbf{p}_1(t)\| = \|\mathbf{R}(t_0)\mathbf{p}_1(t_0) - \mathbf{R}^{[1]}(t,t_0)\mathbf{v}(t_0)\|,$$

is shown to be the same given two different initial states. Squaring this expression and expanding the result yields

$$r_1^2(t) = \|\mathbf{R}(t_0)\mathbf{p}_1(t_0)\|^2 + \|\mathbf{R}^{[1]}(t,t_0)\mathbf{v}(t_0)\|^2 - 2\int_{t_0}^{t} \left(v_d(\tau)\mathbf{d} + v_{d^{\perp_1}}(\tau)\mathbf{d}^{\perp_1}\right)\mathbf{R}(t_0)\mathbf{p}_1(t_0)\mathrm{d}\tau, \tag{4.20}$$

where the definitions of $\mathbf{R}^{[1]}(t,t_0)$ and ${}^{E}\mathbf{v}(\tau) = \mathbf{R}(\tau)\mathbf{v}(t_0)$ were employed. Furthermore, the linear velocity in $\{E\}$ was also replaced in (4.20) by the negation of the conditions of the theorem expressed by (4.13). Consider two initial states defined as

$$\tilde{\xi}(t_0) = \begin{bmatrix} \mathbf{R}^T(t_0)\left(\mathbf{d}\times\mathbf{d}^{\perp_1}\right) \\ \mathbf{v}(t_0) \end{bmatrix}$$

and

$$\bar{\xi}(t_0) = \begin{bmatrix} \mathbf{R}^T(t_0)\left(\mathbf{d}^{\perp_1} \times \mathbf{d}\right) \\ \mathbf{v}(t_0) \end{bmatrix}.$$

Then, it is possible to see that the integral in (4.20) vanishes for both cases, and that these two initial states lead to the same output, which implies that if the conditions of the theorem do not hold, the system is unobservable. This concludes the proof of the necessity of the conditions of the theorem for the nonlinear system to be observable. □

The previous results show that, if it is possible to design a globally exponentially stable observer for the LTV system, it will also be suitable for the original nonlinear system. This has established the ground to the design of such an observer, using a linear time-varying Kalman filter, which, to assure the GES nature of the estimation error dynamics, requires the pair $(\mathbf{A}(t), \mathbf{C})$ to be uniformly completely observable, as shown in Theorem 2.2. The following theorem addresses the uniform complete observability of the pair $(\mathbf{A}(t), \mathbf{C})$. However, an additional assumption on the linear velocity of the vehicle is required.

Assumption 4.2. *The norm of the linear velocity of the vehicle in the Earth-fixed frame* $\{E\}$ *is always bounded, i.e.,*

$$\underset{t \geq t_0}{\forall} \quad \underset{V_M > 0}{\exists} : \quad \|{}^E\mathbf{v}(t)\| \leq V_M.$$

Although imposing definite bounds on the linear velocity, this assumption is still a mild one, as it is physically impossible to reach arbitrarily large speeds. Moreover, the value of V_M is not required for the filter design.

Theorem 4.4 (Uniform complete observability of the RO-SLAM LTV system)**.** *The pair* $(\mathbf{A}(t), \mathbf{C})$ *is uniformly completely observable if and only if Assumption 4.2 is true and there exist* $\delta > 0$ *and* $\alpha^* > 0$ *such that, for all* $t \geq t_0$*, it is possible to choose a set of instants* $\{t_1, t_2, t_3\} \in \mathcal{T}_\delta$*, with* $\mathcal{T}_\delta := [t, t+\delta]$*, for which the linear velocity of the vehicle in the Earth-fixed frame respects*

$$\left|\det\left(\begin{bmatrix} {}^E\mathbf{v}(t_1) & {}^E\mathbf{v}(t_2) & {}^E\mathbf{v}(t_3) \end{bmatrix}\right)\right| > \alpha^*. \tag{4.21}$$

Proof. The concept of uniform complete observability implies uniform bounds on the observability Gramian in time intervals of length δ. Considering that the pair $(\mathbf{A}(t), \mathbf{C})$ is related to $(\boldsymbol{\mathcal{A}}(t), \mathbf{C})$ by a Lyapunov transformation, the uniform complete observability of the latter implies the uniform complete observability of the former, and thus the proof will focus on this pair.

Consider the previous definitions of the observability Gramian in (2.3) and of the arbitrary unitary vector $\mathbf{c}$, and note that these are bounded. Then, the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C})$ may be expressed through the following statement,

$$\underset{\substack{\delta > 0 \\ \alpha > 0}}{\exists} \quad \underset{t \geq t_0}{\forall} \quad \underset{\substack{\mathbf{c} \in \mathbb{R}^{n_\mathcal{X}} \\ \|\mathbf{c}\| = 1}}{\forall} : \quad \mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} \geq \alpha, \tag{4.22}$$

reproduced here from Definition 2.4 for convenience. The proof follows by exhaustion, i.e., it consists of studying $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$ for every possible case of $\mathbf{c}$ and showing that (4.22) is true for all of them.

Recall the proof of Theorem 4.1, and the expansion of $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0,t_f)\mathbf{c}$ therein, with the definition of $\mathbf{f}(\tau,t)$ and $\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau,t)$ in (4.11) and (4.12). Then, it is possible to write

$$\|\mathbf{f}(\tau,t)\|^2 = \|\mathbf{c}_v\|^2 + \left(\mathbf{v}^{[0]}(\tau,t)\mathbf{c}_m - \mathbf{v}^{[1]}(\tau,t)\mathbf{c}_v - c_r\right)^2. \tag{4.23}$$

Consider Proposition A.1 from Appendix A. It is possible to see that it applies to $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t_0,t_f)\mathbf{c}$, and therefore it suffices to show that $\|\mathbf{f}(\tau,t)\|$ is lower bounded for every possible $\mathbf{c}$.

The first condition to study is $\|\mathbf{c}_v\| \geq \alpha_v > 0$ without any restriction on both $\mathbf{c}_m$ or c_r. In that case, the norm of $\mathbf{f}(\tau,t)$ respects

$$\|\mathbf{f}(\tau,t)\| \geq \|\mathbf{c}_v\| \geq \alpha_v,$$

and thus $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} \geq \alpha_1 > 0$. The second case is $|c_r| \geq \alpha_r > 0$, again with no restriction on the remaining quantities. Then, it is possible to evaluate $\mathbf{f}(\tau,t)$ for $\tau = t$ which leads to

$$\|\mathbf{f}(t,t)\| \geq |c_r| \geq \alpha_r.$$

The last case involves setting $\|\mathbf{c}_v\| < \alpha_v$ and $|c_r| < \alpha_r$, while imposing $\|\mathbf{c}_m\| \geq \alpha_m > 0$. For the purpose of analysing this case, consider the new function

$$f_{pv}(\tau,t) := \mathbf{v}^{[0]}(\tau,t)\mathbf{c}_m - \mathbf{v}^{[1]}(\tau,t)\mathbf{c}_v$$

such that $f_r(\tau,t) = -f_{pv}(\tau,t) + c_r$. Then, it is possible to write

$$\begin{aligned}\|\mathbf{f}(\tau,t)\|^2 &\geq f_r^2(\tau,t)\\ &\geq |f_{pv}(\tau,t)|(|f_{pv}(\tau,t)| - 2|c_r|)\\ &\geq |f_{pv}(\tau,t)|(|f_{pv}(\tau,t)| - 2\alpha_r).\end{aligned}$$

Considering that $f_{pv}(\tau,t)$ is lower bounded by some α_{pv}, if the upper bound on c_r is chosen to be smaller than $\frac{1}{4}\alpha_{pv}$, then

$$\|\mathbf{f}(\tau,t)\|^2 \geq \frac{\alpha_{pv}^2}{2},$$

and it suffices to show that the norm of the function $f_{pv}(\tau,t)$ is lower bounded. Then, one has $f_{pv}(t,t) = 0$, and Proposition A.1 applies once more, i.e., if it is shown that $\left|\frac{\mathrm{d}}{\mathrm{d}\tau}f_{pv}(\tau,t)\right| \geq \alpha_2$, then $\|\mathbf{f}(\tau,t)\|$ is lower bounded as intended. The norm of the derivative of $f_{pv}(\tau,t)$ is given by

$$\left|\frac{\mathrm{d}}{\mathrm{d}\tau}f_{pv}(\tau,t)\right| = \left|\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{c}_m - \frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{R}^{[1]}(\tau,t)\mathbf{c}_v\right|, \tag{4.24}$$

where, in the definitions of $\mathbf{v}^{[0]}(\tau,t)$ and $\mathbf{v}^{[1]}(\tau,t)$, $\mathbf{R}(\tau)\mathbf{y}_v(\tau)$ was replaced by ${}^E\mathbf{v}(\tau)$ and $y_{R_1}(\tau)$ by $r_1(\tau)$. The square of (4.24) can be expanded to read

$$\left|\frac{\mathrm{d}}{\mathrm{d}\tau}f_{pv}(\tau,t)\right|^2 = \left|\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{c}_m\right|^2 + \left|\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{R}^{[1]}(\tau,t)\mathbf{c}_v\right|^2 - 2\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{c}_m\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{R}^{[1]}(\tau,t)\mathbf{c}_v.$$

Consider now Assumptions 4.1 and 4.2. Using the bounds defined therein, it is possible to write

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}\tau} f_{pv}(\tau,t)\right|^2 &\geq \left|\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{c}_m\right| \left(\left|\frac{{}^E\mathbf{v}^T(\tau)}{r_1(\tau)}\mathbf{c}_m\right| - 2\frac{\|{}^E\mathbf{v}(t)\|\|\mathbf{c}_v\|}{r_1(\tau)} \right) \\ &\geq \frac{\left|{}^E\mathbf{v}^T(\tau)\mathbf{c}_m\right|}{R_M^2} \left(\left|{}^E\mathbf{v}^T(\tau)\mathbf{c}_m\right| - 2\frac{R_M V_M \alpha_v}{R_m} \right).\end{aligned}$$

Assuming that $\left|{}^E\mathbf{v}^T(\tau)\mathbf{c}_m\right| \geq \alpha_3$ for some τ, if it is defined that $\alpha_v < \frac{R_m}{4R_M V_M}\alpha_3$, then the norm of this derivative respects

$$\left|\frac{\mathrm{d}}{\mathrm{d}\tau} f_{pv}(\tau,t)\right|^2 \geq \frac{1}{2R_M^2}\alpha_3^2.$$

The proof follows now by showing that there is at least a t^* such that the absolute value of the inner product of the linear velocity in $\{E\}$ with $\mathbf{c}_m$ in that instant is greater than some α_3 for all $\|\mathbf{c}_m\| \geq \alpha_m$. Consider the sum of the mentioned inner product for the time instants t_1, t_2, and t_3. Given that $\mathbf{c}_m$ cannot be perpendicular to all velocities, as it is impossible according to the conditions of the theorem, the worst case occurs when $\mathbf{c}_m$ is perpendicular to the two velocity vectors that have the greatest cross product, say ${}^E\mathbf{v}(t_2)$ and ${}^E\mathbf{v}(t_3)$. The condition of the theorem can be written as

$$\left|{}^E\mathbf{v}^T(t_1)\mathbf{S}\big({}^E\mathbf{v}(t_2)\big){}^E\mathbf{v}(t_3)\right| \geq \alpha. \tag{4.25}$$

In the worst case now defined, $\mathbf{c}_m = \pm\frac{\mathbf{S}({}^E\mathbf{v}(t_2)){}^E\mathbf{v}(t_3)}{\|\mathbf{S}({}^E\mathbf{v}(t_2)){}^E\mathbf{v}(t_3)\|}\|\mathbf{c}_m\|$, which can be substituted in (4.25) to yield

$$\left|{}^E\mathbf{v}^T(t_1)\mathbf{c}_m\right| \geq \alpha\frac{\alpha_m}{V_M^2}.$$

Hence, in the conditions of the theorem, the sum of the absolute values of the inner products of the linear velocity in $\{E\}$ with $\mathbf{c}_m$ at t_1, t_2, and t_3 respects

$$\sum_{i=1}^{3}\left|{}^E\mathbf{v}^T(t_i)\mathbf{c}_m\right| \geq \alpha\frac{\alpha_m}{V_M^2}.$$

This means that there exists at least a t^* equal to t_1, t_2, t_3 or any combination of the three such that

$$\left|{}^E\mathbf{v}^T(t^*)\mathbf{c}_m\right| \geq \alpha\frac{\alpha_m}{3V_M^2},$$

which implies that

$$\left|\frac{\mathrm{d}}{\mathrm{d}\tau} f_{pv}(t^*,t)\right| \geq \frac{\alpha\alpha_m}{3\sqrt{2}R_M V_M^2},$$

and, by Proposition A.1, $|f_{pv}(\tau,t)| \geq \alpha_3$, which in turn implies $\|\mathbf{f}(\tau,t)\| \geq \alpha_4$.

This concludes the enumeration of all the possible values for $\mathbf{c}$, and the proof that, for every possibility, $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$ is lower bounded, thus proving the sufficiency of the conditions of the theorem for the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t),\mathbf{C})$. Being related to the LTV system (4.5) through a Lyapunov transformation, this result implies that $(\mathbf{A}(t),\mathbf{C}(t))$ is also uniformly completely observable, and thus the proof of the sufficiency part of the theorem is concluded.

In the same manner of the previous necessity theorems, the proof is made by contraposition, i.e, by showing that the negation of the conditions implies that the pair $(\mathbf{A}(t), \mathbf{C})$ cannot be uniformly completely observable. Consider then the negation of the conditions of the theorem, stated by

$$\underset{\substack{\delta>0\\ \alpha>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{t_1,t_2,t_3\in\mathcal{T}_\delta}{\forall} : \quad \left|\det\left(\begin{bmatrix} {}^E\mathbf{v}^T(t_1) \\ {}^E\mathbf{v}^T(t_2) \\ {}^E\mathbf{v}^T(t_3)\end{bmatrix}\right)\right| < \alpha$$

This means that the linear velocity of the vehicle in the Earth-fixed frame can move in $\mathbb{R}^3$ even though the quality of the tridimensional space it spans is degraded, i.e., considering that the Earth-fixed linear velocity is, in general, given by ${}^E\mathbf{v}(t) = v_d(t)\mathbf{d} + v_{d^{\perp_1}}(t)\mathbf{d}^{\perp_1} + v_{d^{\perp_2}}(t)\mathbf{d}^{\perp_2}$, there is at least one direction $\mathbf{d}^{\perp_2}$ in $\mathbb{R}^3$ upon which the linear velocity is upper bounded and as small as wanted, i.e.,

$$\underset{t\in\mathcal{T}_\delta}{\forall} \quad \underset{\beta>0}{\forall} : \quad |{}^E\mathbf{v}(t)\cdot\mathbf{d}^{\perp_2}| < \beta.$$

The proof follows by showing that this limitation on the linear velocity implies that the pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C})$ is not uniformly completely observable, i.e.,

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_X}\\ \|\mathbf{c}\|=1}}{\exists} : \quad \mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} < \epsilon. \tag{4.26}$$

For that purpose, consider the substitution of (4.23) in the definition of $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$ given by (4.10), which yields

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t_0)\mathbf{c} = \int_t^{t+\delta} \|\mathbf{c}_v\|^2 \mathrm{d}\tau + \int_t^{t+\delta} \left(\mathbf{v}^{[0]}(\tau,t)\mathbf{c}_m - \mathbf{v}^{[1]}(\tau,t)\mathbf{c}_v - c_r\right)^2 \mathrm{d}\tau. \tag{4.27}$$

The statement (4.26) requires the existence of a single unit vector $\mathbf{c}$ that satisfies it. Therefore, it is possible to choose particular values for $\mathbf{c}_m$, $\mathbf{c}_v$, and c_r that aid the development of the proof. For that purpose, consider that $\mathbf{c}_v$ and c_r are both zero. In this case, (4.27) becomes

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} = \int_t^{t+\delta}\left[\int_t^{\tau} \frac{\mathbf{y}_v^T(\sigma)\mathbf{R}^T(\sigma)}{y_{R_1}(\sigma)}\mathbf{c}_m \mathrm{d}\sigma\right]^2 \mathrm{d}\tau,$$

which, noting that ${}^E\mathbf{v}(t) = \mathbf{R}(t)\mathbf{y}_v(t)$ and $y_{R_1}(t) = r_1(t)$ and using the Cauchy-Schwartz inequality, allows to write

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} \leq \int_t^{t+\delta}\int_t^{\tau}\left[\frac{{}^E\mathbf{v}^T(\sigma)\mathbf{c}_m}{r_1(\sigma)}\right]^2 \mathrm{d}\sigma\mathrm{d}\tau.$$

If $\mathbf{c}_m$ is chosen as the unit vector $\mathbf{d}^{\perp_2}$ previously introduced, it is possible to further simplify this expression, yielding

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_X}\\ \|\mathbf{c}\|=1}}{\exists} : \quad \mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} < \frac{\beta^2\delta^2}{2R_m^2} := \epsilon,$$

where $\beta := \sqrt{2\epsilon}\frac{R_m}{\delta}$ and Assumption 4.1 was used. From this it is possible to conclude that if the conditions of the theorem do not hold, the pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C})$ is not uniformly completely observable. This pair belongs to system (4.7), that is related to the LTV system (4.5) by the Lyapunov transformation (4.6) that maintains observability properties. Hence, it follows that if the pair $(\mathbf{A}(t), \mathbf{C})$ is uniformly completely observable, the conditions of the theorem must hold, thus concluding the proof of the necessity part of the theorem. ☐

Remark 4.2. *The determinant in* (4.21) *can be written as* ${}^{E}\mathbf{v}(t_1) \cdot {}^{E}\mathbf{v}(t_2) \times {}^{E}\mathbf{v}(t_3)$*. Together with Assumption 4.2, this condition can be understood as a persistent excitation condition, i.e., the velocity at* t_1 *must be sufficiently out of the plane defined by the velocity at* t_2 *and* t_3 *for the vector space defined by them not to degenerate in time.*

4.4 Filter design and implementation

This section addresses the design of the sensor-based RO-SLAM filter. The theoretical results of the previous section were established in a deterministic setting, and thus the presence of measurement noise raises the need for a filtering solution. Theorems 4.2, 4.3, and 4.4 show that it is possible to design an observer with globally exponentially stable error dynamics for the nonlinear system (4.1). Hence, a Kalman filter follows naturally for the augmented nonlinear system (4.4), in a discrete-time setting. Recall the discrete-time system (2.9). For the implementation of RO-SLAM, the process employed to obtain it was the Euler forward discretization, with a small detail regarding the rotation of a landmark from one instant to the following. For the purpose of obtaining this rotation, it is considered that the angular velocity is constant over each sampling interval and the exponential map (see Definition 2.1) is used. It follows that

$$\mathbf{R}_{k+1}^{T}\mathbf{R}_k = \exp\left(-\mathbf{S}(\boldsymbol{\omega}_k)\, T_s\right),$$

and the dynamics matrix becomes

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{F}_{M_k} & T_s\mathbf{A}_{MV} & \mathbf{0}_{n_L \times n_R} \\ \mathbf{0}_{n_V \times n_M} & \mathbf{I}_3 & \mathbf{0}_{n_L \times n_R} \\ T_s\mathbf{A}_{RM_k} & \mathbf{0}_{n_R \times n_V} & \mathbf{I}_{n_R} \end{bmatrix},$$

where $\mathbf{F}_{M_k} = \operatorname{diag}\left(\mathbf{R}_{k+1}^{T}\mathbf{R}_k, \ldots, \mathbf{R}_{k+1}^{T}\mathbf{R}_k\right)$.

The generic description of subsection 2.3.2 serves as the basis for the work detailed in this section, and, since the signals received from the beacon landmarks are tagged, the association of measured data with state data is trivial. For this reason, there is no need for a data association algorithm, nor for a loop closure procedure, which means that the algorithm is a standard discrete Kalman filter, (see [Gel74]), with the detail that, when a landmark is invisible and its range is unavailable, the estimated range is used in the predict step, allowing the propagation in open loop of the invisible landmarks.

4.5 Simulation results

In this section, results from a typical run in a simulation setting are presented. The simulated environment consists of 20 landmarks spread randomly throughout a 16m×16m×3m map. The trajectory of the vehicle was designed in order to satisfy the observability conditions, and it can be seen as the blue line in Figure 4.3. For more details on the simulation setup, see Appendix C. All the measurements are assumed to be perturbed by zero-mean Gaussian white noise, with standard deviations of $\sigma_\omega = 0.05$ °/s for the angular rates, $\sigma_v = 0.03$ m/s for the linear velocity, and $\sigma_r = 0.03$ m for the ranges. The Kalman filter parameters were chosen as indicated in Table 4.1.

Table 4.1: RO-SLAM Kalman filter parameters

	Position $[m^2]$	Velocity $[(m/s)^2]$	Range $[m^2]$
$\boldsymbol{\Xi}$	$T_s 10^{-3}\mathbf{I}_{3N}$	$T_s 10^{-2}\mathbf{I}_{3N}$	$T_s 10^{-5}\mathbf{I}_{3N}$
$\boldsymbol{\Theta}$	–	$10^{-3}\mathbf{I}_3$	$\mathbf{I}_{N_O}$

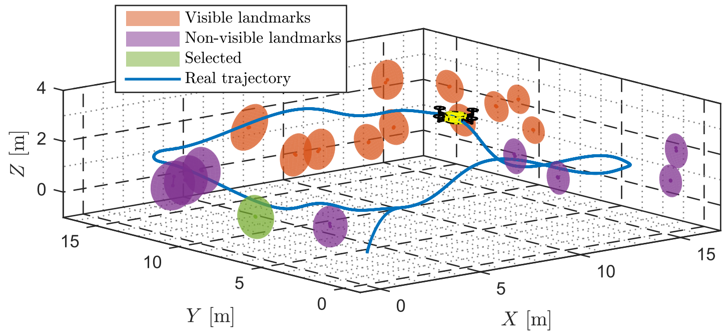
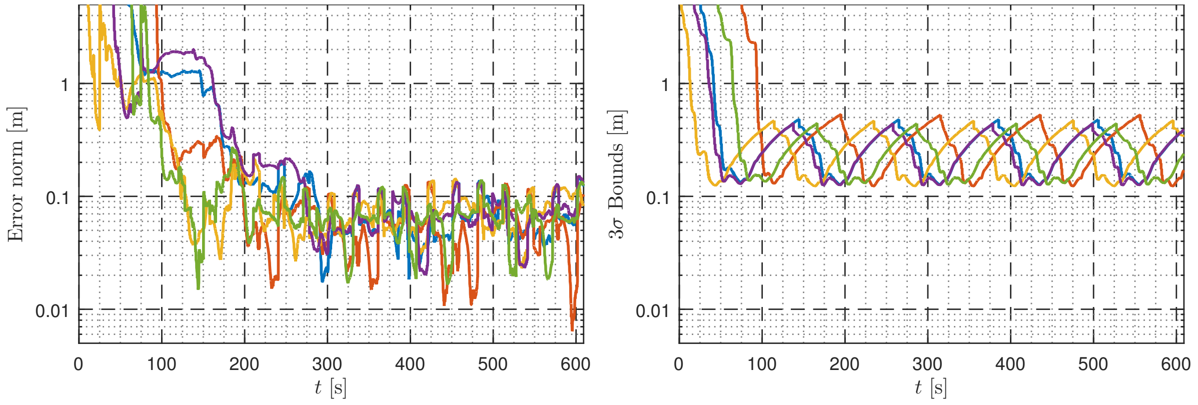


Figure 4.3: Picture of the estimated map rotated and translated using the true transformation at $t = 300$ s. The real trajectory is the blue line, the red ellipsoids represent the 3σ uncertainty of the currently visible landmark estimates, the purple ellipsoids represent the uncertainty of the non-visible landmarks, and the green ellipsoid is the landmark selected for closer inspection on Figure 4.5.



(a) Norm of the error.

(b) Uncertainty.

Figure 4.4: Evolution of the estimation of 5 landmarks in time.

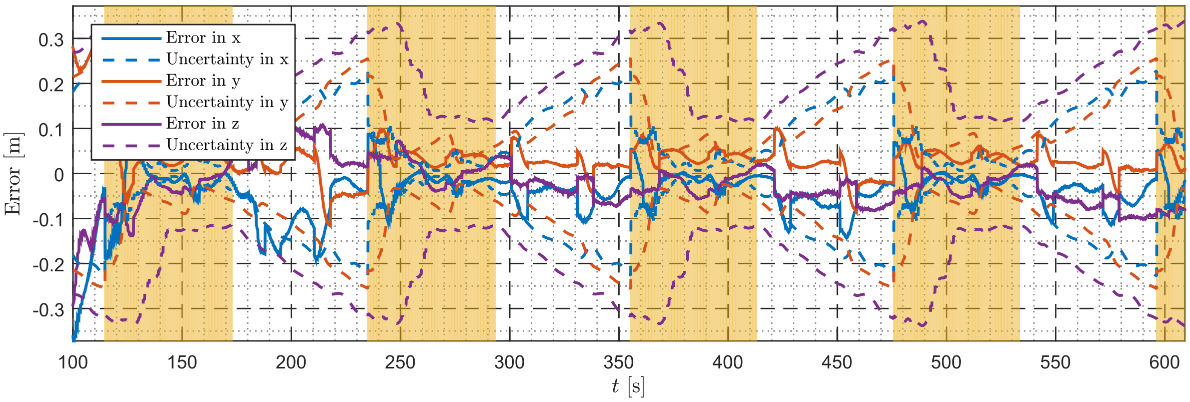


Figure 4.5: The estimation error of a single landmark for all coordinates with 3σ uncertainty bounds. The yellow bars indicate the moments when the landmark was observed.

The performance of the RO-SLAM filter can be assessed through Figure 4.4, where the norms of the estimation error of 5 landmarks are presented: after the initial transient period, the error stays within an interval with magnitude of 10 cm. The estimation error of the velocity is understandably small, as the quantity is directly observed. Its mean is below 10^{-4} m/s and its standard deviation below 10^{-3} m/s. The range error, that grows for invisible landmarks, has a mean of 0.0266 m and its standard deviation is 0.0435 m. For better visualization, in Figure 4.5 it is presented in detail the estimated errors and uncertainty bounds of a single sensor-based landmark. These results are in accordance with the theoretical results of Section 4.3, as the visible landmarks, as well as the other estimated quantities, converge both in uncertainty and in error. The periodic pattern that can be observed there is due to the increase and decrease of uncertainty that occurs when a landmark is not seen and suddenly reappears in the sensors field-of-view (FOV), or the persistent excitation conditions of Theorem 4.4 are not fulfilled.

Finally, an example of the estimated map is given in Figure 4.3, where the coloured ellipsoids represent the uncertainty associated ($3\sigma_p$) and the small circles mark the true coordinates of each landmark. The yellow quadrotor marks the position of the vehicle at the time of the estimation and the blue solid line is the real trajectory. The estimated map was rotated and translated to the Earth-fixed frame. Note that the 3σ uncertainty ellipsoids surround the true values, as they should in a consistent filter.

This simulation was designed to attest the validity of the theoretical results presented in this chapter, as well as the convergence properties of the RO-SLAM filter here proposed. It was shown that the algorithm is able to produce a consistent map, depicted in Figure 4.3. Moreover, the good performance of the algorithm for a sufficiently rich trajectory was demonstrated.

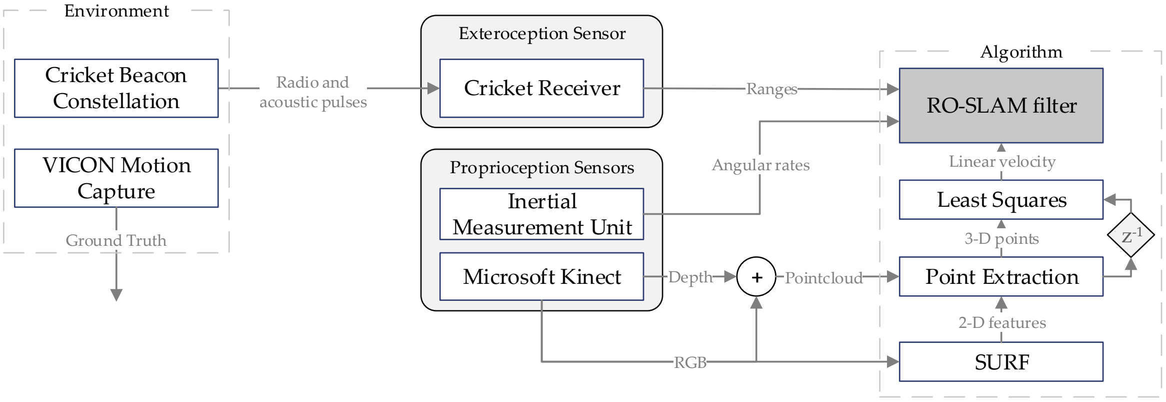

Environment
Cricket Beacon Constellation
VICON Motion Capture
Ground Truth
Radio and acoustic pulses
Exteroception Sensor
Cricket Receiver
Proprioception Sensors
Inertial Measurement Unit
Microsoft Kinect
Ranges
Angular rates
Depth
+
Pointcloud
RGB
Algorithm
RO-SLAM filter
Linear velocity
Least Squares
3-D points
Point Extraction
2-D features
SURF
z-1

Figure 4.6: The flow of information in the proposed algorithm. The process employed to obtain linear velocity measurements is also explained.

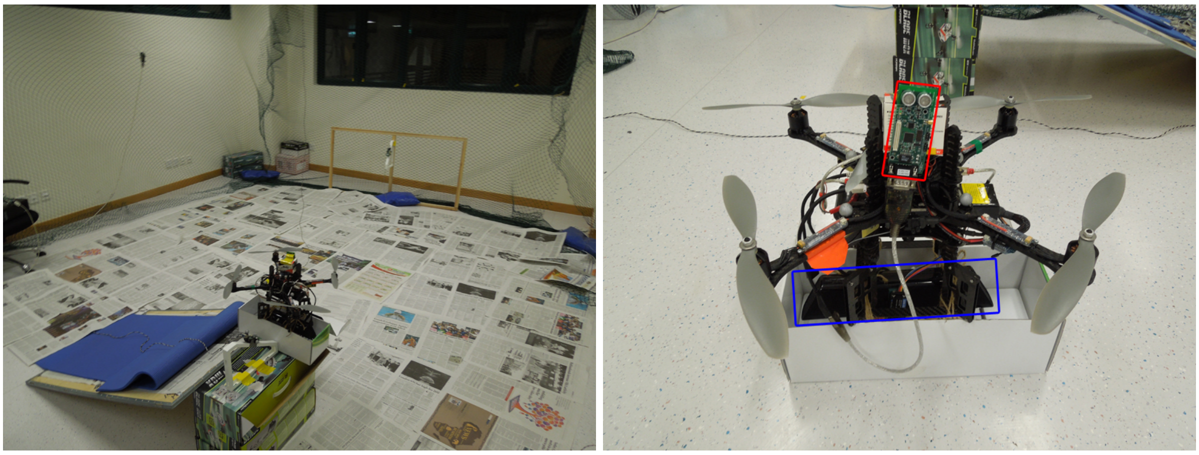
4.6 Experimental results

4.6.1 Setup

This section details an experiment that took place in the Sensor-based Cooperative Robotics Research (SCORE) Lab of the Faculty of Science and Technology of the University of Macau. The experimental setup consists of an *AscTec Pelican* quadrotor instrumented with a *Microstrain 3DM-GX3-25* IMUs working at 200 Hz, a *Microsoft Kinect*, at 10 Hz, a *Crossbow Cricket* receiver, and *VICON* markers (see Appendix D for further details on the experimental setup). Furthermore, the lab was equipped with 7 more *Crossbow Cricket* motes, emitting sequentially one at a time at 10 Hz (each beacon emits every 700 ms), as well as with a *VICON* Bonita motion capture system, providing accurate estimates of the linear and angular motion quantities of the vehicle, that was used for validation of the estimates provided by the RO-SLAM algorithm.

The cricket (highlighted in red in Figure 4.7(b)) is a small hardware platform consisting of a RF transceiver, a micro-controller, and other associated hardware for generating and receiving ultrasonic signals and interfacing with a host device [PCB00]. Each cricket beacon emits simultaneously a radio and an acoustic pulse that are received by the cricket placed in the vehicle, thus allowing the computation of the ranges through difference of time of arrival of the two pulses using the speed of sound corrected by on-board temperature sensors.

The facing down camera (highlighted in blue in Figure 4.7(b)) is used to compute the linear velocity of the vehicle, through the process depicted in Figure 4.6. An implementation of SURF, see [BETG08], detects features in the red-green-blue (RGB) images which correspond to tridimensional points in the Kinect pointclouds. The N_k tridimensional features ($\mathbf{f}_k^{(i)}$) of two subsequent frames are associated using a Sequential Compatibility Nearest Neighbour [NT01] algorithm, and then used in a Least Squares procedure to obtain the linear velocity in 3-D



(a) The experiment area.

(b) The instrumented quadrotor.

Figure 4.7: The experimental setup. The Cricket beacons can be seen on the left spread around the room. The Cricket receiver is mounted on the quadrotor on the right.

through the following equation

$$\mathbf{v}_k = \begin{bmatrix}\mathbf{I}_3 & \mathbf{0}_3\end{bmatrix}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\frac{\mathbf{f}_{k+1}-\mathbf{f}_k}{T_s}$$

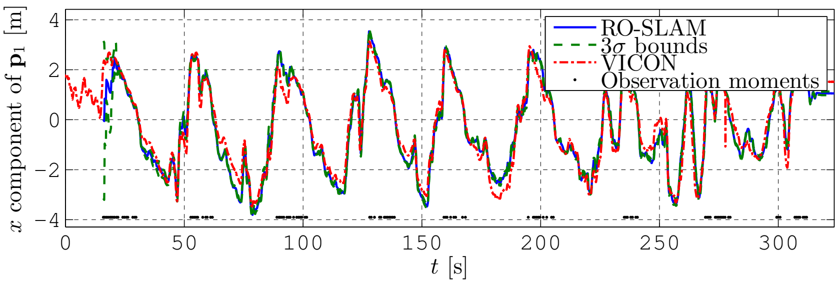
where $\mathbf{H} \in \mathbb{R}^{3N_k \times 6}$ is composed of $\mathbf{H}_i = \begin{bmatrix}-\mathbf{I}_3 & \mathbf{S}\left(\mathbf{f}_k^{(i)}\right)\end{bmatrix}$. As seen in Figure 4.7(a), to try to circumvent possible lack of features, the floor was covered with newspaper thus providing a good ground for feature detection.

It must be stressed that the experiments detailed in this section were designed as proof of concept, and as such, alternative sensors and processes for obtaining the linear velocity of the platform could have been employed. One possible example of alternative techniques are optical flow procedures, a common approach for velocity estimation [BSL+11].

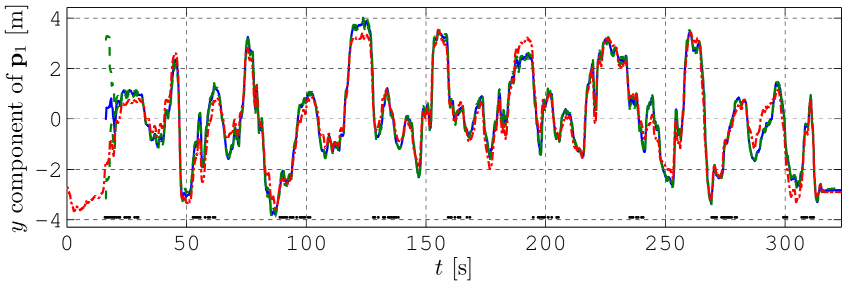
4.6.2 Results

The experiment here detailed consists of series of hand-driven circular-like laps of the quadrotor in a 6m×4m area covered by the Cricket constellation (see Figure 4.7). The run lasts for 5 minutes, comprising a total distance of around 90 meters at 0.270 m/s. The trajectory, shown in Figure 4.10 was intended to maximize the exposure to each of the beacons, as well as to provide sufficient excitation to the filter.

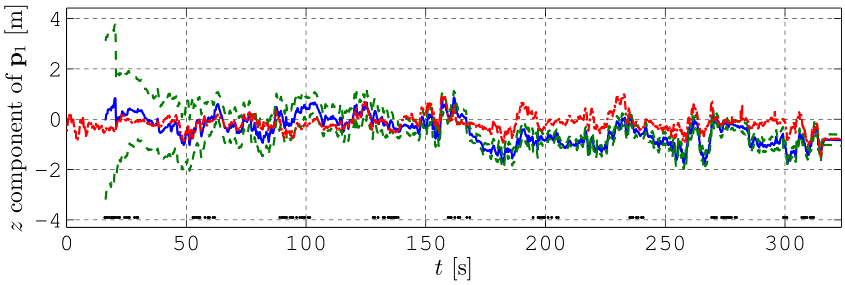
Figures 4.8 and 4.9 depict the estimated position (solid blue) with 3σ uncertainty bounds of two different landmarks in the body-fixed frame against the ground truth provided by *VICON* (dashed red) representing the best (landmark #1) and worst (landmark #7) estimation performance in this experiment. These quantities are accompanied by black dots indicating the observation moments, to provide better understanding of what is going on. As mentioned, each beacon emits once every 700 ms, and, due to the nature of the receiver, only one beacon pulse is received at a time. Furthermore, neither of the beacons nor receiver are omnidirectional, which means that there may occur long time intervals without any measurement from



(a) x component



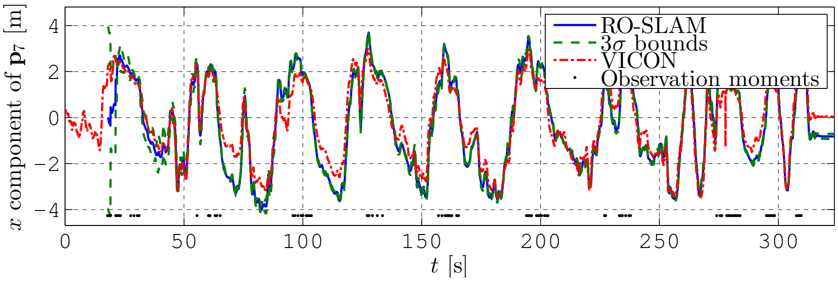
(b) y component



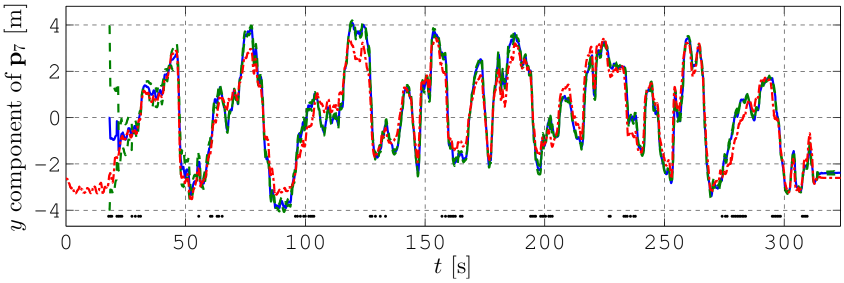
(c) z component

Figure 4.8: The sensor-based landmark #1 estimate against ground truth with 3σ uncertainty bounds and observation instants.

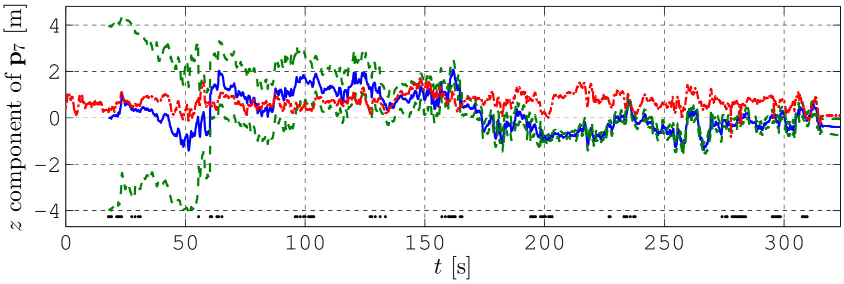
one beacon. The black dots in both Figure 4.8 and Figure 4.9 show the moments were each landmark is observed. As explained in the introductory chapter, the initialization of the landmarks is one of the more challenging issues in RO-SLAM procedures. In this work, however, the global convergence results imply that this issue is solved as whichever the initial guess the filter will converge. Figures 4.8 and 4.9, where each landmark was initialized at a random point in the sphere defined by the corresponding range measurement, depict exactly this. It can be seen that the convergence is very fast in the horizontal plane, represented by Figures 4.8(a) and 4.8(b), for the first landmark, and Figures 4.9(a) and 4.9(b) for the seventh. Moreover, after converging, the estimation is very close to ground truth in the first landmark and with higher error in the seventh. However, in the vertical axis, the estimation is much worse, and the convergence is also slower, as it can be seen in Figures 4.8(c) and 4.9(c) that represent the quantities associated with the vertical coordinate. That is due to the less rich trajectory in that axis, as Figure 4.10 shows. The visual odometry procedure employed is quite noisy, and



(a) x component



(b) y component



(c) z component

Figure 4.9: The sensor-based landmark #7 estimate against ground truth with 3σ uncertainty bounds and observation instants.

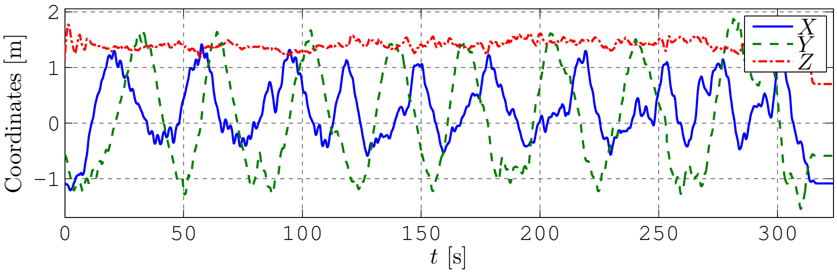


Figure 4.10: The position of the vehicle in time.

as its measurements of the linear velocity are directly used in the dynamics matrix in (2.9) as if they were the true value, the noise can make that direction appear observable, even if the information is sparse. Furthermore, this noise has an influence on the consistency of a landmark estimate whenever that landmark is being propagated in open loop.

Finally, an example of the estimated map in the body-fixed frame is presented in Figure

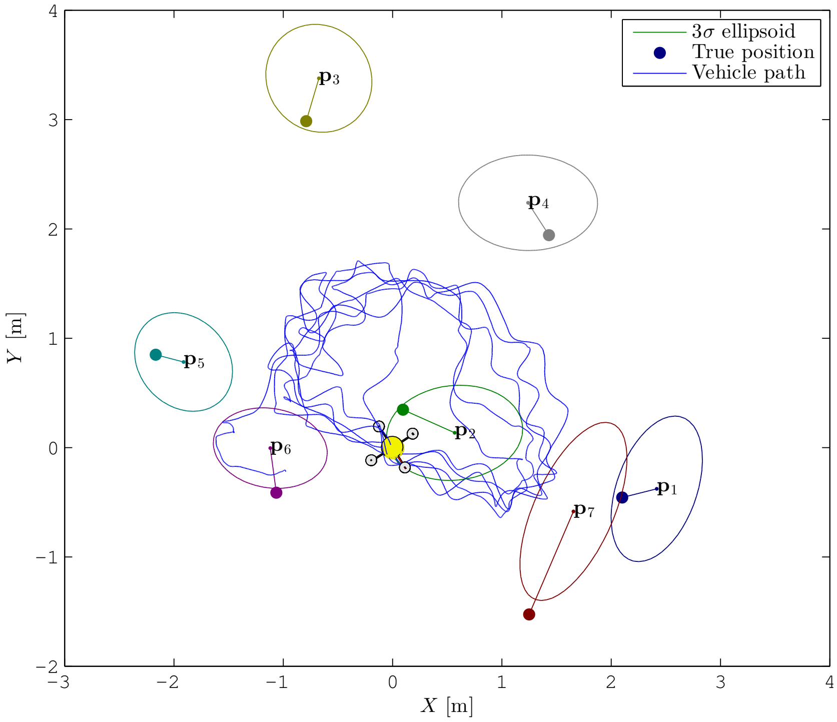


Figure 4.11: Top view of the estimated sensor-based map at $t = 175$ s.

4.11. The top view of the sensor-based map is shown along with the true landmark positions and the vehicle path rotated and translated to the body-fixed frame. The coloured ellipses represent orthogonal cross-sections of 3σ uncertainty ellipsoids, i.e., the estimation uncertainty, and the small circles mark the true landmark positions.

These experiments show the good performance of the proposed algorithm in realistic conditions, especially in the horizontal variables. The filter has some problems in the vertical coordinates due to the less rich velocity profile and noisy visual odometry measurements, although in Section 4.5 a proper trajectory was designed and the algorithm was shown to behave well. Therefore, these experiments underpin the need for appropriate trajectories.

4.7 Conclusions

This chapter presented a novel sensor-based range-only simultaneous localization and mapping filter with globally exponentially stable error dynamics. This was achieved through state augmentation of a nonlinear system, which, along with the disposal of the non-visible landmarks, enabled regarding the resulting system as linear time-varying. The work focused on the observability analysis of the resulting system, providing theoretical observability guarantees, and equivalence between the systems used in each step of the analysis. The theoretical results include the derivation of necessary and sufficient conditions for observability, stability and convergence of the algorithm, establishing a constructive basis for trajectory design. These results were followed by the design of a Kalman filter with globally exponentially stable error dynamics. Simulations allowed the validation of the results, and real world experiments showed also the good performance of the proposed algorithm in realistic conditions while demonstrating the need for a properly designed trajectory. This is due to the fact that a landmark becomes unobservable when, during the observation time, the vehicle does not respect the observability conditions, i.e., for this three-dimensional range-only problem, when the vehicle does not consistently move in all the directions. Even though this may seem a fairly restrictive imposition, it is inherent to any range-only algorithm. With aerial or underwater vehicles it is simple to accommodate this condition on trajectory design. However, for ground vehicles, which usually only move in two dimensions, either the beacons must be all positioned at the same height of the receiver and a 2-D version of the algorithm used, or, in order to use the 3-D version of the algorithm the heights of the beacons must be known. In this case, the vehicle only needs to consistently move in 2-D.

5

SIMULTANEOUS LOCALIZATION AND MAPPING IN SENSOR NETWORKS

THIS chapter presents the design, analysis, and validation of a globally exponentially stable (GES) filter for three-dimensional (3-D) range-only simultaneous localization and mapping in sensor networks with moving object tracking. Following the successful application of the sensor-based framework and the methodology that are at the core of this thesis to the range-only problem in the previous chapter, the work in the current chapter extends it by introducing beacon-to-beacon communication, as well as the possibility of having mobile nodes in the network. This is done by building on the systems designed in Chapter 4 and introducing a motion model for the mobile node or target and then exploiting a state augmentation that allows the proposed formulation to be considered as linear time-varying without linearizing the original nonlinear systems. Constructive observability results can then be established, leading naturally to the design of a Kalman Filter with globally exponentially stable (GES) error dynamics. These results also provide valuable insight on the motion planning of the vehicle. Simulation results demonstrate the good performance of the algorithm and help validate the theoretical results presented, as well as illustrate the necessity of a proper trajectory.

5.1 Introduction

SLAM in sensor networks is a relevant problem, since, in many cases, nodes in a sensor network do not have sufficient connectivity due to a lack of neighbours. In that situation, a vehicle with ego-motion sensors moving in the environment and connecting to different nodes in time can largely improve the overall localization scheme in the network [DS12]. If there are no other mobile nodes in the network, this is also an interesting scenario for range-only simultaneous localization and mapping (RO-SLAM). However, if there are nodes in the network attached to, for example, a person performing an inspection on site, then a moving object tracking scheme is necessary.

In this chapter it is proposed a fusion of the concepts of SLAM, localization of sensor networks, and moving object tracking in a RO-SLAM framework. With this approach, a mapping of the sensor network is achieved aided by the motion of the vehicle. Taking advantage of the sensor-based approach that allows the whole network to be observable without anchors (as opposed to the absolute or inertial approach that requires anchors [AM10]), this chapter introduces a novel algorithm for RO-SLAM in sensor networks with moving object tracking (SN-SLAMMOT) that eliminates the initialization problem through the establishment of global convergence results with a 3-D sensor-based formulation that avoids the representation of the pose of the vehicle in the state, as it becomes deterministic and available by construction. This SN-SLAMMOT solution builds on the version that presented only RO-SLAM capabilities described in Chapter 4, and is grounded on the source-localization algorithm proposed in [BSO11b] and the long baseline (LBL) navigation algorithm presented in [BSO14], as the global convergence results are achieved through similar state augmentations.

The main contributions of this chapter are the design, analysis, and experimental validation of a 3-D simultaneous localization and mapping in sensor networks with moving object tracking (SN-SLAMMOT) algorithm that i) has globally exponentially stable (GES) error dynamics; ii) resorts to the exact linear and angular motion kinematics; iii) uses as odometry-like measurements the linear and angular velocities; iv) solves a nonlinear problem with no linearizations whatsoever; and v) builds on the well-established linear time-varying Kalman filtering theory.

Chapter Structure

The chapter is organized as follows: in Section 5.2, the problem addressed in this chapter is stated; the systems to be filtered are designed and their observability analysed in Section 5.3 and filter implementation issues are detailed in Section 5.4. The results of simulation experiments are presented in Sections 5.5, and, finally, Section 5.6 addresses some concluding remarks.

5.2 SLAM in sensor networks

5.2.1 Problem statement

Consider the existence of a sensor network composed of several static nodes, or beacons, and two mobile nodes/agents, one of them being the main vehicle and another an agent or target. One of the agents is a full-fledged vehicle with several sensing capabilities, such as ranging, as well as angular and linear odometry. The other agent is merely passive, i.e., it does not measure any motion-related variables and its distance to other nodes is only measured by them. This could represent a person or another vehicle in an enemy setting, within many other possibilities. The sensor network distributes among its nodes the range measurements acquired by each of its nodes. Therefore, each beacon can maintain an average filter for the range from its static neighbours, thus reducing the uncertainty of that information. The vehicle is the only node

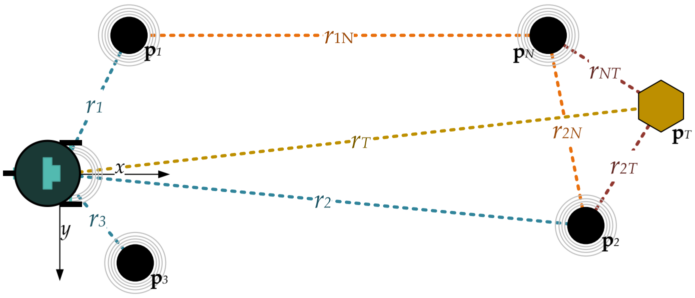


Figure 5.1: Schematic of a possible network scenario. Circles indicate static nodes, the wheeled circle is the vehicle and the hexagon represents the target. The movement of the vehicle allows for the connection of isolated nodes or groups of nodes, which otherwise would not communicate with the rest of the network.

with real computing capabilities, and therefore the localization of itself and the remaining network is performed in the vehicle. This renders the problem at hand a simultaneous localization and mapping/tracking problem. One advantage of looking at this in a SLAM framework is that with the regular algebraic solutions for sensor-network localization data loss is a very relevant issue as is network connectivity. With a SLAM filter, as long as certain observability conditions are fulfilled, the whole network can be localized in time.

Recall the scenario described in Sections 2.1 and 4.2. Further consider the existence of a second agent, henceforth denominated the target, which is denoted as ${}^{E}\mathbf{p}_T(t) \in \mathbb{R}^3$ or $\mathbf{p}_T(t) \in \mathbb{R}^3$ depending on whether it is expressed in frame $\{E\}$ or $\{B\}$.

Throughout the network, the following range measurements are available: beacon-to-beacon (B2B), denoted as $r_{ij}(t) > 0$ for all $i < j$ and $i, j \in \mathcal{M}$, vehicle-to-beacon (V2B), denoted as $r_i(t) > 0$ for all $i \in \mathcal{M}$, vehicle-to-target (V2T), or $r_T(t) > 0$, and beacon-to-target (B2T), denoted as $r_{iT}(t) > 0$ $i \in \mathcal{M}$. The vehicle has available its linear and angular velocities as expressed in its own frame, respectively $\mathbf{v}(t) \in \mathbb{R}^3$ and $\boldsymbol{\omega}(t) \in \mathbb{R}^3$. This can be seen as three-dimensional odometry.

5.2.2 Filtering concept

The main idea behind this chapter is the sensor-based approach that allows to solve one of the main nonlinearities that affects the SLAM problem: the presence of the vehicle pose in the state. In this particular formulation of SLAM, with range-only measurements, there are further nonlinearities that impair the development of filtering solutions with convergence guarantees, namely the nonlinear relation between the output (the ranges) and the state (the beacons positions). This nonlinearity can be bypassed through the use of state augmentation, as was successfully done in [BSO11b] and in the SLAM context in previous chapters. In those works, the state augmentation and the sensor-based approach allowed to design systems that resemble linear time-varying systems, and therefore permit the usage of the Kalman filter with its convergence and stability properties. Following the same reasoning, this work aims at achieving global convergence and stability results in a filter for the problem described above.

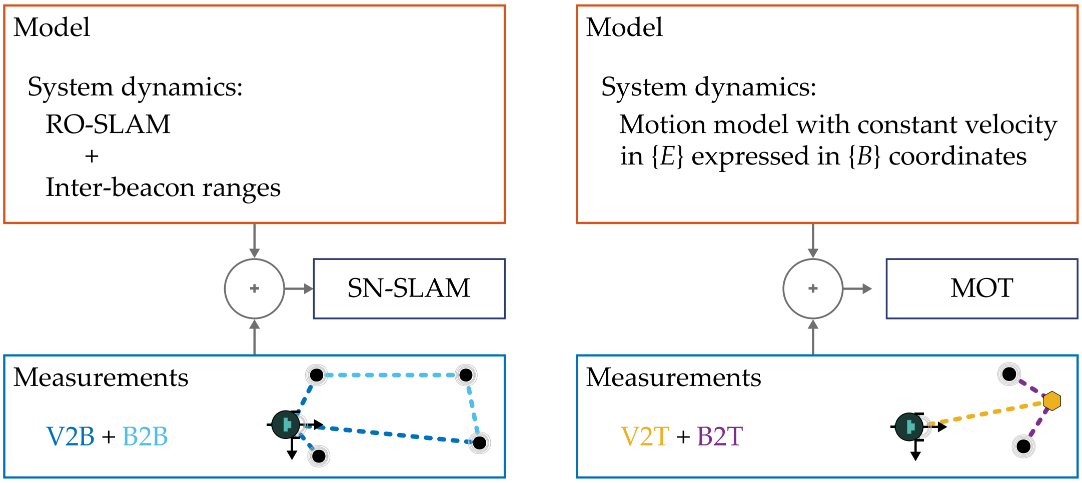

Model
System dynamics:
RO-SLAM
+
Inter-beacon ranges
SN-SLAM
Measurements
V2B + B2B
Model
System dynamics:
Motion model with constant velocity in {E} expressed in {B} coordinates
MOT
Measurements
V2T + B2T

Figure 5.2: The concepts behind the algorithm proposed in this chapter, showcasing the strategy used for design, as well as the relation with Chapter 4.

SLAM in Sensor Networks The problem is modelled as a nonlinear dynamical system, expressed in local coordinates, fixed to the vehicle similarly to the RO-SLAM problem in Chapter 4. In fact, if there are no agents besides the vehicle, the only difference to RO-SLAM is the presence of sensor-to-sensor ranges, and the fact that each beacon is capable of measuring distances and save them (see Figure 5.2). This communication allows for the nodes without direct connectivity to a given node to have information about that node with a few timesteps of delay – allowing for multi-hop filtering when that information reaches the vehicle. Hence, the nonlinear system that encodes that part of the problem is

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t))\,\mathbf{p}_i(t) - \mathbf{v}(t), \forall i \in \mathcal{M} \\ \dot{\mathbf{v}}(t) = \mathbf{0} \\ \mathbf{y}_V(t) = \mathbf{v}(t) \\ r_j(t) = \|\mathbf{p}_j(t)\|, \forall j \in \mathcal{M}_O \\ r_{kl}(t) = \|\mathbf{p}_k(t) - \mathbf{p}_l(t)\|, \forall k,l \in \mathcal{M}_I \end{cases}, \tag{5.1}$$

where the linear velocity is directly measured, even though it is modelled as constant for filtering purposes. The set $\mathcal{M}_O := \{1,\ldots,N_O\}$ represents the beacons directly observed by the vehicle (in contrast to the set $\mathcal{M}_U := \{N_O+1,\ldots,m\}$), and the set $\mathcal{M}_I$ contains the beacons indirectly observed (note that, in general, $\mathcal{M}_O \cap \mathcal{M}_I \neq \emptyset$).

Tracking moving agents Given that the second agent is passive in the network, a model must be assumed for its motion. For simplicity and to allow the use of an LBL-like structure, the constant velocity model is used in the Earth-fixed frame, resulting in

$$\begin{cases} \frac{d}{dt}\Big(\mathbf{R}(t)\mathbf{p}_T(t)\Big) = {}^E\mathbf{v}_T(t) - \mathbf{R}(t)\mathbf{v}(t) \\ {}^E\dot{\mathbf{v}}_T(t) = \mathbf{0} \\ r_T(t) = \|\mathbf{p}_T(t)\| \\ r_{iT}(t) = \|\mathbf{p}_i(t) - \mathbf{p}_T(t)\| \end{cases}, \tag{5.2}$$

where ${}^E\mathbf{v}_T(t) \in \mathbb{R}^3$ is the velocity of the target expressed in the Earth-fixed frame, and $\mathcal{M}_T := \{1, \ldots, N_T\}$ is the set of beacons whose field-of-view includes the target.

In opposition to the beacons, ranges to the target can only be used in the instant they are acquired, i.e., only the range to the target whose originating beacons are directly seen by the robot matter, as multi-hop ranges are older and cannot be easily inserted in a forward-time filter. A possibility is their use in a smoother in post-processing to better recover the full trajectory of the target.

The ultimate idea behind this chapter is to design a linear time-varying system that mimics the nonlinear dynamics derived in this section, while allowing for a LTV Kalman filter to serve as the filtering engine for this problem, taking advantage of its good stability and convergence properties.

5.3 System design

This section presents the design of new systems that can be regarded as LTV for observability purposes, while mimicking the underlying nonlinear systems. Without loss of generality, the two systems in analysis will be treated separately, for the sake of clarity. Note that the first system is independent of the second, while the second relies on the first. For all the derivations to come, the following assumption is necessary.

Assumption 5.1. *Any measured range from any node to any other node of the network is upper and lower bounded by some positive constants,*

$$R_m < r_i(t), r_{ij}(t), r_T(t), r_{iT}(t) < R_M, \forall i, j \in \mathcal{M}, \forall t > t_0.$$

Note that, as in similar assumptions in previous chapters, the actual values of the bounds are not relevant for design/implementation. Furthermore, in practice, two nodes will never coincide with each other, nor will they be arbitrarily distant and hence the measured ranges will always be bounded.

5.3.1 SLAM in Sensor Networks

The system introduced in the previous section is evidently nonlinear and, as such, cannot be used directly in a Kalman filter as intended. For that reason, and given that the nonlinearity presents itself on the output equations, a strategy of state augmentation to obtain a linear relation between the system state and output is proposed. Apart from the B2B ranges, the system (5.1) is in every way similar to the underlying system designed in Chapter 4, and, therefore, the idea of adding the ranges to beacons to the system state proposed there will be used here as well.

In order to be able to incorporate the beacon-to-beacon ranges in this system, consider the expansion of the range from beacon i to beacon j, given by

$$r_{ij}(t) = \frac{1}{r_{ij}(t)}\left(\mathbf{p}_i^T(t)\mathbf{p}_i(t) + \mathbf{p}_j^T(t)\mathbf{p}_j(t) - 2\mathbf{p}_i^T(t)\mathbf{p}_j(t)\right). \tag{5.3}$$

Note that the first two parcels in (5.3) are in fact the square of the ranges to beacons i and j and therefore could be substituted by those quantities whenever available. However, to allow for the use of these ranges even when one or more of the corresponding beacons are not directly visible, the squared ranges are added to the state of the system, in contrast to the augmented state in (4.2) which consisted of the ranges. The last parcel, the dot product between the two beacons, is also nonlinear on the state of the original system (5.1), and therefore it is also added to the state.

To summarize all this information, the new state is $\mathbf{x}(t) := \left[\mathbf{x}_M^T(t)\ \mathbf{x}_V^T(t)\ \mathbf{x}_R^T(t)\ \mathbf{x}_D^T(t)\right]^T$, where each component is given by

$$\begin{cases} \mathbf{x}_M(t) := \{\mathbf{p}_i(t), \forall i \in \mathcal{M}\} \in \mathbb{R}^{3N} \\ \mathbf{x}_V(t) := \mathbf{v}(t) \in \mathbb{R}^3 \\ \mathbf{x}_R(t) := \{\|\mathbf{p}_i(t)\|^2, \forall i \in \mathcal{M}\} \in \mathbb{R}^N \\ \mathbf{x}_D(t) := \{\mathbf{p}_i^T(t)\mathbf{p}_j(t), \forall i,j \in \mathcal{M}, i < j\} \in \mathbb{R}^{C_2^N} \end{cases}. \tag{5.4}$$

The augmented system dynamics are

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}(t, \mathbf{y}_v(t))\mathbf{x}(t) \\ \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) \end{cases}, \tag{5.5}$$

where the dynamics matrix is

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{A}_M(t) & \mathbf{A}_{MV} & \mathbf{0}_{3N\times N} & \mathbf{0}_{3N\times C_2^N} \\ \mathbf{0}_{3\times 3N} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times C_2^N} \\ -2\mathbf{A}_{RM}(t) & \mathbf{0}_{N\times 3} & \mathbf{0}_{N\times N} & \mathbf{0}_{N\times C_2^N} \\ -\mathbf{A}_{DM}(t) & \mathbf{0}_{C_2^N\times 3} & \mathbf{0}_{C_2^N\times N} & \mathbf{0}_{C_2^N\times C_2^N} \end{bmatrix},$$

with $\mathbf{A}_M(t) := -\operatorname{diag}(\mathbf{S}(\boldsymbol{\omega}(t)), \ldots, \mathbf{S}(\boldsymbol{\omega}(t))))$, $\mathbf{A}_{MV} := \left[-\mathbf{I}_3\ \cdots\ -\mathbf{I}_3\right]^T$, $\mathbf{A}_{RM}(t) := \operatorname{diag}\left(\mathbf{y}_V^T(t), \ldots, \mathbf{y}_V^T(t)\right)$, and $\mathbf{A}_{DM}(t)$ is the matrix that encodes all the possible combinations of two beacons, such that $\frac{d}{dt}\mathbf{p}_i(t)^T\mathbf{p}_j(t) = -\mathbf{y}_V^T(t)\left(\mathbf{p}_i(t) + \mathbf{p}_j(t)\right)$ for all $i < j$, $i, j \in \mathcal{M}$. The output matrix is

$$\mathbf{C}(t) = \begin{bmatrix} \mathbf{0}_{3\times 3N} & \mathbf{I}_3 & \mathbf{0}_{3\times N_O} & \mathbf{0}_{3\times N_U} & \mathbf{0}_{3\times C_2^N} \\ \mathbf{0}_{N\times 3N} & \mathbf{0}_{N\times 3} & \mathbf{C}_R(t) & \mathbf{0}_{N_O\times N_U} & \mathbf{0}_{3\times C_2^N} \\ \mathbf{0}_{C_2^{N_I}\times 3N} & \mathbf{0}_{C_2^{N_I}\times 3} & \mathbf{C}_{DR}(t) & \mathbf{0}_{C_2^{N_I}\times N_U} & -2\mathbf{C}_D(t) \end{bmatrix},$$

where the first N_O range states were separated from the N_U non-visible ones. Its components are $\mathbf{C}_R(t) := \operatorname{diag}\left(r_1^{-1}(t),\ldots,r_{N_O}^{-1}(t)\right)$, $\mathbf{C}_D(t) := \left[\operatorname{diag}\left(r_{12}^{-1}(t),\ldots,r_{(N_I-1)N_I}^{-1}(t)\right)\ \mathbf{0}_{C_2^{N_I}\times(C_2^N-C_2^{N_I})}\right]$, and $\mathbf{C}_{DR}(t)$ is the matrix that encodes all the possible combinations of two observed beacons, such that $r_{ij}(t) = r_{ij}^{-1}(t)\left(x_{R_i}+x_{R_j}-2x_{D_{ij}}\right)$ for all $i<j$, $i,j\in\mathcal{M}_I$.

5.3.2 Tracking

The target tracking part of the system is more complicated to tackle than the previous one. Let $\mathbf{z}_{T_P}(t) = \mathbf{R}(t)\mathbf{p}_T(t)$ and $\mathbf{z}_{T_V}(t) := {}^E\mathbf{v}_T(t)$ be its states. Consider the two types of ranges to the target that the vehicle has access to: vehicle-to-target and beacon-to-target. Following the reasoning in [BSO14], a series of new states are added to yield a linear-like structure to the output-state relation. Take, for example, the vehicle-to-target range, $r_T(t) = \|\mathbf{z}_{T_P}(t)\|$. If it is added to the state, the output will be linear-like with respect to the state, with dynamics given by

$$\frac{d}{dt}(r_T^2(t)) = 2\mathbf{z}_{T_V}^T(t)\mathbf{z}_{T_P} - 2{}^E\mathbf{v}^T(t)\mathbf{z}_{T_P}(t).$$

The first parcel can then be added to a new state, $z_{T_{VP}}(t) := \mathbf{z}_{T_V}^T(t)\mathbf{z}_{T_P}$, yielding

$$\dot{z}_{T_{VP}}(t) := \mathbf{z}_{T_V}^T(t)\mathbf{z}_{T_V}(t) - {}^E\mathbf{v}^T(t)\mathbf{z}_{T_V}(t),$$

and denoting the first parcel as $z_{T_{VV}}(t) := \mathbf{z}_{T_V}^T(t)\mathbf{z}_{T_V}(t)$, its derivative is

$$\dot{z}_{T_{VV}}(t) = 0.$$

To tackle the inclusion of the beacon-to-target ranges in the system, consider its expansion as in (5.3),

$$\begin{aligned} r_{iT}^2(t) &= \|\mathbf{R}(t)\mathbf{p}_i(t) - \mathbf{z}_{T_P}(t)\|^2 \\ &= x_{R_i}(t) + z_{T_R}(t) - 2\mathbf{z}_{T_P}^T(t)\mathbf{R}(t)\mathbf{p}_i(t), \end{aligned}$$

where $x_{R_i}(t)$ and $z_{T_R}(t)$ are replacing the corresponding squared ranges. The last parcel can then be added as a new state, $z_{T_{Pi}}(t)$, with dynamics given by

$$\dot{z}_{T_{Pi}}(t) = \mathbf{z}_{T_V}^T(t)\mathbf{R}(t)\mathbf{p}_i(t) - {}^E\mathbf{v}^T(t)\left(\mathbf{R}(t)\mathbf{p}_i(t) + \mathbf{z}_{T_P}(t)\right).$$

Once again, the first nonlinear parcel is replaced by a new state, $z_{T_{Vi}}(t)$, whose derivative is

$$\dot{z}_{T_{Vi}}(t) = -{}^E\mathbf{v}^T(t)\mathbf{z}_{T_V}(t).$$

The quantities introduced here can then be stacked in the target state,

$$\mathbf{z}_T(t) = \left[\mathbf{z}_{T_P}^T(t)\ \ \mathbf{z}_{T_V}^T(t)\ \ z_{T_R}(t)\ \ z_{T_{VP}}(t)\ \ z_{T_{VV}}(t)\ \ \mathbf{z}_{T_{PM}}(t)\ \ \mathbf{z}_{T_{VM}}(t)\right]^T, \tag{5.6}$$

where $\mathbf{z}_{T_{PM}}(t) \in \mathbb{R}^N$ and $\mathbf{z}_{T_{VM}}(t) \in \mathbb{R}^N$ are the stacking of all the $z_{T_{Pi}}(t)$ and $z_{T_{Vi}}(t)$ respectively.

It must be noted that although present in all the expressions in this subsection, the rotation matrix is not observed nor estimated. To avoid this issue, consider the state $\mathbf{x}_T(t) :=$

$\operatorname{diag}\left(\mathbf{R}^T(t), \mathbf{R}^T(t), 1, 1, 1, \mathbf{I}_N, \mathbf{I}_N\right)\mathbf{z}_T(t)$. The new augmented system that can be joined with (5.5) is

$$\begin{cases} \dot{\mathbf{x}}_T(t) = \mathbf{A}_T(t)\mathbf{x}_T(t) + \mathbf{A}_{TB}(t)\mathbf{x}(t) \\ \mathbf{y}_T(t) = \mathbf{C}_T(t)\mathbf{x}_T(t) + \mathbf{C}_{TB}(t)\mathbf{x}(t) \end{cases}, \tag{5.7}$$

where the dynamics matrix is composed of

$$\mathbf{A}_T(t) = \begin{bmatrix} -\mathbf{S}(\boldsymbol{\omega}(t)) & \mathbf{I}_3 & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times N} \\ \mathbf{0}_{3\times3} & -\mathbf{S}(\boldsymbol{\omega}(t)) & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times N} \\ -2\mathbf{y}_V^T & \mathbf{0}_{1\times3} & 0 & 2 & 0 & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times N} \\ \mathbf{0}_{1\times3} & -\mathbf{y}_V^T & 0 & 0 & 1 & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times N} \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & 0 & 0 & 0 & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times N} \\ -\mathbf{1}_N\mathbf{y}_V^T & \mathbf{0}_{N\times3} & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times N} & \mathbf{I}_N \\ \mathbf{0}_{N\times3} & -\mathbf{1}_N\mathbf{y}_V^T & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times1} & \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \end{bmatrix}$$

and

$$\mathbf{A}_{TB}(t) = \begin{bmatrix} \mathbf{0}_{3\times3N} & -\mathbf{I}_3 & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times C_2^N} \\ \mathbf{0}_{3\times3N} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times N} & \mathbf{0}_{3\times C_2^N} \\ \mathbf{0}_{1\times3N} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times C_2^N} \\ \mathbf{0}_{1\times3N} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times C_2^N} \\ \mathbf{0}_{1\times3N} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times C_2^N} \\ -\mathbf{A}_{RM}(t) & \mathbf{0}_{N\times3} & \mathbf{0}_{N\times N} & \mathbf{0}_{N\times C_2^N} \\ \mathbf{0}_{1\times3N} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times C_2^N} \end{bmatrix}.$$

Finally, the output matrices of this part of the system are

$$\mathbf{C}_T(t) = \begin{bmatrix} \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & r_T^{-1}(t) & 0 & 0 & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times N} \\ \mathbf{0}_{N_T\times3} & \mathbf{0}_{N_T\times3} & \mathbf{C}_{T_R}(t) & 0 & 0 & -2\mathbf{C}_{T_P}(t) & \mathbf{0}_{1\times N} \end{bmatrix}$$

and

$$\mathbf{C}_{TB}(t) = \begin{bmatrix} \mathbf{0}_{1\times3N} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times C_2^N} \\ \mathbf{0}_{N_T\times3N} & \mathbf{0}_{N_T\times3} & \mathbf{C}_{T_P}(t) & \mathbf{0}_{N_T\times C_2^N} \end{bmatrix},$$

where $\mathbf{C}_{T_R}(t) = \left[r_{1T}^{-1} \; \cdots \; r_{N_T T}^{-1}\right]^T$ and $\mathbf{C}_{T_P}(t) = \left[\operatorname{diag}\left(r_{1T}^{-1}, \ldots, r_{N_T T}^{-1}\right) \; \mathbf{0}_{N_T\times(N-N_T)}c\right]$.

These new systems, that mimic the dynamics of (5.1) and (5.2), are still nonlinear in form as, in fact, the time dependence of the dynamics and output matrices hides the dependence on the system output through the linear velocity of the vehicle and the measured ranges, two known signals. However, for that reason they are considered linear time-varying systems for observability purposes. Furthermore, the algebraic constraints that helped define $\mathbf{x}(t)$ and $\mathbf{x}_T(t)$ are not imposed anywhere in the system dynamics.

Note that in the systems designed here, it is assumed that the first N_O beacons are directly observed, the first N_I are within the field-of-view of each other, and the first N_T can measure ranges to the target. In practice, any beacon can be in either sets.

5.3.3 Observability analysis

The augmented system that does not include the dynamics of the target is, as mentioned, very similar to the augmented system derived in Chapter 4. In fact, the inclusion of the beacon-to-beacon ranges does not lighten the observability requirements established therein. Therefore, this first part of the observability analysis is quite straightforward.

In this analysis, the invisible beacons, ranges, and dot products are removed from the states, as they cannot be observable. Therefore, the set $\mathcal{M}_U$ is considered to be empty and the sets $\mathcal{M}_I$ and $\mathcal{M}_T$ are contained in $\mathcal{M}_O$.

Theorem 5.1 (Observability of the SN-SLAM LTV system)**.** *Consider the system* (5.5) *and let* $\mathcal{T} := [t_0, t_f]$. *If there exist three instants* $\{t_1, t_2, t_3\} \in \mathcal{T}$ *such that the linear velocity of the vehicle expressed in the Earth-fixed frame is linearly independent in those instants, i.e.,* $\det\left(\begin{bmatrix} {}^E\mathbf{v}(t_1) & {}^E\mathbf{v}(t_2) & {}^E\mathbf{v}(t_3)\end{bmatrix}\right) \neq 0$, *then the system is observable in the sense that, given the system output* $\{\mathbf{y}(t), t \in \mathcal{T}\}$, *the initial condition* $\mathbf{z}(t_0)$ *is uniquely defined.*

Proof. A system is observable if its initial condition is uniquely defined given the system output and input. Therefore, if it is shown that it is possible to recover the initial state from the outputs of the system, the proof is done. It was shown in Theorem 4.1 that the system with states $\mathbf{x}_M(t)$, $\mathbf{x}_V(t)$ and $\mathbf{x}_R(t)$, and outputs $\mathbf{y}_V(t)$ and $\mathbf{y}_R(t)$ is observable in the conditions of this theorem. For the observability of the remaining states, $\mathbf{x}_D(t)$, consider the time evolution of the corresponding output, for all $i < j$, $i, j \in \mathcal{M}_I$

$$r_{ij}^2(t) = x_{R_i}(t_0) + x_{R_j}(t_0) - 2x_{D_{ij}}(t_0). \tag{5.8}$$

As the first two parcels are given by the outputs $r_i(t_0)$ and $r_j(t_0)$, the last parcel is directly recoverable too. Hence, given the system output and input, all the initial states are uniquely defined if the conditions of the theorem hold, thus concluding the proof. ☐

The following theorem is necessary to establish a clear relation between the augmented system (5.5) that mimics the nonlinear dynamics with the actual nonlinear system (5.1).

Theorem 5.2 (Equivalence of the SN-SLAM LTV and nonlinear systems)**.** *Consider the LTV system* (5.5) *and the original nonlinear system* (5.1). *If the conditions of Theorem 5.1 hold, then the state of the original nonlinear system and that of the LTV system are the same and uniquely defined, provided that the invisible beacons are discarded. Furthermore the constraints expressed by* (5.4) *become naturally imposed by the dynamics.*

Proof. Similarly to the proof of Theorem 5.1, the states associated with the positions of the beacons and the velocity are shown to be equivalent by Theorem 4.1, while establishing the algebraic constraint $x_{R_i}(t) = \|\mathbf{x}_{M_i}(t)\|^2$ for all $i \in \mathcal{M}_O$. For the dot-product states, consider first the transition expression for the position of a beacon

$$\mathbf{R}(t)\mathbf{p}_i(t) = \mathbf{R}(t_0)\mathbf{p}_i(t_0) - \mathbf{R}^{[0]}(t, t_0)\mathbf{v}(t_0), \tag{5.9}$$

where $\mathbf{R}^{[0]}(t, t_0) = \int_{t_0}^{t} \mathbf{R}(\tau)\mathrm{d}\tau$. This can be substituted in $r_{ij}^2(t) = \|\mathbf{p}_i(t) - \mathbf{p}_j(t)\|^2$, yielding

$$r_{ij}^2(t) = \|\mathbf{R}(t)\mathbf{p}_i(t)\|^2 + \|\mathbf{R}(t)\mathbf{p}_i(t)\|^2 - 2\left(\mathbf{R}(t)\mathbf{p}_i(t)\right)^T \mathbf{R}(t)\mathbf{p}_j(t).$$

Comparing this with (5.8) yields $x_{D_{ij}}(t_0) = \left(\mathbf{R}(t)\mathbf{p}_i(t)\right)^T \mathbf{R}(t)\mathbf{p}_j(t)$, thus concluding the proof. □

The tracking of a target with unknown motion in a SLAM framework is a more complex problem than simply mapping, and therefore it will impose stronger requirements on the motion of the vehicle. The following theorem addresses this issue, but requires the definition of two functions of the linear velocity, $\boldsymbol{\mathcal{V}}^{[0]}(t,t_0) = {}^E\mathbf{v}(t) - {}^E\mathbf{v}(t_0)$ and $\boldsymbol{\mathcal{V}}^{[1]}(t,t_0) = \int_{t_0}^{t} {}^E\mathbf{v}(\tau)\mathrm{d}\tau + (t - t_0){}^E\mathbf{v}(t)$, each composed of three components $\mathcal{V}_1^{[.]}(t,t_0)$, $\mathcal{V}_2^{[.]}(t,t_0)$, and $\mathcal{V}_3^{[.]}(t,t_0)$. In order to write the transition equations for the state of the system, a few other definitions are needed. They are summarized here,

$$\begin{cases} \mathbf{v}^{[n]}(t,t_0) = \displaystyle\int_{t_0}^{t} (\tau - t_0)^{n\,E}\mathbf{v}^T(\tau)\mathrm{d}\tau \\ \mathbf{v}^{[1,n]}(t,t_0) = \displaystyle\int_{t_0}^{t}\int_{t_0}^{\tau_2} \mathbf{v}^{[n]}(\tau_1,t_0)\mathrm{d}\tau_1 \ldots \mathrm{d}\tau_2 \\ \mathbf{v}_r(t,t_0) = \displaystyle\int_{t_0}^{t} {}^E\mathbf{v}^T(\tau)\mathbf{R}^{[0]}(\tau,t_0)\mathrm{d}\tau \\ \mathbf{v}^*(t,t_0) = \mathbf{v}^{[1,0]}(t,t_0) + \mathbf{v}^{[1]}(t,t_0) \end{cases}.$$

To simplify the analysis, the set $\mathcal{M}_I$ is assumed to be empty, i.e., there are no B2B ranges. Further consider the Lyapunov transformation (see Definition A.1 in Appendix A) $\mathbf{z}_F(t) = \mathbf{T}(t)\mathbf{x}_F(t)$, where

$$\mathbf{T}(t) = \operatorname{diag}\left(\mathbf{R}(t),\ldots,\mathbf{R}(t),\mathbf{I}_3,\mathbf{I}_N,\mathbf{R}(t),\mathbf{R}(t),1,1,1,\mathbf{I}_N,\mathbf{I}_N\right),$$

which is a concatenation of the transformation used in Chapter 4 and the inverse transformation that lead to (5.7). The transformed full system dynamics are given by

$$\begin{cases} \dot{\mathbf{z}}_F(t) = \boldsymbol{\mathcal{A}}(t)\mathbf{z}_F(t) \\ \mathbf{y}_F(t) = \mathbf{C}_F(t)\mathbf{z}_F(t) \end{cases}, \tag{5.10}$$

where $()_F(t)$ quantities are the concatenation of $(.)(t)$ and $(.)_T(t)$, and the dynamics matrix is $\boldsymbol{\mathcal{A}}(t) = \left[\mathbf{T}(t)\mathbf{A}_F(t) + \dot{\mathbf{T}}(t)\right]\mathbf{T}^{-1}(t)$.

Theorem 5.3 (Observability of the SN-SLAMMOT LTV system)**.** *The systems* (5.5) *and* (5.7) *are observable if the conditions of Theorem* 5.1 *hold, if the functions* $\mathcal{V}_1^{[0]}(t,t_0)$, $\mathcal{V}_2^{[0]}(t,t_0)$, $\mathcal{V}_3^{[0]}(t,t_0)$, $\mathcal{V}_1^{[1]}(t,t_0)$, $\mathcal{V}_2^{[1]}(t,t_0)$, *and* $\mathcal{V}_3^{[1]}(t,t_0)$ *are linearly independent in* $\mathcal{T}$ *and at least one B2T range is available three times in* $\mathcal{T}$. *If there are no B2T ranges, the* $\mathcal{V}$ *functions must also be linearly independent to* $(t - t_0)$.

Proof. If the conditions of this theorem hold, the conditions of Theorem 5.1 also hold, and therefore the first part of the joint system is rendered observable. Therefore, only the tracking state and observations will be addressed in this proof.

The proof will focus on the system (5.10), which is related to the systems (5.7) and (5.5) through a Lyapunov transformation, which preserves observability properties. Suppose that the system (5.7) is not observable. In that case, the associated observability Gramian $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is singular, which implies

$$\underset{\substack{\mathbf{c} \in \mathbb{R}^{n_Z} \\ \|\mathbf{c}\| = 1}}{\exists} : \quad \mathbf{c}^T \boldsymbol{\mathcal{W}}(t_0, t_f)\mathbf{c} = \int_{t_0}^{t_f} \|\mathbf{f}(\tau, t_0)\|^2 \mathrm{d}\tau = 0,$$

where $\mathbf{f}(\tau, t_0) = \mathbf{C}_F(\tau)\boldsymbol{\phi}(\tau, t_0)\mathbf{c}$ and $\mathbf{c}$ is

$$\mathbf{c} = \left[\mathbf{c}_M^T \ \mathbf{c}_V^T \ \mathbf{c}_R^T \ \mathbf{c}_{T_P}^T \ \mathbf{c}_{T_V}^T \ c_{T_R} \ c_{T_{VP}} \ c_{T_{VV}} \ \mathbf{c}_{T_{PM}}^T \ \mathbf{c}_{T_{VM}}^T\right]^T.$$

$\boldsymbol{\phi}(\tau, t_0)$ is the full system transition matrix, computed using the Peano-Baker equation or by solving $\boldsymbol{\phi}(\tau, t_0)\mathbf{z}_F(t_0) = \mathbf{z}_F(t_0) + \int_{t_0}^{\tau} \boldsymbol{\mathcal{A}}_F(\tau)\mathbf{z}_F(\sigma)\mathrm{d}\sigma$. If the Gramian is singular, then $\mathbf{f}(\tau, t_0)$ and its derivative must also be zero for all $\tau \in \mathcal{T}$. From Theorem 5.1, it is known that $\mathbf{c}_M$, $\mathbf{c}_V$, and $\mathbf{c}_R$ are all zero or the conditions of the theorem cannot hold. This comes from the first $3 + 3N_O$ rows of $\mathbf{f}(\tau, t_0)$. The remaining rows are

$$\mathbf{f}(\tau, t_0) = \begin{bmatrix} \mathbf{c}_V \\ \mathbf{f}_R(\tau, t_0) \\ f_T(\tau, t_0) \\ \mathbf{f}_{TM}(\tau, t_0), \end{bmatrix} \tag{5.11}$$

with

$$r_T(\tau) f_T(\tau, t_0) = -2\mathbf{v}^{[0]}(\tau, t_0)\mathbf{c}_{T_P} - 2\mathbf{v}^{*}(\tau, t_0)\mathbf{c}_{T_V} + c_{T_R} + 2(\tau - t_0)c_{T_{VP}} + (\tau - t_0)^2 c_{T_{VV}}, \tag{5.12}$$

and

$$r_{iT}(\tau) f_{TM_i}(\tau, t_0) = c_{T_R} + 2(\tau - t_0)c_{T_{VP}} + (\tau - t_0)^2 c_{T_{VV}} - 2c_{T_{Pi}} + 2(\tau - t_0)c_{T_{Vi}}.$$

Setting both expressions equal to zero at $t = t_0$ yields $c_{T_R} = c_{T_{Pi}} = 0$ for all $i \in \mathcal{M}_T$. As $(\tau - t_0)$ and $(\tau - t_0)^2$ are linearly independent functions, $\mathbf{f}_{TM_i}(\tau, t_0) = 0$ yields $c_{T_{VV}} = 0$ and $c_{T_{VP}} = c_{T_{Vi}}$. Consider now the first time derivative of (5.12) equal to zero,

$${}^{E}\mathbf{v}^T(\tau)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(\tau, t_0)\mathbf{c}_{T_V} - c_{T_{VP}} = 0.$$

When evaluated at $\tau = t_0$, this becomes ${}^{E}\mathbf{v}^T(t_0)\mathbf{c}_{T_P} = c_{T_{VP}}$, which, when substituted above, yields

$$\boldsymbol{\mathcal{V}}^{[0]}(\tau, t_0)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(\tau, t_0)\mathbf{c}_{T_V} = 0.$$

In the hypothesis of the proof, this implies that the functions $\mathcal{V}_i^{[0]}(\tau, t_0)$ and $\mathcal{V}_i^{[1]}(\tau, t_0)$, $i = 1, 2, 3$, are not linearly independent in $\mathcal{T}$, which contradicts the conditions of the theorem, or that $\mathbf{c}_{T_P} = \mathbf{c}_{T_V} = \mathbf{0}$ and $c_{T_{VP}} = c_{T_{Vi}} = 0$. Consider now that the set $\mathcal{M}_T$ is empty, and therefore, the states $\mathbf{z}_{T_{PM}}(t)$ and $\mathbf{z}_{T_{VM}}(t)$ disappear, as do the ranges $r_{iT}(t)$. In that case, all the information must be extracted from $r_T(t)$ and, consequently, $f_T(\tau, t)$. Hence, the first time derivative of (5.12) becomes

$$\boldsymbol{\mathcal{V}}^{[0]}(\tau, t_0)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(\tau, t_0)\mathbf{c}_{T_V} - (\tau - t_0)c_{T_{VV}} = 0.$$

This adds $(\tau - t_0)$ to the set of functions that cannot be linearly independent for the system to be not observable. Otherwise, $\mathbf{c}_{T_P}$, $\mathbf{c}_{T_V}$, $c_{T_{VP}}$, and $c_{T_{VV}}$ are all zero.

Therefore, if the system is not observable, the conditions of the theorem cannot hold, or, conversely, it the conditions of the theorem hold, the system (5.10) is observable. As it is related to the augmented system (5.7) through a Lyapunov transformation which preserves observability, this system is also observable, thus concluding the proof. □

Theorems 5.1 and 5.3 established conditions in which a reduced version of the augmented systems (5.5) and (5.7) are observable. As the discarded states are not observable, and do not influence the non discarded ones, these results also apply to the augmented systems even when there are non-visible beacons. However, this does not immediately establish an equivalence between the augmented systems and the original nonlinear ones. Furthermore, the state relations stated in 5.4 and those described in subsection 5.3.2 are not imposed. Theorem 5.2 addressed the first constraints, and the following result will address the latter.

Theorem 5.4 (Equivalence of the SN-SLAMMOT LTV and nonlinear systems)**.** *If the conditions of Theorem 5.3 hold, the systems* (5.2) *and* (5.7) *are equivalent, when discarding the non-visible beacons and respective ranges, and:*

(i) the state of the original nonlinear system and that of the LTV system are the same and uniquely defined, provided that the invisible beacons are discarded. Furthermore the constraints described in subsection 5.3.2 become naturally imposed by the dynamics; and

(ii) a state observer with uniformly globally exponentially stable error dynamics for the LTV system is also a state observer for the underlying nonlinear system, and the estimation error converges exponentially fast for all initial conditions.

Proof. The proof is made by comparison of the outputs of both systems. The time evolution of the V2T range in the original nonlinear system depends on the transition for the target position, given by

$$\mathbf{R}(t)\mathbf{p}_T(t) = \mathbf{R}(t_0)\mathbf{p}_T(t_0) + (t - t_0)^E\mathbf{v}_T(t_0) - \mathbf{R}^{[0]}(t, t_0)\mathbf{v}(t_0). \tag{5.13}$$

Replacing this into $r_T(t) = \|\mathbf{p}_T(t)\|$ and squaring the result yields

$$\begin{aligned} r_T^2(t) = {} & \|\mathbf{R}(t_0)\mathbf{p}_T(t_0)\|^2 + (t - t_0)^2\|^E\mathbf{v}_T(t_0)\|^2 \\ & + \|\mathbf{v}^{[0]}(t, t_0)\|^2 + 2(t - t_0)^E\mathbf{v}_T^T(t_0)\mathbf{R}(t_0)\mathbf{p}_T(t_0) \\ & - 2\mathbf{v}^{[0]}(t, t_0)\mathbf{R}(t_0)\mathbf{p}_T(t_0) - 2(t - t_0)\mathbf{v}^{[0]}(t, t_0)^E\mathbf{v}_T(t_0), \end{aligned} \tag{5.14}$$

where the identity $\left(\mathbf{R}^{[0]}(t, t_0)\mathbf{v}(t_0)\right)^T = \mathbf{v}^{[0]}(t, t_0)$ was used. For the B2T range, recall (5.9) and (5.13), consider the expansion of $r_{iT}^2(t)$ given by

$$\begin{aligned} r_{iT}^2(t) = {} & \|\mathbf{R}(t_0)\mathbf{p}_i(t_0)\|^2 - 2\left(\mathbf{R}(t_0)\mathbf{p}_i(t_0)\right)^T\mathbf{R}(t_0)\mathbf{p}_T(t_0) \\ & + \|\mathbf{R}(t_0)\mathbf{p}_T(t_0)\|^2 + (t - t_0)^2\|^E\mathbf{v}_T(t_0)\|^2 \\ & + 2(t - t_0)^E\mathbf{v}_T^T(t_0)\left(\mathbf{R}(t_0)\mathbf{p}_T(t_0) - \mathbf{R}(t_0)\mathbf{p}_i(t_0)\right). \end{aligned} \tag{5.15}$$

For the purpose of comparing the outputs, the same exercise must be performed for the target-related outputs of the augmented system (5.5). However, for simplicity and without loss of generality, the transformed system with state (5.6) will be used. First, consider the transition equation for the target range state

$$\begin{aligned} z_{T_R}(t) = z_{T_R}(t_0) &+ 2(t-t_0)z_{T_{VP}}(t_0) + (t-t_0)^2 z_{T_{VV}}(t_0) \\ &- 2\mathbf{v}^{[0]}(t,t_0)\mathbf{z}_{T_P}(t_0) - 2\mathbf{v}^*(t,t_0)\mathbf{z}_{T_V}(t_0) \qquad , \\ &+ 2\mathbf{v}_r(t,t_0)\mathbf{v}(t_0) \end{aligned}$$

where $\mathbf{v}(t_0)$ replaces $\mathbf{z}_V(t_0)$ according to Theorem 4.2. Given that this state is directly observed as $r_T(t_0)$, this can be compared with (5.14). Setting $t = t_0$, this automatically yields $z_{T_R}(t_0) = \|\mathbf{R}(t_0)\mathbf{p}_T(t_0)\|^2$. Using the transition equation for the beacon-target dot product, it is possible to write

$$\begin{aligned} r_{iT}^2(t) = r_i^2(t_0) &+ z_{T_R}(t_0) - 2z_{T_{Pi}}(t_0) \\ &+ 2(t-t_0)z_{T_{VP}}(t_0) - 2(t-t_0)z_{T_{Vi}}(t_0) \\ &+ (t-t_0)^2 z_{T_{VV}}(t_0). \end{aligned}$$

Comparing this to (5.15) at $t = t_0$ yields $z_{T_{Pi}}(t_0) = \left(\mathbf{R}(t_0)\mathbf{p}_i(t_0)\right)^T \mathbf{R}(t_0)\mathbf{p}_T(t_0)$. Given that $(t-t_0)$ and $(t-t_0)^2$ are linearly independent functions, it is also possible to conclude that $z_{T_{VV}}(t_0) = \|{}^E\mathbf{v}_T(t_0)\|^2$, and that

$$z_{T_{VP}}(t_0) - z_{T_{Vi}}(t_0) = {}^E\mathbf{v}_T^T(t_0)\left(\mathbf{R}(t_0)\mathbf{p}_T(t_0) - \mathbf{R}(t_0)\mathbf{p}_i(t_0)\right).$$

Consider now the first time derivative of the compared outputs, given by

$${}^E\mathbf{v}_T^T(t_0)\mathbf{R}(t_0)\mathbf{p}_T(t_0) - z_{T_{VP}}(t_0) - {}^E\mathbf{v}^T(t)\left(\mathbf{R}(t_0)\mathbf{p}_T(t_0) - \mathbf{z}_{T_P}(t_0)\right) - \boldsymbol{\mathcal{V}}^{[1]}(t,t_0)\left({}^E\mathbf{v}_T(t_0) - \mathbf{z}_{T_V}(t_0)\right) = 0.$$

Setting $t = t_0$ this yields

$$z_{T_{VP}}(t_0) - {}^E\mathbf{v}_T^T(t_0)\mathbf{R}(t_0)\mathbf{p}_T(t_0) - {}^E\mathbf{v}^T(t_0)\left(\mathbf{z}_{T_P}(t_0) - \mathbf{R}(t_0)\mathbf{p}_T(t_0)\right) = 0,$$

which, consequently leads to

$$\boldsymbol{\mathcal{V}}^{[0]}(t,t_0)\left(\mathbf{R}(t_0)\mathbf{p}_T(t_0) - \mathbf{z}_{T_P}(t_0)\right) + \boldsymbol{\mathcal{V}}^{[1]}(t,t_0)\left({}^E\mathbf{v}_T(t_0) - \mathbf{z}_{T_V}(t_0)\right) = 0.$$

It can be seen that if the conditions of the theorem hold, following the same reasoning employed in the proof of Theorem 5.3, this yields

$$\begin{aligned} \mathbf{z}_{T_P}(t_0) &= \mathbf{R}(t_0)\mathbf{p}_T(t_0) \\ \mathbf{z}_{T_V}(t_0) &= {}^E\mathbf{v}_T(t_0) \\ z_{T_{VP}}(t_0) &= {}^E\mathbf{v}_T^T(t_0)\mathbf{R}(t_0)\mathbf{p}_T(t_0). \end{aligned}$$

Noting that $\mathbf{z}_T(t)$ is related to $\mathbf{x}_T(t)$ by a Lyapunov transformation, this concludes the proof that the states of system (5.2) and (5.7) are equivalent, as their initial conditions were shown to be one and the same and the dynamics of both systems are equal.

The second part of the theorem follows naturally from the first part. An observer for (5.7) with globally exponentially stable error dynamics provides estimates that converge exponentially fast to the true state. Therefore, if the state of the LTV system and that of the original nonlinear system, when the invisible beacons are discarded, are one and the same, the estimates of the observer will also tend exponentially fast to the true state of system (5.2) with the same error dynamics. □

5.4 Filter design

The previous results have established the ground to the design of a GES observer, using a linear time-varying Kalman filter, which, to assure the GES nature of the estimation error dynamics, requires the pair $(\mathbf{A}_F(t), \mathbf{C}_F(t))$ to be uniformly completely observable, as shown in Theorem 2.2 and has been used throughout this dissertation. This last theorem addresses the uniform complete observability of the pair $(\mathbf{A}(t), \mathbf{C}(t))$. However, an additional assumption on the linear velocity of the vehicle is required, as well as the definition of $\boldsymbol{\mathcal{V}}(\tau,t) := \left[\boldsymbol{\mathcal{V}}^{[0]}(\tau,t)\ \boldsymbol{\mathcal{V}}^{[1]}(\tau,t)\right]^T$.

Assumption 5.2. *The norm of the linear velocity of the vehicle in the Earth-fixed frame* $\{E\}$ *is always bounded, i.e.,*

$$\underset{t\geq t_0}{\forall}\ \underset{V_M>0}{\exists}:\quad \|^E\mathbf{v}(t)\| \leq V_M.$$

Although imposing bounds on the linear velocity, this assumption is still a mild one, as it is physically impossible to reach arbitrarily large speeds. Moreover, the value of V_M is not required for the filter design.

Theorem 5.5 (Uniform complete observability of the SN-SLAMMOT LTV system)**.** *The pair* $(\mathbf{A}(t), \mathbf{C}(t))$ *is uniformly completely observable if Assumption 5.2 is true and there exist* $\delta > 0$ *and* $\alpha^* > 0$ *such that, for all* $t \geq t_0$*, it is possible to choose a set of instants* $\{t_1, t_2, t_3, t_4, t_5, t_6, t_7\} \in \mathcal{T}_\delta$*, with* $\mathcal{T}_\delta := [t, t+\delta]$*, for which the linear velocity of the vehicle in the Earth-fixed frame respects*

$$\left|\det\left[\boldsymbol{\mathcal{V}}(t_1,t)\ \cdots\ \boldsymbol{\mathcal{V}}(t_6,t)\right]\right| \geq \alpha^* \tag{5.16}$$

and there is a B2T range available, or if there is not,

$$\left|\det\begin{bmatrix}\boldsymbol{\mathcal{V}}(t_1,t) & \cdots & \boldsymbol{\mathcal{V}}(t_7,t)\\ (t_1-t) & \cdots & (t_7-t)\end{bmatrix}\right| \geq \alpha^*. \tag{5.17}$$

Proof. Uniform complete observability implies uniform bounds on the observability Gramian in time intervals of length δ. It is equivalent to having $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} \geq \alpha$ for $t \geq t_0$ and all unit vectors $\mathbf{c}$. Hence, the proof consists of studying $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$ for every possible cases of $\mathbf{c}$.

Consider again system (5.10) and recall that it is related by a Lyapunov transformation to the system composed of (5.5) and (5.7). Thus, the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C}_F(t))$ implies the uniform complete observability of the pair $(\mathbf{A}(t), \mathbf{C})$. Hence, the proof will continue with the transformed system.

A pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C}_F(t))$ is uniformly completely observable if and only if there exist positive constants δ and α such that for all $t \geq t_0$ and for all unit vectors $\mathbf{c}$ the quadratic form $\mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c}$ is greater than or equal to α, i.e., if and only if

$$\underset{\substack{\delta > 0\\ \alpha > 0}}{\exists} \quad \underset{t \geq t_0}{\forall} \quad \underset{\substack{\mathbf{c} \in \mathbb{R}^{n_z}\\ \|\mathbf{c}\| = 1}}{\forall} : \quad \mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} \geq \alpha,$$

meaning that, in contrast with the observability definition used in Theorem 5.3, the Gramian must have uniform bounds at all times. The proof follows by exhaustion, by analysing the quadratic form in the previous expression for all the possible cases of unit vectors $\mathbf{c}$ for all time. The cases $\mathbf{c}_V \geq \alpha_V$, $\mathbf{c}_R \geq \alpha_R$, and $\mathbf{c}_M \geq \alpha_M$ for some $\alpha_V, \alpha_R, \alpha_M \in (0,1)$ have all been treated in the proof of Theorem 4.4, where it was shown that in the conditions of that theorem, naturally enforced by the conditions of this theorem, $\mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} \geq \alpha$ in those cases. Therefore, in the sequel, all these cases are to be negated.

Consider Proposition A.1 from Appendix A. It is possible to see that it applies to $\mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c}$, and therefore it states conditions upon which the quadratic form in analysis is lower bounded if the norm of (5.11) is lower bounded, uniformly in time. Then, it suffices to prove that $\|\mathbf{f}(\tau, t)\| \geq \alpha_1$ with $\tau \in \mathcal{T}_\delta$ and $\alpha_1 > 0$ for every possible $\mathbf{c}$, provided that the conditions of Proposition A.1 are satisfied uniformly in time.

Consider then, that there exists a $|c_{D_{ij}}| \geq \alpha_D$ and that $|c_{R_i}|, |c_{R_i}| < \alpha_R$. Then,

$$\|\mathbf{f}(\tau, t)\|^2 \geq r_{ij}^{-2}(\tau)\left(c_{R_i} + c_{R_j} - 2c_{D_{ij}}\right)^2 \geq \frac{\alpha_D^2}{4R_M^2},$$

where Assumption 5.1 was used to show that $r_{ij} \leq 2R_M$, and α_D was chosen such that $\alpha_D > 4\alpha_R$. Suppose now that $|c_{T_R}| \geq \alpha_{T_R}$, and nothing is imposed to the remaining components of $\mathbf{c}$. Then, if it is evaluated at $\tau = t$, the norm of $\mathbf{f}(\tau, t)$ is $\||\mathbf{f}(t,t)\| \geq r_T^2(t)c_{T_R}^2 \geq \frac{\alpha_{T_R}^2}{R_M^2}$. The next case is defined by $|c_{T_{P_i}}| \geq \alpha_{T_{Pi}}$ for some $0 < \alpha_{T_{Pi}} < 1$ while $|c_{T_R}| < \alpha_{T_R}$ and $|c_R| < \alpha_R$. Then it is possible to write

$$\|\mathbf{f}(t,t)\|^2 \geq r_{iT}^2(t)(c_{R_i} + c_{T_R} - 2c_{T_{Pi}})^2 \geq \frac{\alpha_{T_{Pi}}^2}{4R_M^2}$$

where $\alpha_{T_{Pi}}$ was chosen to respect $\alpha_{T_{Pi}} > 4(\alpha_R + \alpha_{T_R})$. To study the remaining cases, new functions need to be introduced. Let $f_{T_0}(\tau, t) = -r_T^{-1}(\tau) f_T(\tau, t) + c_{T_R}$ and $|c_{T_R}| < \alpha_{T_R}$. It is now possible to write

$$\|\mathbf{f}(t^*, t)\|^2 \geq r_T^{-2}(t^*)\left(f_{T_0}(t^*, t) - c_{T_R}\right)^2 \geq \frac{\alpha_1^2}{2R_M^2},$$

if it is established that $|f_{T_0}(\tau, t)| \geq \alpha_1$ for some $\tau = t^*$ and $\alpha_1 > 4\alpha_{T_R}$. Proposition A.1 can now be used again to show that such a t^* does indeed exist. Consider the first derivative of $f_{T_0}(\tau, t)$, given by

$$\frac{\partial}{\partial \tau} f_{T_0}(\tau, t) = -2^E\mathbf{v}^T(\tau)\mathbf{c}_{T_P} - 2\boldsymbol{\mathcal{V}}^{[1]}(\tau, t_0)\mathbf{c}_{T_V} + 2c_{T_{VP}} + 2(\tau - t_0)c_{T_{VV}} + 2^E\mathbf{v}^T(\tau)\mathbf{R}^{[0]}(\tau, t)\mathbf{c}_V. \quad (5.18)$$

Consider the case $|c_{T_{VP}}| \geq \alpha_{T_{VP}}$ for some $0 < \alpha_{T_{VP}} < 1$ and $\|\mathbf{c}_{T_P}\| < \alpha_{T_P}$. Evaluating the previous expression at $\tau = t$ yields

$$\begin{aligned}\left|\frac{\partial}{\partial\tau} f_{T_0}(t,t)\right|^2 &= 4\left|{}^E\mathbf{v}^T(t)\mathbf{c}_{T_P} - c_{T_{VP}}\right|^2 \\ &\geq 4|c_{T_{VP}}|\left(|c_{T_{VP}}| - 2|{}^E\mathbf{v}^T(t)\mathbf{c}_{T_P}|\right) \\ &\geq 2\alpha_{T_{VP}},\end{aligned}$$

where Assumption 5.2 was used, and $\alpha_{T_{VP}} > 4\alpha_{T_P} V_M$. The next case to study is where either $\|\mathbf{c}_{T_P}\|$ or $\|\mathbf{c}_{T_V}\|$ are lower bounded, i.e., $\|\mathbf{c}_{T_P}\|^2 + \|\mathbf{c}_{T_V}\|^2 \geq \alpha_T$. Adding and subtracting ${}^E\mathbf{v}^T(t)\mathbf{c}_{T_P}$ to (5.18) allows to write

$$\begin{aligned}\left|\frac{\partial}{\partial\tau} f_{T_0}(\tau,t)\right|^2 =& 4\Big[\boldsymbol{\mathcal{V}}^{[0]}(\tau,t)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(\tau,t)\mathbf{c}_{T_V} \\ &- c_{T_{VP}} + {}^E\mathbf{v}^T(t)\mathbf{c}_{T_P} \\ &- {}^E\mathbf{v}^T(\tau)\mathbf{R}^{[0]}(\tau,t)\mathbf{c}_V - (\tau - t_0)c_{T_{VV}}\Big]^2.\end{aligned} \tag{5.19}$$

Noting that in the finite interval $\mathcal{T}_\delta$ it is possible to find an upper bound for $(\tau - t_0)$ and making $|c_{T_{VV}}| < \alpha_{T_{VV}}$, it is possible to define a bound α_2 for the absolute value of the third line above. It can also be noted that, as $\|\mathbf{c}_{T_P}\| < 1$, the second line can also be upper bounded by $\alpha_3 := \sqrt{\alpha_{T_{VP}} + 3V_M^2}$ if $|c_{T_{VP}}|$ is upper bounded by $\alpha_{T_{VP}} \in (0,1)$. Expanding the square in (5.19) and discarding some positive parcels yields

$$\begin{aligned}\left|\frac{\partial}{\partial\tau} f_{T_0}(t^*,t)\right|^2 &\geq 4\alpha_4(\alpha_4 - 2\alpha_2 - 2\alpha_3) - 8\alpha_2\alpha_3 \\ &\geq \frac{3}{4}\alpha_4^2\end{aligned},$$

where $|\boldsymbol{\mathcal{V}}^{[0]}(\tau,t)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(\tau,t)\mathbf{c}_{T_V}|$ was assumed to be lower bounded by some $\alpha_4 \in (0,1)$ at some time $t^* \in \mathcal{T}_\delta$, chosen so that the bounds α_3 and α_2 respect $8(\alpha_2 + \alpha_3) < 3\alpha_4$ and $32\alpha_3\alpha_2 < \alpha_4^2$. From this, it remains to show that there exists in fact a t^* such that $|\boldsymbol{\mathcal{V}}^{[0]}(t^*,t)\mathbf{c}_{T_P} + \boldsymbol{\mathcal{V}}^{[1]}(t^*,t)\mathbf{c}_{T_V}| \geq \alpha_4$ or, defining the vector $\mathbf{c}_T := \left[\mathbf{c}_{T_P}^T \ \mathbf{c}_{T_V}^T\right]^T$,

$$|\boldsymbol{\mathcal{V}}^T(t^*,t)\mathbf{c}_T| \geq \alpha_4.$$

The condition of the theorem (5.16) can be rewritten as

$$|\boldsymbol{\mathcal{V}}^T(t_6,t)\boldsymbol{\eta}| \geq \alpha^*, \tag{5.20}$$

with $\boldsymbol{\eta} \in \mathrm{Null}\{\boldsymbol{\mathcal{V}}^T(t_1,t),\ldots,\boldsymbol{\mathcal{V}}^T(t_5,t)\}$ and $\|\boldsymbol{\eta}\| = \prod_{i=1}^{5}\sigma_i$, where σ_i are the singular values of the matrix composed of the first 5 columns in (5.16). To address the possibility of existence of t^*, consider the worst case. Given that, under the conditions of the theorem, the vector $\boldsymbol{\mathcal{V}}(\tau,t)$ spans $\mathbb{R}^6$ in $\mathcal{T}_\delta$, it is impossible to have $\mathbf{c}_T$ perpendicular to all $\boldsymbol{\mathcal{V}}(t_i,t)$. Therefore, the worst case occurs when it is simultaneously perpendicular to the five $\boldsymbol{\mathcal{V}}(t_i,t)$ whose stacking matrix has the highest singular values. These are assumed to be, without loss of generality, the first five.

Hence, the worst case is given by $\mathbf{c}_T = \boldsymbol{\eta}\|\mathbf{c}_T\|\|\boldsymbol{\eta}\|^{-1}$. To proceed, a bound to $\|\boldsymbol{\eta}\|$ is necessary. Note that the square of the singular values of that matrix are the eigenvalues of

$$\begin{bmatrix} \boldsymbol{\mathcal{V}}^T(t_1)\boldsymbol{\mathcal{V}}(t_1) & \cdots & \boldsymbol{\mathcal{V}}^T(t_1)\boldsymbol{\mathcal{V}}(t_5) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{V}}^T(t_5)\boldsymbol{\mathcal{V}}(t_1) & \cdots & \boldsymbol{\mathcal{V}}^T(t_5)\boldsymbol{\mathcal{V}}(t_5) \end{bmatrix},$$

which respect

$$\lambda_i \leq \sum_{i=1}^{5} \boldsymbol{\mathcal{V}}^T(t_i)\boldsymbol{\mathcal{V}}(t_i) \leq 5\alpha_\lambda.$$

with $\alpha_\lambda \geq 2V_M^2 + 4\delta$. Hence, $\|\boldsymbol{\eta}\| \leq \sqrt{\alpha_\lambda^5}$. Using (5.20), one may write

$$\left|\boldsymbol{\mathcal{V}}^T(t_6, t)\mathbf{c}_T\right| \geq \frac{\alpha^*\alpha_T}{\sqrt{\alpha_\lambda^5}},$$

which is the worst case for any $t^* = t_1$, t_2, t_3, t_4, t_5, or t_6 and any lower norm bounded $\mathbf{c}_T$. In the case where there is no B2T range available, the vector $\boldsymbol{\mathcal{V}}(\tau, t)$ must be extended to $\mathbb{R}^7$ and include $(\tau - t)$, the vector $\mathbf{c}_T$ is extended to include $c_{T_{VV}}$, and condition (5.17) applies in a similar way. Otherwise, a new function must be defined. Let $f_{T_i}(\tau, t) = -r_{iT}^{-1}(\tau)f_{TM_i}(\tau, t) + c_{T_R} + c_{R_i} - 2c_{T_{Pi}}$. It is now possible to write

$$\begin{aligned} \|\mathbf{f}(t^*, t)\|^2 &\geq r_{iT}^{-2}(t^*)\left(f_{T_i}(t^*, t) - c_{T_R} - c_{R_i} + 2c_{T_{Pi}}\right)^2 \\ &\geq \frac{\alpha_5^2}{8R_M^2}, \end{aligned}$$

if it is established that $|f_{T_i}(\tau, t)| \geq \alpha_5$ for some $\tau = t^*$ and $\alpha_5 > 4(\alpha_{T_R} + \alpha_{T_{Pi}})$. Once again, Proposition A.1 can be used. Consider the first derivative of $f_{T_i}(\tau, t)$, given by

$$\frac{\partial}{\partial\tau}f_{T_i}(\tau, t) = 2c_{T_{VP}} + 2(\tau - t_0)c_{T_{VV}} - 2c_{T_{Vi}} + 2^E\mathbf{v}^T(\tau)\mathbf{c}_{M_i} - 2^E\mathbf{v}^T(\tau)\mathbf{R}^{[0]}(\tau, t)\mathbf{c}_V. \tag{5.21}$$

When evaluated at $\tau = t$, for the case $|c_{T_{Vi}}| \geq \alpha_{T_{Vi}}$ and $|c_{T_{VP}}| < \alpha_{T_{VP}}$ this yields,

$$\begin{aligned} \left|\frac{\partial}{\partial\tau}f_{T_i}(t, t)\right| &= 4(c_{T_{Vi}} - c_{T_{VP}})^2 \\ &\geq 2\alpha_{T_{Vi}}^2 \end{aligned},$$

where $\alpha_{T_{Vi}}$ was chosen so that $\alpha_{T_{Vi}} > 4\alpha_{T_{VP}}$. The final case to address is $|c_{VV}| \geq \alpha_{T_{VV}}$ for some $\alpha_{T_{VV}} \in (0, 1)$. Noting that all other components of $\mathbf{c}$ are upper bounded, all the parcels in (5.21) are also bounded except for $(\tau - t_0)c_{T_{VV}}$. Therefore, denoting α_6 as the bound for the absolute value of the sum of all the other parcels, one can write

$$\left|\frac{\partial}{\partial\tau}f_{T_i}(t + \delta, t)\right| \geq 2\delta^2|\alpha_{T_VV}^2,$$

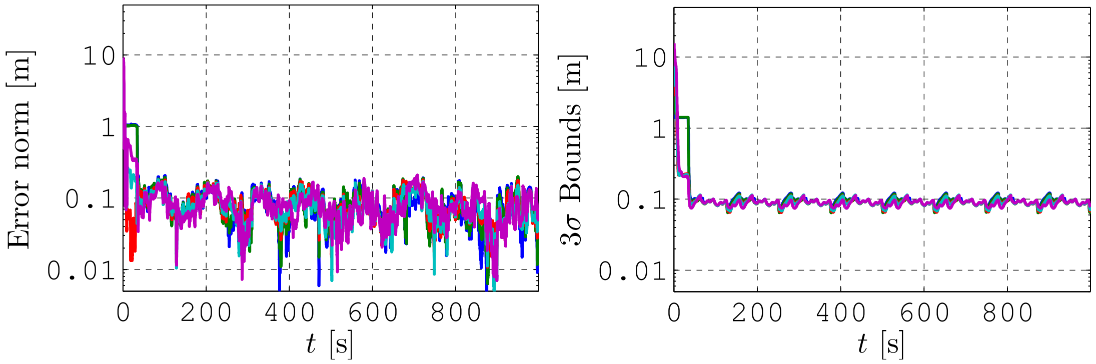
where $\alpha_{T_{VV}} \geq 4\alpha_6$. By Proposition A.1, $|f_{T_i}(\tau, t)| \geq \alpha_3$, which in turn implies $\|\mathbf{f}(\tau, t)\| \geq \alpha$.

This concludes the enumeration of all the possible values for $\mathbf{c}$, and the proof that, for every possibility, $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c}$ is lower bounded, thus proving the sufficiency of the conditions of the theorem for the uniform complete observability of the pair $(\boldsymbol{\mathcal{A}}(t), \mathbf{C}_F(t))$. Being related to the LTV system (5.7) through a Lyapunov transformation, this result implies that $(\mathbf{A}(t), \mathbf{C}(t))$ is also uniformly completely observable, and thus the proof of the sufficiency part of the theorem is concluded. ☐

The theoretical results of this chapter were established in a deterministic setting, and thus the presence of measurement noise raises the need for a filtering solution. Hence, a Kalman filter follows naturally for the augmented nonlinear systems (5.5) and (5.7), in a discrete time framework, obtained using the forward Euler discretization as in previous chapters. The algorithm is therefore the standard discrete time LTV Kalman filter [Gel74], as described in Section 4.4.

5.5 Simulation results

This section details the realistic simulations performed to validate the algorithm proposed in this chapter and to assess its performance. The results of a typical run in the simulated environment detailed in Appendix C are presented and discussed. The environment consists of a 16 m by 16 m by 3 m corridor.



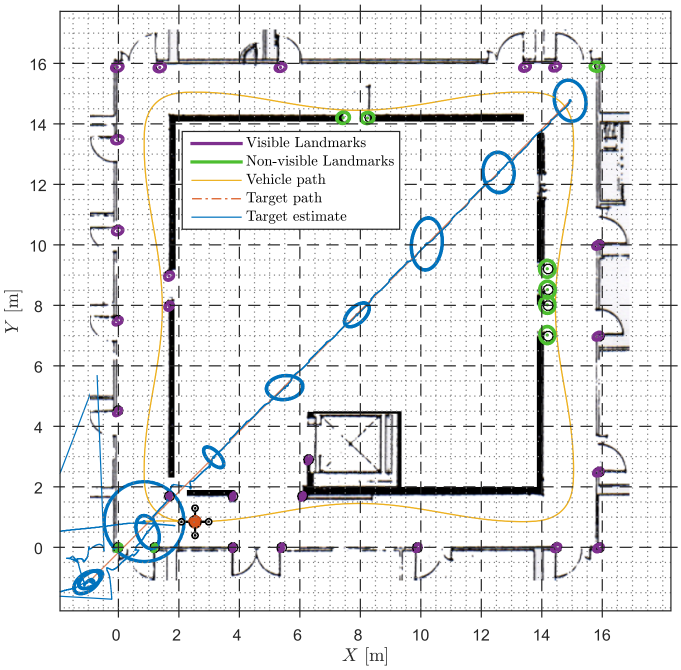
(a) Norm of the error.

(b) Uncertainty.

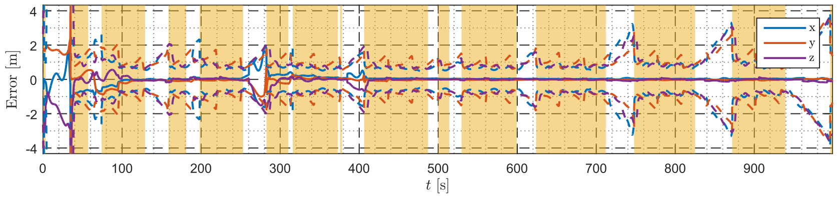
Figure 5.3: Evolution of the estimation of 5 beacons in time.

The chosen environment was modified to include a large open space in the section inside the corridor. 36 beacons were put in notable places such as corners and doors, with random heights. All the measurements are assumed to be perturbed by zero-mean Gaussian white noise, with standard deviations of $\sigma_\omega = 0.05$ °/s for the angular rates, $\sigma_v = 0.03$ m/s for the linear velocity, and $\sigma_r = 0.03$ m for the ranges. The trajectory of the vehicle is depicted in yellow in Figure 5.4(a), and was designed to meet the observability requirements.

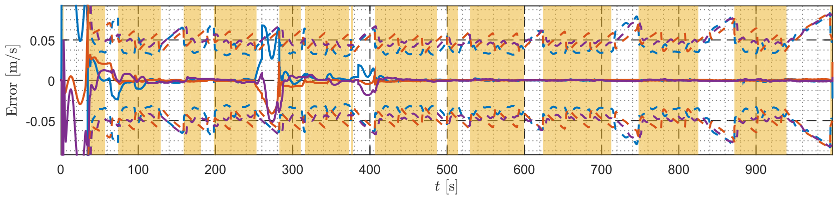
Firstly, the results of a simulation without a target to track are presented in Figure 5.3. These aim at demonstrating the good performance of the RO-SLAM algorithm in a sensor



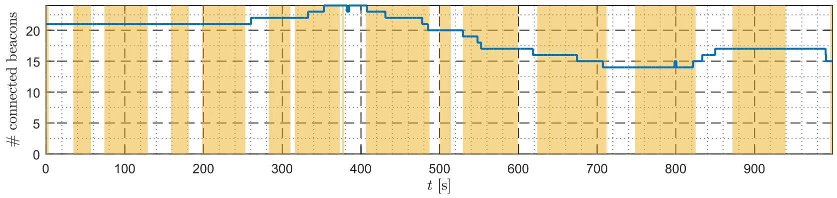
(a) 2-D map



(b) Position error.



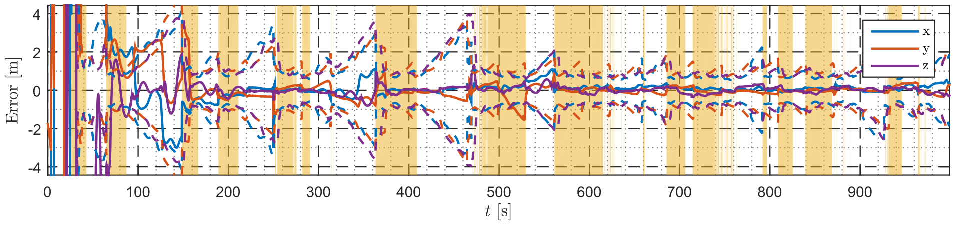
(c) Velocity error.



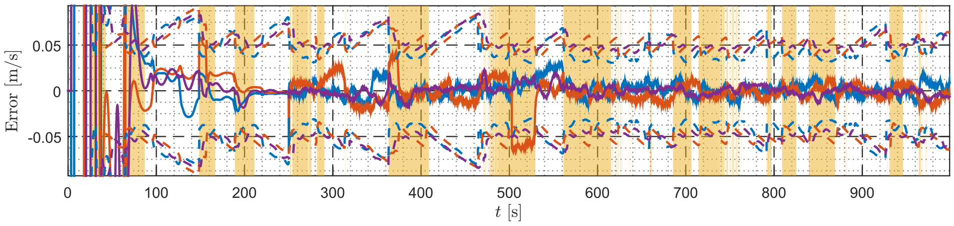
(d) Beacons connected to the target .

Figure 5.4: On top, a picture of the estimated map rotated and translated using the true transformation, with 3σ ellipsoids, and the path of the target, both real (blue) and estimated (red). Non-visible beacons in green; Visible beacons in purple; Vehicle path in yellow. Below, the estimation error and 3σ uncertainty bounds of the target position and velocity when it follows a straight line. Solid lines indicate the estimation error for each coordinate and dashed lines the uncertainty bounds. The vertical yellow bars indicate the time intervals where the vehicle is directly observing the target.

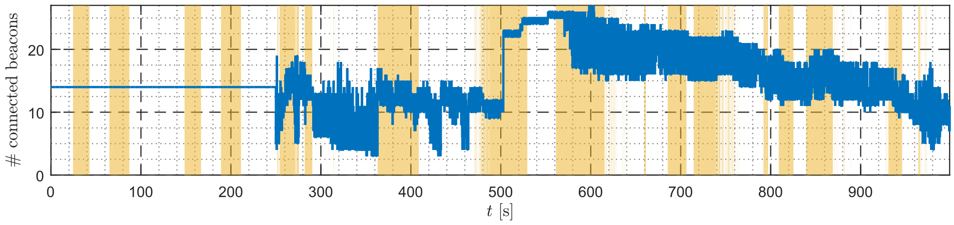
network, and are complemented in that purpose by the simulation and experimental results detailed in Sections 4.5 and 4.6. Comparison with these results shows that the uncertainty does not grow as much when the beacons are not observed, although its lowest value when the estimation converges is very near the achieved with the simple RO-SLAM algorithm (10 cm). The next figures depict the results of two different simulations of target tracking. In the first, the target starts in the bottom right and is stopped for the first 250 seconds. Then it accelerates until reaching constant velocity while describing a straight line as depicted in Figure 5.4(a). This trajectory was chosen because it is the only one that respects the constant velocity model. The results of this simulation are presented in Figure 5.4. Note that the algorithm takes around 100 seconds to converge, but when it does, the error is kept very low until the target starts moving



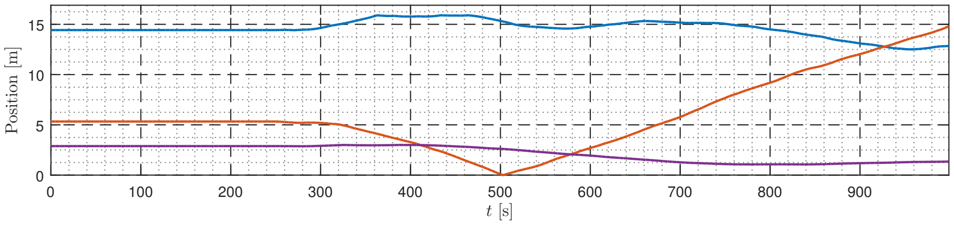
(a) Position error.



(b) Velocity error.



(c) Beacons connected to the target .



(d) Target trajectory.

Figure 5.5: The estimation error and 3σ uncertainty bounds of the target position and velocity when it follows a random trajectory. The vertical yellow bars indicate the time intervals where the vehicle is directly observing the target.

and the error grows, just to converge again when the algorithm tends to the new velocity of the target. The 3σ uncertainty is consistent with the errors, except for a few moments when the target starts moving. It can be seen that even when the vehicle is directly observing the target (for example, in interval between 400 and 500 seconds in Figure 5.4) the uncertainty may grow, which means that in those moments the observability conditions of Theorem 5.5 are not consistently met.

To illustrate a case in which the constant velocity model is not correct, in the second simulation, the target is placed randomly in the environment and, after 250 seconds stopped, starts moving in a random walk (${}^{E}\mathbf{v}_T(t_{k+1}) = {}^{E}\mathbf{v}_T(t_k) + \mathbf{u}(t_k)$ where $\mathbf{u}(t_k)$ is zero-mean Gaussian white noise), while restricted to the confines of the environment. The results of a typical run are shown in Figure 5.5. Once again, the theoretical convergence properties are confirmed by the results, as, albeit being initialized at the origin of the body-fixed frame, the estimation converges fairly fast. Even when the vehicle starts moving, the filter accompanies the motion and the error does not increase significantly except when the velocity is inverted at around 500 seconds into the run. The actual trajectory followed can be seen in Figure 5.5(d).

5.6 Conclusions

This chapter presented a novel sensor-based filter for range-only simultaneous localization and mapping in sensor networks with moving target tracking. The filter is shown to have globally exponentially stable error dynamics, through state augmentation of two nonlinear systems, which, along with the disposal of the non-visible beacons, enabled regarding the resulting system as linear time-varying. The work focused on the observability analysis of the resulting system, providing theoretical observability guarantees, and equivalence between the systems used in each step of the analysis. The theoretical results include the derivation of sufficient conditions for observability, stability and convergence of the algorithm, establishing a constructive basis for trajectory design. These results were followed by the design of a Kalman filter with globally exponentially stable error dynamics. Simulations allowed the validation of the results, even when the mobile node trajectory does not respect the constant Earth-fixed velocity model used. As in its parent framework, RO-SLAM, these simulations also show the relevance of the trajectory in the performance of the algorithm.

6

SIMULTANEOUS LOCALIZATION AND MAPPING WITH MONOCULAR VISION

THIS chapter proposes a novel filter for sensor-based bearing-only simultaneous localization and mapping in three dimensions with globally exponentially stable (GES) error dynamics. Building on the generic system described in Chapter 2, a nonlinear system is designed, its output transformed, and its dynamics augmented so that the proposed formulation can be considered as linear time-varying for the purpose of observability analysis. This allows the establishment of observability results related to the original nonlinear system that naturally lead to the design of a Kalman filter with GES error dynamics. The performance of the proposed algorithm is assessed resorting to real experiments based on the *Rawseeds* dataset, where a wheeled ground robot equipped with a monocular camera is used, and is accompanied by results of simulations to explore specific details of the algorithm.

6.1 Introduction

In opposition to range-and-bearing SLAM, a bearing-only formulation using monocular vision is less demanding both in terms of equipment and computational power. A single camera has usually a low weight, requires low power, is a less expensive sensor, and can also serve other purposes aside from mapping and localization. For example, an onboard camera can be used for teleoperation, surveillance, within other purposes. On the other hand, bearing-only simultaneous localization and mapping (BO-SLAM) requires a measure of the linear velocity of the vehicle, which demands an additional sensor that is not needed for range-and-bearing simultaneous localization and mapping (RB-SLAM). All in all, SLAM with monocular vision has many possible applications where a camera is useful, from aerial vehicles mapping 3-D environments, finding survivors in search-and-rescue operations, to household robots[1].

This work proposes a solution to the initialization problem of bearing-only SLAM explained in the introductory chapter by introducing an algorithm with exponentially fast global convergence, which allows for undelayed initialization at any arbitrary depth, provided that the vehicle respects certain motion conditions. With its three-dimensional (3-D) sensor-based approach, the pose of the vehicle is eliminated from the filter state and the inclusion of odometry-like measurements and relative bearings is straightforward. This aspect, coupled with a state augmentation and output transformation, leads to the design of an LTV system whose observability is analysed in this chapter, resulting in constructive conditions with clear physical insight that are important for motion planning. The underlying idea of this chapter is influenced by the source-localization algorithm presented in [BSO13b], as the proposed filter results from similar state and output transformations. This work is validated experimentally using a monocular camera and data from a widely available dataset, the *Rawseeds* dataset [BBF⁺06, CFG⁺09].

---

[1] An interesting example is the *Dyson 360 Eye*™ that uses an omnidirectional camera, see `http://www.dyson.com/vacuum-cleaners/robot/dyson-360-eye.aspx`

Chapter Structure

The chapter is organized as follows. A short description of the problem, with the definition of the system dynamics, is presented in Section 6.2. Section 6.3 details the proposed solution, including the observability analysis and filter design. The implementation of the algorithm is described in Section 6.4 including landmark detection, data association, and loop closing procedures. Section 6.5 presents results of typical simulations for different noise levels. Experimental results based on the *Rawseeds* dataset are detailed and discussed in Section 6.6, whereas in Section 6.7 the concluding remarks are provided.

6.2 The bearing-only SLAM problem

Consider a vehicle operating in a static environment, capable of measuring the relative azimuth and elevation of landmarks installed in unknown locations, as well as its linear and angular velocities in its own reference frame. Possible examples are ground robots equipped with wheel encoders, or aerial vehicles equipped with optical flow sensors and an altimeter. Both of these assisted by an IMU to measure the angular velocities and a camera or an acoustic/electromagnetic receiver to detect landmarks. The landmarks can be artificial or natural, i.e., previously installed or extracted from the scenery. This situation falls under the scope of BO-SLAM, which is the problem of navigating a vehicle in an unknown environment, building a map of metric landmarks by measuring bearings and using this map to deduce its location, without the need for *a priori* information about landmark Earth-fixed location.

6.2.1 The sensor-based approach

Recall the generic system presented in (2.1). From the problem definition, it is known that both the linear and the angular velocities are measured, as well as relative bearings to the landmarks (see Figure 6.1). This last quantity is described by the unit vector $\mathbf{b}_i(t)$ that defines

Figure 6.1: In bearing-only SLAM, the vehicle is equipped with sensors that are capable of measuring the relative direction to a landmark. For example, a single camera.

the line between the position of the vehicle and landmark i, and is given by

$$\mathbf{b}_i(t) = \frac{\mathbf{p}_i(t)}{\|\mathbf{p}_i(t)\|} \in \mathcal{S}(2),$$

with $i \in \mathcal{M}_O$. As the information this measurement carries is limited, not only several measurements from the same landmark are needed to unambiguously determine its position but also some measure of scale is required. This is provided by the linear velocity measurements and it is the reason why they must be available.

This section culminates naturally with the nonlinear system that puts all this information together. The positions of the landmarks in the body-fixed frame are its states, the linear and angular velocities are the inputs, and the measured quantities are its outputs ($\mathbf{b}_i(t)$). The resulting system model is

$$\begin{cases} \dot{\mathbf{p}}_i(t) = -\mathbf{S}(\boldsymbol{\omega}(t))\,\mathbf{p}_i(t) - \mathbf{v}(t), & i \in \mathcal{M} \\ \mathbf{b}_j(t) = \dfrac{\mathbf{p}_j(t)}{\|\mathbf{p}_j(t)\|}, & j \in \mathcal{M}_O \end{cases}. \tag{6.1}$$

The output $\mathbf{b}_j(t)$ can be stacked in a column vector to obtain $\mathbf{b}(t) = \left[\mathbf{b}_1^T(t) \cdots \mathbf{b}_{N_O}^T(t)\right]^T$. This stacked vector will be used to simplify the presentation of later quantities.

6.2.2 Problem statement

The problem addressed in this chapter is that of designing a navigation system for a vehicle operating in the environment here described, by means of a filter for the dynamics in (6.1), assuming noisy measurements. The algorithm consists of a BO-SLAM filter in the space of sensors, and, therefore, the pose of the vehicle is deterministic as, by construction, it corresponds to the position and attitude of the body-fixed frame expressed in that same frame.

6.3 Proposed solution: GES BO-SLAM

The system presented in the last section is still nonlinear, even though the sensor-based approach allowed one to avoid including the pose of the vehicle in the dynamics. In some problems, where the nonlinearity occurs in the output equation, a state augmentation can help to remove the nonlinearity, as was done successfully in the two previous chapters, where the idea was applied to range-only simultaneous localization and mapping (RO-SLAM). In this chapter, the proposed solution relies on an output transformation that leads to a state augmentation, inspired by the results presented in [BSO13b].

6.3.1 State augmentation and output transformation

The objective of this subsection is to obtain a linear-like system that mimics the dynamics of the original nonlinear system while avoiding the nonlinearity on the bearing output. For that purpose, consider then the following physically sensible mild assumption.

Assumption 6.1. *The position of the vehicle cannot coincide with a landmark, i.e., a visible bearing vector is always defined.*

With this in mind, the manipulation of the output of (6.1) yields

$$\mathbf{p}_i(t) - \mathbf{b}_i(t)\|\mathbf{p}_i(t)\| = \mathbf{0}, \quad i \in \mathcal{M}_O. \tag{6.2}$$

If the norm of the i-th landmark is added as a state, this expression becomes in fact linear. That is the idea behind the augmented state

$$\mathbf{x}(t) := \left[\mathbf{x}_M^T(t)\ \mathbf{x}_R^T(t)\right]^T,$$

where $\mathbf{x}_M(t) \in \mathbb{R}^{n_M}$ is the stacking of all landmarks, both visible and non-visible, and $\mathbf{x}_R(t) \in \mathbb{R}^{n_R}$ agglomerates all the norms of the landmarks, i.e., the distance from each landmark to the vehicle. These correspondences are summarized by the state constraints

$$\begin{cases} \mathbf{x}_{M_i}(t) := \mathbf{p}_i(t) \\ x_{R_i}(t) := \|\mathbf{x}_{M_i}(t)\| \end{cases}, \tag{6.3}$$

for all $i \in \mathcal{M}$, where $\mathbf{x}_{M_i}(t) \in \mathbb{R}^3$ and $x_{R_i}(t) \in \mathbb{R}$ are i-th components of the landmark and range state vectors, respectively. Note that both the landmark and range states are composed by visible and non-visible parts, denoted by subscripts O and U respectively.

Consider the derivative of the range state, given by

$$\dot{x}_{R_i}(t) = -\frac{\mathbf{x}_{M_i}^T(t)}{x_{R_i}(t)}\mathbf{v}(t)$$

which is needed to write the full state dynamics. When a landmark is observed and its bearing is available, the quotient $\frac{\mathbf{x}_{M_i}(t)}{x_{R_i}(t)}$ can be replaced by the bearing $\mathbf{b}_i(t)$ for all $i \in \mathcal{M}_O$. Knowing this, the resulting system reads

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_F(t)\,\mathbf{x}(t) + \mathbf{B}_F(t, \mathbf{b}(t), \mathbf{x}_{M_U}(t), \mathbf{x}_{R_U}(t))\mathbf{v}(t) \\ \mathbf{y}(t) = \mathbf{C}_F(t, \mathbf{b}(t))\,\mathbf{x}(t) \end{cases}, \tag{6.4}$$

where the dynamics matrix is

$$\mathbf{A}_F(t) = \begin{bmatrix} \mathbf{A}_M(t) & \mathbf{0}_{n_M \times n_R} \\ \mathbf{0}_{n_R \times n_M} & \mathbf{0}_{n_R \times n_R} \end{bmatrix}$$

with component $\mathbf{A}_M(t) = -\operatorname{diag}(\mathbf{S}(\boldsymbol{\omega}(t)), \ldots, \mathbf{S}(\boldsymbol{\omega}(t)))$. The input matrix is given by

$$\mathbf{B}_F(t, \mathbf{b}(t), \mathbf{x}_{M_U}(t), \mathbf{x}_{R_U}(t)) = \begin{bmatrix} \mathbf{B}_M \\ \mathbf{B}_R(t, \mathbf{b}(t), \mathbf{x}_{M_U}(t), \mathbf{x}_{R_U}(t)) \end{bmatrix}$$

where $\mathbf{B}_M = -\begin{bmatrix} \mathbf{I} & \cdots & \mathbf{I} \end{bmatrix}^T \in \mathbb{R}^{n_M \times 3}$ and

$$\mathbf{B}_R(t, \mathbf{b}(t), \mathbf{x}_{M_U}(t), \mathbf{x}_{R_U}(t)) = -\begin{bmatrix} \mathbf{b}_1(t) & \cdots & \mathbf{b}_{N_O}(t) & \frac{\mathbf{x}_{M_{N_O+1}}(t)}{x_{R_{N_O+1}}(t)} & \cdots & \frac{\mathbf{x}_{M_N}(t)}{x_{R_N}(t)} \end{bmatrix}^T.$$

The output matrix is

$$\mathbf{C}_F(t, \mathbf{b}(t)) = \begin{bmatrix} \mathbf{I}_{n_O} & \mathbf{0}_{n_O \times n_U} & \mathbf{C}_b(t, \mathbf{b}(t)) & \mathbf{0}_{n_O \times N_U} \end{bmatrix},$$

with $\mathbf{C}_b(t, \mathbf{b}(t)) = -\operatorname{diag}\left(\mathbf{b}_1(t), \ldots, \mathbf{b}_{N_O}(t)\right)$. Finally, the output is $\mathbf{y}(t) = \begin{bmatrix} \mathbf{0}_{n_O \times 1} \end{bmatrix}$.

Remark 6.1. *The measured quantities $\mathbf{b}_i(t)$ are signals external to this augmented system that tie it to the original nonlinear one. They can be seen as inputs, available for observer design purposes, and they provide the information necessary in the output equation, which is identically zero. The observability analysis in the next subsection addresses these issues.*

Even though the output nonlinearity as first brought up disappeared with the state augmentation and output transformation proposed in this section, the process introduced two new non-linearities. The first is on the input matrix, as it depends both on a measured quantity, the bearing, and on the state, when the measurement is not available. The second is on the output matrix that also depends on a measured quantity. However, the presence of the measurement in the input and output matrices is not really a problem, as, for observability purposes, a system whose dynamics matrix depends on the output can be seen as a linear time-varying system. The presence of the state in the input matrix only affects the non-visible landmarks ($\mathbf{x}_{M_U}(t)$ and $\mathbf{x}_{R_U}(t)$). These are not observable, and therefore will be propagated in open loop.

Another important aspect that must be stressed is the fact that there is nothing in the augmented system (6.4) that imposes the constraints (6.3), particularly the nonlinear relation $x_{R_i}(t) = \|\mathbf{x}_{M_i}(t)\|$, and as such the relation between the nonlinear and the augmented systems must be carefully analysed.

The current state-of-the-art approach in BO-SLAM filters, inverse depth-based algorithms [CDM08], also employs a state augmentation to try to cope with the nonlinearity in the measurement model. In that case, each feature is represented by a 6-D vector containing the camera coordinates from the first sighting, the inverse depth and the image coordinates. The approach presented in this chapter is also not minimal, but has lower dimensionality and ensures that the system is truly linear time-varying which is important for observability, stability, and convergence purposes as explained in the subsequent subsections.

6.3.2 Observability analysis

The subject of this subsection is the observability analysis of the nonlinear and augmented systems presented previously. The augmented system (6.4) contains non-visible landmarks and associated ranges that are clearly not observable as the corresponding bearing is not available. Hence, in the observability analysis, the non-visible landmarks and the associated ranges are discarded, following the successful approach first used by in [GBSO13]. Furthermore, since each landmark-range-bearing group is independent of the others, it is possible to simplify the analysis greatly by assuming that only one landmark is visible, i.e., $\mathcal{M}_O := \{1\}$.

The new reduced system is given by

$$\begin{cases} \dot{\mathbf{z}}(t) = \mathbf{A}(t)\mathbf{z}(t) + \mathbf{B}(t)\mathbf{v}(t) \\ \mathbf{y}(t) = \mathbf{C}(t)\mathbf{z}(t) \end{cases} \tag{6.5}$$

where the dynamics matrix is

$$\mathbf{A}(t) = \begin{bmatrix} -\mathbf{S}(\omega(t)) & \mathbf{0}_{3\times1} \\ \mathbf{0}_{1\times3} & 0 \end{bmatrix},$$

the input matrix is

$$\mathbf{B}(t) = \begin{bmatrix} -\mathbf{I}_3 \\ -\mathbf{b}_1^T(t) \end{bmatrix},$$

and the output matrix is

$$\mathbf{C}(t) = \begin{bmatrix} \mathbf{I}_3 & -\mathbf{b}_1(t) \end{bmatrix}.$$

The dependence of the input matrix on the non-visible landmarks and ranges has disappeared with the reduction of the state of this system. However, both the input and the output matrices still depend on the visible bearing. Given that this quantity is a known function of time, the reduced system can be considered as linear time-varying for observability analysis and observer design. Lemma 2.1, which addresses this issue, will be exploited throughout this subsection.

The forthcoming analysis requires the definition of ${}^E\mathbf{b}_1(t) = \mathbf{R}(t)\mathbf{b}_1(t)$ as the Earth-fixed or absolute bearing. This theorem addresses the observability analysis of system (6.5) regarded as LTV.

Theorem 6.1 (Observability of the BO-SLAM LTV system)**.** *Take system* (6.5)*, regarded as LTV, and let $\mathcal{T} := [t_0, t_f]$. The system is observable in $\mathcal{T}$ if and only if the absolute bearing associated with the visible landmark is not constant in $\mathcal{T}$, i.e., there exists a $t_1 \in \mathcal{T}$ such that ${}^E\dot{\mathbf{b}}_1(t_1) \neq 0$.*

Proof. Consider the Lyapunov transformation (see Definition A.1 in Appendix A for details)

$$\chi(t) = \mathbf{T}(t)\mathbf{z}(t)$$

where $\mathbf{T}(t) = \operatorname{diag}(\mathbf{R}(t), 1)$ preserves the observability properties of the original system. It is a matter of computation to obtain the transformed system, given by

$$\begin{cases} \dot{\chi}(t) = \boldsymbol{\mathcal{B}}(t)\mathbf{v}(t) \\ \mathbf{y}(t) = \boldsymbol{\mathcal{C}}(t)\chi(t) \end{cases} \tag{6.6}$$

where the input matrix is

$$\boldsymbol{\mathcal{B}}(t) = \begin{bmatrix} -\mathbf{R}(t) \\ -\mathbf{b}_1^T(t) \end{bmatrix},$$

and the output matrix is

$$\boldsymbol{\mathcal{C}}(t) = \begin{bmatrix} \mathbf{R}^T(t) & -\mathbf{b}_1(t) \end{bmatrix}.$$

The proof, made by contraposition, will follow with the transformed system for simplicity of analysis. This system is assumed not observable, which, using Lemma 2.1 implies that the observability Gramian is singular. Then it is shown that the conditions of the theorem cannot hold.

Given the nature of (6.6), it is clear that its transition matrix is $\boldsymbol{\phi}(t,t_0) = \mathbf{I}$. With the transition matrix at hand, the observability Gramian given by (2.3) can now be analysed.

Suppose that the LTV system (6.5) is not observable, meaning that the transformed system is also not observable, which implies that the observability Gramian is singular. Then, there exists a unit vector $\mathbf{c} := \begin{bmatrix} \mathbf{c}_m^T & c_r \end{bmatrix} \in \mathbb{R}^{n_x}$ such that

$$\mathbf{c}^T \boldsymbol{\mathcal{W}}(t_0,t_f)\mathbf{c} = \int_{t_0}^{t_f} \|\mathbf{f}(\tau,t_0)\|^2 \mathrm{d}\tau \tag{6.7}$$

is zero. In that case, the function $\mathbf{f}(\tau,t_0)$, given by

$$\mathbf{f}(\tau,t_0) = \mathbf{R}^T(\tau)\mathbf{c}_m - \mathbf{b}_1(\tau)c_r, \tag{6.8}$$

must be zero too for all $\tau \in \mathcal{T}$. The following condition arises after left multiplication of (6.8) by $\mathbf{R}(\tau)$

$${}^{E}\mathbf{b}_1(\tau)c_r = \mathbf{c}_m, \quad \forall \tau \in \mathcal{T}.$$

From this it follows that ${}^{E}\mathbf{b}_1(\tau)$ must be constant for all $\tau \in \mathcal{T}$, which contradicts the conditions of the theorem. Hence, if the system is not observable, the conditions of the theorem cannot hold. Therefore, by contraposition, if the conditions of the theorem do hold, the transformed system (6.6) is observable, and, as it is related to the LTV system (6.5) through a Lyapunov transformation, the latter is also observable, thus concluding the first part of the proof.

For the necessity part, assume that the conditions of the theorem do not hold, i.e., ${}^{E}\mathbf{b}_1(t) = {}^{E}\mathbf{b}_1(t_0)$ for all $t \geq t_0$. In that case, if it is possible to find a unit vector $\mathbf{c}$ such that (6.7) is zero, then the system is not observable, thus proving the necessity of the conditions. Under this assumption, (6.8) becomes

$$\mathbf{f}(\tau,t_0) = \mathbf{R}^T(\tau)(\mathbf{c}_m - {}^{E}\mathbf{b}_1(t_0)c_r).$$

Choosing $\mathbf{c}_m = {}^{E}\mathbf{b}_1(t_0)c_r$ and $c_r = \frac{\sqrt{2}}{2}$, it is clear that $\mathbf{f}(\tau,t_0) = \mathbf{0}$ for all $\tau \in \mathcal{T}$. Then, it is possible to conclude that (6.7) is zero, which means that the system is not observable. Hence, by contraposition, if the system is observable, the conditions of the theorem must hold, thus concluding the proof of the necessity part of the theorem. □

Remark 6.2. *The sufficient and necessary condition introduced by this theorem is in fact a requirement on the motion of the vehicle. For the system to be observable, i.e., in order to be possible to obtain the initial condition of a landmark, the trajectory of the vehicle must not be restricted to the line described by the absolute bearing as exemplified in Figure 6.2, or, in other words, landmark parallax must be observed.*

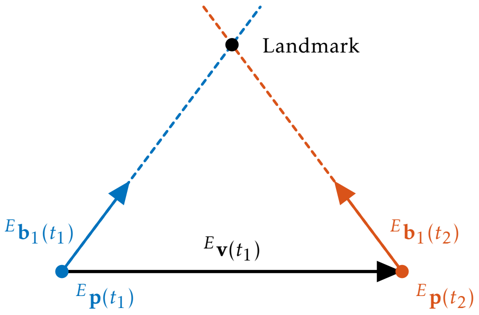


Figure 6.2: Triangulation for positioning a landmark in 2-D.

This theorem established sufficient and necessary conditions for the observability of the system (6.5) that is a reduced version of the augmented nonlinear system (6.4). Given that the discarded states are not observable and do not influence the others, the two systems are equivalent in what concerns observability, when discarding the non-visible landmarks. Hence, this result also applies to the augmented system with a note of caution: when dealing with multiple landmarks, the observability condition requires that, in the considered time interval, all landmarks must be visible in instants where the vehicle did not travel in their direction. This does not mean that all landmarks should be visible at the same time, but that in the time intervals in which each landmark is visible the conditions of Theorem 6.1 must hold. As to the original nonlinear system, this observability result cannot be extrapolated without special attention. Recall that although the augmented system (6.4) mimics the dynamics of the nonlinear one, there is nothing imposing the state relations (6.3). The sequel addresses this aspect, following the approach in the previous chapters.

Theorem 6.2 (Equivalence of the BO-SLAM LTV and nonlinear systems)**.** *If the conditions of Theorem 6.1 and Assumption 6.1 hold, then:*

(i) the state of the original nonlinear system (6.4) and that of the LTV system (6.5) are the same and uniquely determined, and the constraints (6.3) are imposed by the dynamics;

(ii) an observer for the LTV system with globally exponentially stable error dynamics is also a state observer for the underlying nonlinear system with error dynamics that converge exponentially.

Proof. Once more, the proof will focus on the transformed system (6.6) for simplicity of analysis. The output of this system is $\mathbf{y}(t) = \mathbf{0}$.

Consider the time evolution of the transformed LTV system

$$\boldsymbol{\chi}(t) = \boldsymbol{\phi}(t,t_0)\boldsymbol{\chi}(t_0) + \int_{t_0}^{t} \boldsymbol{\phi}(t,\tau)\boldsymbol{\mathcal{B}}(\tau)\mathbf{v}(\tau)\mathrm{d}\tau.$$

It is a matter of computation to obtain

$$\boldsymbol{\chi}_{M_1}(t) = \boldsymbol{\chi}_{M_1}(t_0) - \boldsymbol{\Delta}\mathbf{p}(t,t_0) \tag{6.9}$$

where $\boldsymbol{\Delta}\mathbf{p}(t,t_0) = \int_{t_0}^{t} \mathbf{R}(\tau)\mathbf{v}(\tau)\mathrm{d}\tau$, and

$$\chi_{R_1}(t) = \chi_{R_1}(t_0) - \int_{t_0}^{t} \mathbf{b}_1^T(\tau)\mathbf{v}(\tau)\mathrm{d}\tau. \tag{6.10}$$

Noting that $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{p}_1(t)\| = -\mathbf{b}_1^T(t)\mathbf{v}(t)$, it is also possible to write

$$\mathbf{R}(t)\mathbf{p}_1(t) = \mathbf{R}(t_0)\mathbf{p}_1(t_0) - \boldsymbol{\Delta}\mathbf{p}(t,t_0) \tag{6.11}$$

and

$$\|\mathbf{p}_1(t)\| = \|\mathbf{p}_1(t_0)\| - \int_{t_0}^{t} \mathbf{b}_1^T(\tau)\mathbf{v}(\tau)\mathrm{d}\tau. \tag{6.12}$$

From the output of the transformed system (6.6) and using (6.9) and (6.10), it follows that

$$\mathbf{R}^T(t)\boldsymbol{\chi}_{M_1}(t_0) - \mathbf{R}^T(t)\boldsymbol{\Delta}\mathbf{p}(t,t_0) - \mathbf{b}_1(t)\chi_{R_1}(t_0) - \mathbf{b}_1(t)\int_{t_0}^{t} \mathbf{b}_1^T(\tau)\mathbf{v}(\tau) = \mathbf{0},$$

which, left-multiplying by $\mathbf{R}(t)$ and using the inverse transformation $\mathbf{z}(t) = \mathbf{T}^{-1}(t)\boldsymbol{\chi}(t)$, can be manipulated to yield

$$\mathbf{R}(t_0)\mathbf{z}_{M_1}(t_0) - \boldsymbol{\Delta}\mathbf{p}(t,t_0) - {}^{E}\mathbf{b}_1(t)z_{R_1}(t_0) - {}^{E}\mathbf{b}_1(t)\int_{t_0}^{t} \mathbf{b}_1^T(\tau)\mathbf{v}(\tau) = \mathbf{0}.$$

Finally, recalling that the output of the original nonlinear system (6.1) can be expressed as (6.2) if Assumption 6.1 holds, and using the explicit relations of the nonlinear state with the initial conditions in (6.11) and (6.12), it is possible to write

$$\mathbf{R}(t_0)\mathbf{p}_1(t_0) - \boldsymbol{\Delta}\mathbf{p}(t,t_0) - {}^{E}\mathbf{b}_1(t)\|\mathbf{p}_1(t_0)\| - {}^{E}\mathbf{b}_1(t)\int_{t_0}^{t} \mathbf{b}_1^T(\tau)\mathbf{v}(\tau) = \mathbf{0}.$$

The comparison of these two expressions yields

$$\mathbf{R}(t_0)\Big(\mathbf{z}_{M_1}(t_0) - \mathbf{p}_1(t_0)\Big) = {}^{E}\mathbf{b}_1(t)\Big(z_{R_1}(t_0) - \|\mathbf{p}_1(t_0)\|\Big),$$

for all $t \in \mathcal{T}$. If the conditions of the theorem hold, then ${}^{E}\mathbf{b}_1(t)$ is not the same for all time, and thus both the left-hand and right-hand sides must be zero, which implies that

$$\mathbf{z}_{M_1}(t_0) = \mathbf{p}_1(t_0)$$

and

$$z_{R_1}(t_0) = \|\mathbf{p}_1(t_0)\|.$$

The initial conditions of both systems are proven to be the same, and, as their dynamics are the same by construction, the states do indeed correspond to each other and the state constraints (6.3) are imposed for visible landmarks by the dynamics when the conditions of the theorem hold, hence concluding the first part of the proof.

The proof of the second part of the theorem follows directly from the first. As the state of the LTV system (6.5) corresponds to that of the underlying nonlinear system when discarding the non-visible landmarks if the conditions of the theorem hold, the estimates of an observer with globally exponentially stable error dynamics will converge exponentially fast not only to the true LTV state but also to the state of the nonlinear system (6.1). □

The two previous results address the observability of the LTV system and the correspondence between the states of that and of the original nonlinear system. Combining the two results, it is possible to determine in which conditions the nonlinear system is observable, when the non-visible landmarks are discarded. The following theorem addresses this issue.

Theorem 6.3 (Observability of the BO-SLAM nonlinear system)**.** *The nonlinear system* (6.1) *is observable, when discarding the non-visible landmarks, if and only if the conditions of Theorem 6.1 and Assumption 6.1 hold.*

Proof. The sufficiency part of the proof is readily provided by Theorems 6.1 and 6.2. The former establishes conditions for the observability of the LTV system (6.5) and the latter relates the state of that system to that of the nonlinear system in analysis. For the necessity part of the proof, note that if, for the same input, there exist two initial conditions that lead to the same output at all times, then the system is not observable. Assume then that the conditions of the theorem do not hold, and recall the output of the nonlinear system (6.1). Given that ${}^{E}\mathbf{b}_1(t)$ is constant, the vehicle can only move in the line defined by its initial position and the landmark, which means that any initial state corresponding to a landmark defined in that line will yield the same output, thus implying that the system is not observable when the conditions do not hold, or, conversely, if the system is observable, the conditions must hold. □

6.3.3 Filter design

The results of the previous subsection show that, in certain conditions with physical insight, the augmented system is equivalent to the nonlinear system, and that if a filter with GES error dynamics can be constructed to the LTV system, it will also be applicable to the original nonlinear system. From Theorem 2.2, it is known that the error dynamics of the Kalman filter for linear time-varying systems are globally exponentially stable if the pair $(\mathbf{A}(t), \mathbf{C}(t))$ is uniformly completely observable, a form of observability stronger than the ones addressed previously. This last theorem addresses the uniform complete observability of the LTV system.

Theorem 6.4 (Uniform complete observability of the BO-SLAM LTV system)**.** *Let $\mathcal{T}_\delta := [t, t+\delta]$. The pair $(\mathbf{A}(t), \mathbf{C}(t))$ associated with the system (6.5), regarded as LTV, is uniformly completely observable if there exist positive constants δ and α_b such that, for all $t \geq t_0$, it is possible to find a $t_1 \in \mathcal{T}_\delta$ for which the absolute bearing to the visible landmark respects*

$$\left\| \int_t^{t_1} {}^E\dot{\mathbf{b}}_1(\tau) \mathrm{d}\tau \right\| \geq \alpha_b. \tag{6.13}$$

Proof. As was done in the proof of Theorem 6.1, the transformed system (6.6) will be used to simplify the analysis, and the obtained results can be applied to the system (6.5).

Recall the definition of uniform complete observability in (2.4) reproduced here

$$\underset{\substack{\delta > 0 \\ \alpha > 0}}{\exists} \quad \underset{t \geq t_0}{\forall} \quad \underset{\substack{\mathbf{c} \in \mathbb{R}^{n_x} \\ \|\mathbf{c}\| = 1}}{\forall} : \quad \mathbf{c}^T \boldsymbol{\mathcal{W}}(t, t+\delta)\mathbf{c} \geq \alpha, \tag{6.14}$$

meaning that, in contrast with the observability definition used in Theorem 6.1, the Gramian must have uniform bounds at all times. The proof follows by exhaustion, by analysing the quadratic form in the previous expression for all the possible cases of unit vectors $\mathbf{c}$ for all time. Proposition A.1 (available on Appendix A) is helpful in that analysis, as it applies to the quadratic form in study, and therefore it suffices to show that the norm of $\mathbf{f}(\tau, t)$, given by

$$\|\mathbf{f}(\tau, t)\|^2 = \left\| \mathbf{c}_m - {}^E\mathbf{b}_1(\tau) c_r \right\|^2 \tag{6.15}$$

is lower bounded uniformly in time by some $\alpha^* > 0$ at some time $\tau \in \mathcal{T}_\delta$ for all the possible cases of unit vectors $\mathbf{c}$. It is necessary to address $\|\mathbf{f}(\tau, t)\|$ differently depending on $\mathbf{c}_m$ and $\mathbf{c}_r$. If (6.15) is evaluated at $\tau = t$, it yields

$$\begin{aligned} \|\mathbf{f}(t, t)\|^2 &\geq \|\mathbf{c}_m\|^2 + |c_r|^2 - 2\|\mathbf{c}_m\| \, |c_r| \\ &\geq \|\mathbf{c}_m\| \left(\|\mathbf{c}_m\| - |c_r| \right) + |c_r| \left(|c_r| - \|\mathbf{c}_m\| \right). \end{aligned} \tag{6.16}$$

Consider now that $\|\mathbf{c}_m\| - |c_r| \geq \alpha_m$ for some positive constant $\alpha_m < 1$. Note that, given that $\mathbf{c}$ is a unit vector and $\|\mathbf{c}_m\|^2 + |c_r|^2 = 1$, the worst case occurs when this difference is at its lowest, thus reducing the positive contribution above and increasing the size of $|c_r|$ and the negative contribution. Replacing this in (6.16) yields

$$\begin{aligned} \|\mathbf{f}(t, t)\|^2 &\geq \|\mathbf{c}_m\| \alpha_m - \left(\|\mathbf{c}_m\| - \alpha_m \right) \alpha_m \\ &\geq \alpha_m^2. \end{aligned}$$

The case where $|c_r| - \|\mathbf{c}_m\| \geq \alpha_m$ is very similar, yielding

$$\begin{aligned}\|\mathbf{f}(t,t)\|^2 &\geq -\left(|c_r| - \alpha_m\right)\alpha_m + |c_r|\alpha_m \\ &\geq \alpha_m^2.\end{aligned}$$

For the final case, where $\left|\|\mathbf{c}_m\| - |c_r|\right| < \alpha_m$, the condition (6.13) of the theorem will be necessary. Consider the norm of $\mathbf{f}(\tau, t)$ once more evaluated at $\tau = t$,

$$\begin{aligned}\|\mathbf{f}(t,t)\|^2 &= \left\| \mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r + \int_t^{t_1} {}^E\dot{\mathbf{b}}_1(\tau)\mathrm{d}\tau\, c_r \right\|^2 \\ &\geq \left\| \int_t^{t_1} {}^E\dot{\mathbf{b}}_1(\tau)\mathrm{d}\tau\, c_r \right\|^2 - 2\left\|\mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r\right\| \left\| \int_t^{t_1} {}^E\dot{\mathbf{b}}_1(\tau)\mathrm{d}\tau\, c_r \right\| \\ &\geq \alpha_b|c_r|\left(\alpha_b|c_r| - 2\left\|\mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r\right\|\right).\end{aligned}$$

Noting that, in the case at hand, $|c_r| > \frac{1}{2}\sqrt{2-\alpha_m^2} - \frac{\alpha_m}{2} := \alpha_r > 0$, which is obtained using $\left|\|\mathbf{c}_m\| - |c_r|\right| < \alpha_m$ and $\|\mathbf{c}_m\|^2 + |c_r|^2 = 1$, it is possible to write

$$\|\mathbf{f}(t,t)\|^2 \geq \alpha_b\alpha_r\left(\alpha_b\alpha_r - 2\left\|\mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r\right\|\right).$$

If $\|\mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r\| < \alpha_1$ with $\alpha_1 < \frac{1}{4}\alpha_b\alpha_r$ this becomes

$$\|\mathbf{f}(t,t)\|^2 > \frac{\alpha_b^2\alpha_r^2}{2}.$$

If, on the other hand, $\|\mathbf{c}_m - {}^E\mathbf{b}_1(t_1)c_r\|$ is lower bounded by α_1, then

$$\|\mathbf{f}(t_1,t)\|^2 \geq \alpha_1^2.$$

This concludes the enumeration of all the possible cases. Positive lower bounds were found for $\|\mathbf{f}(\tau,t)\|$ on different instants of $\mathcal{T}_\delta$ depending on the case, which, using Proposition A.1 implies that the integral in (6.7) is in fact lower bounded for unit vectors $\mathbf{c}$, and hence (6.14) is true. Thus, the pair $(\boldsymbol{\mathcal{A}}(t), \boldsymbol{\mathcal{C}}(t))$ is uniformly completely observable, and, as its associated system is related to system (6.5) through a Lyapunov transformation, the pair $(\mathbf{A}(t), \mathbf{C}(t))$ is also uniformly completely observable, thus concluding the proof of the sufficiency part of the theorem.

The proof of the necessity part of the theorem follows by contraposition. The objective is to negate the conditions of the theorem, and show that under that hypothesis the pair $(\mathbf{A}(t), \mathbf{C}(t))$ is not uniformly completely observable. For that purpose consider the negation of the condition of the theorem, given by

$$\underset{\substack{\delta > 0 \\ \beta > 0}}{\forall} \;\; \underset{t \geq t_0}{\exists} \;\; \underset{t_1 \in \mathcal{T}_\delta}{\forall} : \quad \left\| \int_t^{t_1} {}^E\dot{\mathbf{b}}_1(\tau)\mathrm{d}\tau \right\| < \beta, \tag{6.17}$$

and let us analyse $\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c}$ for a particular $\mathbf{c}$.

$$\begin{aligned}\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} &= \int\limits_{t}^{t+\delta}\left\|\mathbf{c}_m - {}^E\mathbf{b}_1(\tau)c_r\right\|^2 \mathrm{d}\tau \\ &= \int\limits_{t}^{t+\delta}\left\|\mathbf{c}_m - \left({}^E\mathbf{b}_1(t) + \int\limits_{t}^{\tau}{}^E\dot{\mathbf{b}}_1(\sigma)\mathrm{d}\sigma\right)c_r\right\|^2 \mathrm{d}\tau\end{aligned}$$

Choosing $\mathbf{c}_m = {}^E\mathbf{b}_1(t)c_r$ and $c_r = \frac{\sqrt{2}}{2}$, it is possible to write

$$\mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} = \frac{1}{2}\int\limits_{t}^{t+\delta}\left\|\int\limits_{t}^{\tau}{}^E\dot{\mathbf{b}}_1(\sigma)\mathrm{d}\sigma\right\|^2 \mathrm{d}\tau. \tag{6.18}$$

Recall now the negation of the conditions of the theorem, (6.17). It is clear that the norm therein is upper-bounded for any t_1 between t and $t+\delta$ for a given $t \geq t_0$. In particular, this is also true for $t_1 = \tau$, which means that there exists a $t \geq t_0$ such that the norm in the integral in (6.18) is also upper-bounded, i.e.,

$$\underset{\substack{\delta>0\\ \epsilon>0}}{\forall} \quad \underset{t\geq t_0}{\exists} \quad \underset{\substack{\mathbf{c}\in\mathbb{R}^{n_x}\\ \|\mathbf{c}\|=1}}{\exists} : \qquad \mathbf{c}^T\boldsymbol{\mathcal{W}}(t,t+\delta)\mathbf{c} < \frac{\beta^2\delta}{2} := \epsilon, \tag{6.19}$$

where $\beta = \sqrt{\frac{2\epsilon}{\delta}}$. Therefore, when the conditions of the theorem do not hold, it is possible to find a unit $\mathbf{c}$, e.g. $\mathbf{c} = \begin{bmatrix}\frac{\sqrt{2}}{2}{}^E\mathbf{b}_1^T(t) & \frac{\sqrt{2}}{2}\end{bmatrix}^T$, such that (6.19) is true, i.e., the pair $(\mathbf{A}(t), \mathbf{C}(t))$ is not uniformly completely observable. By contraposition, the conditions of the theorem are necessary for the uniform complete observability of the system (6.5). ☐

A Kalman filter for LTV systems can now be implemented for (6.4), and it is done so in its discrete version as described in Subsection 2.3.2. As in the RO-SLAM implementation, described in Section 4.4, the discretization is of the forward Euler variety, using the exponential map and a constant angular velocity to discretize the rotation of the sensor-based frame between instants. With this information, the dynamics matrix of the discrete system (2.9) is, in this case,

$$\hat{\mathbf{F}}_{F_k} = \begin{bmatrix}\mathbf{F}_{M_k} & \mathbf{0}_{n_M\times n_R}\\ \mathbf{0}_{n_R\times n_M} & \mathbf{I}_{n_R}\end{bmatrix}, \tag{6.20}$$

where $\mathbf{F}_{M_k} = \mathrm{diag}\left(\mathbf{R}_{k+1}^T\mathbf{R}_k, \ldots, \mathbf{R}_{k+1}^T\mathbf{R}_k\right)$ and $\mathbf{R}_{k+1}^T\mathbf{R}_k = \mathrm{e}^{-\mathbf{S}(\boldsymbol{\omega}_k)T_s}$.

The remaining aspects of the implementation are as delineated in Subsection 2.3.2 with further details in the following section.

6.4 Algorithm implementation

This section addresses the practical implementation of the filter introduced in the previous section. Aside from the Kalman filter that serves as the estimation engine of the algorithm, there are several other relevant components that need to be discussed. Given that the primary exteroceptor in the proposed implementation is a camera, and only natural landmarks are being dealt with, there needs to exist a feature detection process and a consequent landmark association stage. Therefore, this section starts by explaining how the algorithm extracts bearings from the images provided by the camera, bearings that are then used in the update step of the Kalman filter.

6.4.1 Obtaining bearing measurements from a camera

The first step in obtaining bearing measurements to landmarks from an image is *feature detection*. This step takes place every time a new image is available from the installed camera. This image is fed to a SURF implementation [BETG08], which detects points-of-interest (features) on 2-D pictures of the environment (see Section 3.4.1 for further details). Aside from the measured feature locations in image coordinates $\mathbf{f}_{m_i}$, SURF also provides a metric to determine which features have higher *quality* (the Hessian), a characteristic that distinguishes types of features (the Laplacian), and a 64-dimensional descriptor for association. After feature detection, some processing is necessary before the insertion of these measurements in the filter. Note, however, that this processing serves only the purpose of transforming the features from image coordinates to bearings in body-fixed coordinates, as the landmarks are initialized undelayed.

Transforming image coordinates to bearings This step depends greatly on the used camera model. This work uses a model [HS97] that relates the undistorted normalized coordinates $\mathbf{f}_i \in \mathbb{R}^2$ with the measured coordinates, given by

$$\begin{cases} \mathbf{f}_{d_i} = \left(1 + k_1^c \|\mathbf{f}_i\|^2 + k_2^c \|\mathbf{f}_i\|^4 + k_5^c \|\mathbf{f}_i\|^6\right) \mathbf{f}_i + \mathbf{d}_i \\ \mathbf{d}_i = \begin{bmatrix} 2k_3^c x_i y_i + k_4^c \left(\|\mathbf{f}_i\|^2 + 2x_i^2\right) \\ 2k_4^c x_i y_i + k_3^c \left(\|\mathbf{f}_i\|^2 + 2y_i^2\right) \end{bmatrix} \\ \mathbf{f}_{m_i} = \begin{bmatrix} f_1^c & \alpha_c f_1^c & c_1^c \\ 0 & f_2^c & c_2^c \end{bmatrix} \begin{bmatrix} \mathbf{f}_{d_i} \\ 1 \end{bmatrix} \end{cases} . \tag{6.21}$$

where $\mathbf{f}_{d_i} \in \mathbb{R}^2$ are the distorted feature coordinates, $\mathbf{d}_i$ is the tangential distortion, and the remaining quantities are the intrinsic model parameters: f_j^c is the camera focal length on each axis $j = 1, 2$; the c_j^c, with $j = 1, 2$, are the principal point coordinates; k_j^c, for $j = 1, \ldots, 5$, are the distortion coefficients; and α_c is the skew coefficient.

This model was chosen because it is the one used by the authors of the *Rawseeds* dataset [BBF+06, CFG+09] from which the data for the experiments detailed in this chapter was obtained. To solve (6.21) for the normalized feature coordinates, consequently removing distortion and projection, the calibration toolbox [Bou15] is employed.

From the normalized feature locations, given by

$$\mathbf{f}_i = \begin{bmatrix} \frac{x_i}{z_i} & \frac{y_i}{z_i} \end{bmatrix}^T,$$

it is possible to compute a normalized range,

$$\bar{r}_i = \sqrt{\|\mathbf{f}_i\|^2 + 1} = |z_i|^{-1}\sqrt{x_i^2 + y_i^2 + z_i^2}.$$

Noting that z_i is always positive due to the limited field of view of the camera, it can be seen that this is also equal to the norm of the three-dimensional point normalized by the z coordinate. Careful observation of the information available, $(\mathbf{f}_i, \bar{r}_i)$, it is clear that it is possible to compute directions to the landmarks, i.e., $\frac{\mathbf{p}_i}{\|\mathbf{p}_i\|}$, which are then given by

$$\mathbf{b}_i = \frac{\mathbf{p}_i}{\|\mathbf{p}_i\|} = \begin{bmatrix} \frac{x_i}{z_i} \\ \frac{y_i}{z_i} \\ \frac{z_i}{z_i} \end{bmatrix} \frac{z_i}{\|\mathbf{p}_i\|} = \begin{bmatrix} \mathbf{f}_i \\ 1 \end{bmatrix} \frac{1}{\bar{r}_i}.$$

Following this process, the remaining step before the actual update step of the filter is landmark association, the subject of the sequel.

6.4.2 Landmark association

In typical bearing-only algorithms, data association is sometimes not an issue, as the measurements originate from identified acoustic or RF signals. However, when monocular vision is utilized to obtain the bearing measurements, as is the case here, association becomes extremely important. Due to this importance, several algorithms stemming from a fertile research effort in the field have appeared. This is not the focus of this dissertation and any method could be employed. A brief description of the chosen methodology is included for the sake of completeness.

The first step in this process is to define the sets to associate with. A maximum number N_{max} of observed features is defined and the N_{max} features with highest metric are chosen. Then, the set of landmarks in the current map to associate is defined as the ones within the field-of-view of the camera (only azimuth and elevation are checked). Both sets are divided in two, depending on the sign of the Laplacian.

The next step avails itself of the possibilities permitted by the image processing procedure used, namely the description of the landmarks by labels invariant and robust to several transformations. This follows the previous successful implementation in Chapter 3, where the landmarks are first associated with respect to their descriptor, with individual compatibility gating based on the Mahalanobis distance of the measurement (6.2) (see Definition 3.1).

Given that the number of features detected by SURF can be very high, and to avoid the size of the filter state getting prohibitively large, after the association procedure only the associated bearings are directly used as observations in the filter. All non-associated measurements are discarded as long as the total number of *valid* bearings is greater than a predefined threshold. If it is not, then the best new landmarks are added until the threshold is passed. This is the landmark initialization. The guaranteed convergence of the filter frees the user from careful considerations on how to initialize each landmark in the filter. A possible method is depicted in Figure 6.3, and includes the following steps: (i) define a minimum (R_m) and maximum range (R_M) for the camera; (ii) define a circle of error in the image plane that is equivalent to a cone in 3-D space with an aperture angle of α centered on the measured bearing; and (iii) approximate a Gaussian distribution to this volume using the ellipsoid corresponding to the 3σ uncertainty bound, i.e.,

$$\begin{cases} \hat{\mathbf{x}}_{M_i}(k_0) := \dfrac{R_M + R_m}{2}\mathbf{b}_i(k_0) \\ \Sigma_{M_i}(k_0) := \mathbf{U}(\mathbf{b}_i(k_0))\,\mathrm{diag}\left(\left(\dfrac{1}{6}(R_M - R_m)\right)^2, \left(\dfrac{1}{6}\|\hat{\mathbf{x}}_{M_i}(k_0)\|\sin\alpha\right)^2, \left(\dfrac{1}{6}\|\hat{\mathbf{x}}_{M_i}(k_0)\|\sin\alpha\right)^2\right)\mathbf{U}(\mathbf{b}_i(k_0))^T \end{cases},$$

where $\mathbf{U}(\mathbf{b}_i(k_0))$ is the rotation matrix that aligns the ellipsoid with the bearing measurement. In the particular implementation of the filter used in the experimental results of the next section, the above method was used with the detail that the center of the ellipsoid was placed in a random point in the line defined by $\mathbf{b}_i(k_0)$, to better showcase the convergence of the algorithm.

6.4.3 Map maintenance

Even though the association process is designed to try to reduce the number of landmarks in the state, the process employed is rather naïve. Furthermore, it favours the inclusion of spurious measurements, when the non-associated features are inserted in the filter. This can lead to a very large state with a great number of spurious features. To deal with this issue, a simple test was implemented to check whether a landmark is spurious or not, depending on two counters: (i) a visibility counter that is increased each time a landmark is associated with a new measurement; and (ii) a counter that is increased each iteration that a landmark is not associated and reset every time it is, thus counting the time without observations. When the second counter reaches $T_{\max}$ and the first counter is lower than $N_{\min}$, the landmark is considered spurious and discarded.

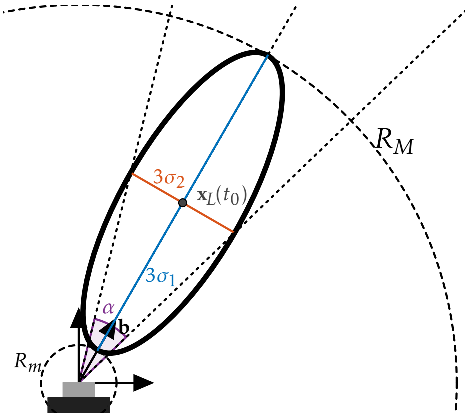


Figure 6.3: A schematic of the construction of the initial covariance. R_m and R_M are the minimum and maximum range of the camera, α is the noise error, and σ_i^2 are the eigenvalues of the new covariance.

6.4.4 Complexity reduction

As mentioned previously, the size of the filter state can get very high very fast which will eventually slow down the algorithm. As the filter operates in the body-fixed frame, all the non-visible landmarks must be propagated in each step according to the motion of the vehicle, This can be very time consuming for too large a state. This is in opposition to the traditional inertial SLAM approach, where the prediction step is trivial (the landmarks are static) but the update step is very computationally expensive as all landmarks (and their covariances) need to be updated.

Since the dynamics of each landmark (and its associated range state) are independent of each other for observability purposes, it can be thought that that independence will be apparent in the stochastic filtering setting. The following result shows in which conditions that insight is correct.

Lemma 6.5 (Preservation of the sparse structure by the Kalman filter.)**.** *Consider the Kalman filter for the linear time-varying discrete system* (2.9) *particularized for the BO-SLAM case. If the initial covariance* Σ_0 *and the process covariance* $\boldsymbol{\Xi}_k$ *for all* $k > 0$ *have the following structure*

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{M_1} & \cdots & \mathbf{0} & \mathbf{M}_{M_1R_1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{M}_{M_N} & \mathbf{0} & \cdots & \mathbf{M}_{M_NR_N} \\ \mathbf{M}_{R_1M_1} & \cdots & \mathbf{0} & M_{R_1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{M}_{R_NM_N} & 0 & \cdots & M_{R_N} \end{bmatrix}. \tag{6.22}$$

and the measurement covariance $\boldsymbol{\Theta}_k$ *is block-diagonal, then both the predicted covariance* $\Sigma_{k+1|k}$ *and the updated covariance* $\Sigma_{k|k}$ *maintain the structure in* (6.22) *for all* $k > 0$.

Proof. The proof is focused in showing that at a given instant $k > 0$ the update and predict steps of the Kalman filter do not change the structure of the covariance if the covariance at $k-1$ already had the structure of (6.22). It can be found in Appendix B.1. □

Using this result, if the design quantities Σ_0, $\boldsymbol{\Theta}_k$, and $\boldsymbol{\Xi}_k$ are chosen accordingly it is possible to simplify the computations by exploiting the sparsity of the covariance matrix, as only a few entries need to be computed, yielding, for all $i \in \mathcal{M}$,

$$\begin{cases} \Sigma_{M_{i\,k+1|k}} = \mathbf{R}_{k+1}^T\mathbf{R}_k\Sigma_{M_{i\,k|k}}\mathbf{R}_k^T\mathbf{R}_{k+1} + \boldsymbol{\Xi}_{M_{i\,k+1|k}} \\ \Sigma_{R_{i\,k+1|k}} = \Sigma_{R_{i\,k|k}} + \boldsymbol{\Xi}_{R_{i\,k+1|k}} \\ \Sigma_{R_iM_{j\,k+1|k}} = \Sigma_{R_iM_{j\,k|k}}\mathbf{R}_k^T\mathbf{R}_{k+1} + \boldsymbol{\Xi}_{R_iM_{j\,k+1|k}} \end{cases}.$$

If it is necessary to include more cross-covariance terms, for example, the structure

$$\hat{\Sigma}_{k+1|k} = \begin{bmatrix} \Sigma_{M_{1\,k+1|k}} & \cdots & \mathbf{0} & \\ \vdots & \ddots & \vdots & \Sigma_{MR_{k+1|k}} \\ \mathbf{0} & \cdots & \Sigma_{M_{N\,k+1|k}} & \\ & \Sigma_{RL_{k+1|k}} & & \Sigma_{R_{k+1|k}} \end{bmatrix}$$

can still be assumed for the propagated covariance. This allows the computation time to decrease greatly, given that the prediction step is the greatest contributor for the computational cost. This becomes more apparent especially for large maps, such as the one presented in the experimental section of this chapter. In this particular unoptimized implementation, designed to demonstrate experimentally the performance and convergence properties of the sensor-based BO-SLAM filter, the algorithm scales with N^2.

6.5 Simulation results and discussion

In order to allow a better understanding of some of the properties of the algorithm proposed in this chapter and assess its performance, a simulated environment was devised and several realistic simulations were performed. This section details those simulations. The results of a typical run in the simulated environment are presented and discussed, followed by the results of simulations with different input noise levels.

6.5.1 Setup

The chosen environment, as in previous chapters and as detailed in Appendix C, tries to emulate the fifth floor of the North Tower at IST. It consists of a 16 m by 16 m by 3 m corridor. 36 landmarks were put in notable places such as corners and doors, with random heights. The aerial vehicle starts stopped at the ground, and after taking off makes several laps around the corridor. It completes a loop of 55 m in 124 s, and the total trajectory is 5 loops at 0.440 m/s.

In order to better approximate the simulation to the reality, the field of view of the vehicle was limited to 90° horizontally and vertically with a range of up to 20 m. The presence of walls was also taken into account, meaning that the landmarks are only visible during a limited period of time in each loop. Furthermore, to demonstrate the performance of the filter, discounting effects of association, the measurements are assumed to be tagged and the association to be perfect.

The bearing measurements are obtained by rotating the true bearing about random vectors of a random zero-mean angle with Gaussian distribution with standard deviation of 1°. The remaining measurements are corrupted with additive zero mean white noise. The standard deviation of the noise corrupting the linear velocity is 0.01 m/s, and that of the angular velocity is 0.15°/s. All measurements are obtained at 20 Hz.

The parameters of the BO-SLAM filter are those of the Kalman filter, and consist of the model disturbance covariance $\mathbf{\Xi} = T_s \operatorname{diag}\left(10^{-2}\mathbf{I}_{n_L}, 10^{-3}\mathbf{I}_{n_R}\right)$, and of the measurement noise covariance $\mathbf{\Theta} = \operatorname{diag}\left(10^{-2}\mathbf{I}_3, \mathbf{I}_{n_O}\right)$.

6.5.2 Typical run results

Figures 6.4 and 6.5 condense the results of a typical run. The first is the top-view of the map with the 3σ uncertainty ellipses in red, purple, and green depending whether they are ob-

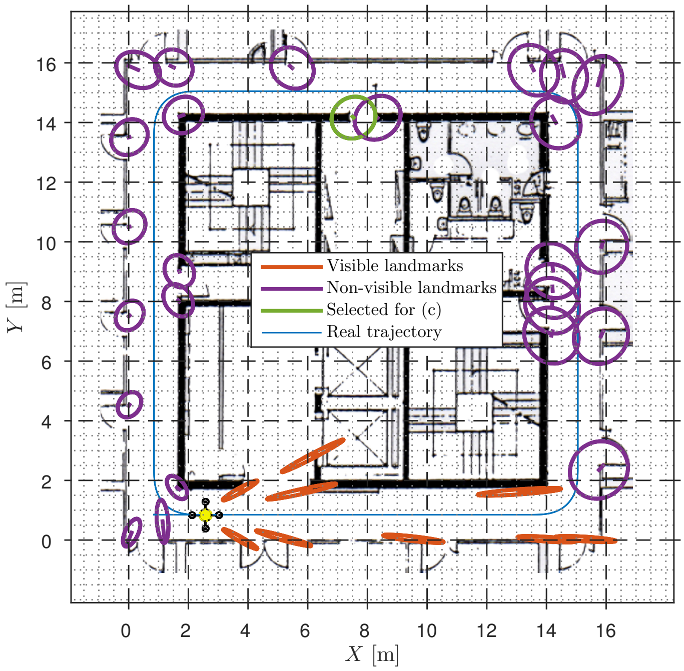
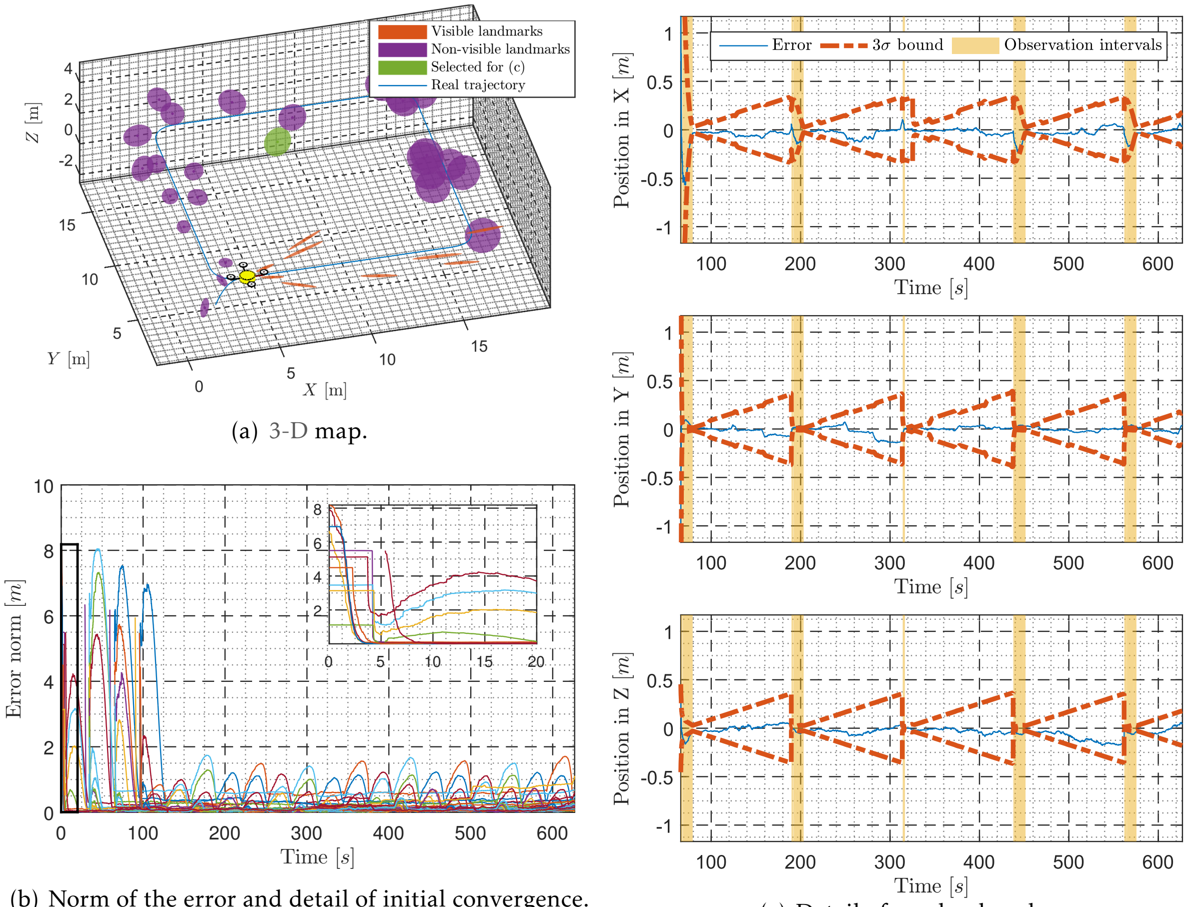


Figure 6.4: Top view of the estimated map in a typical simulation. The ellipsoids represent the 3σ bounds of the current estimated map. The lines inside the ellipsoids represent the distance from the current estimate to the true position of each landmark.

served in that instant or not, including the real trajectory of the vehicle in blue, and the pose of the vehicle at that moment, that is represented by the yellow quadrotor. Note that the ellipsoids surround the true values, as they should in a consistent filter. At the top left of Figure 6.5, Figure 6.5(a) shows the mapped landmarks in 3-D with their 3σ uncertainty ellipsoids and the trajectory of the vehicle in blue. The landmarks that were currently being observed in that instant are depicted in red and the rest are depicted in purple. One selected landmark is shown in green, and the time evolution of its estimates is shown in detail in Figure 6.5(c). There the estimation error is shown in blue for each coordinate and the Kalman filter uncertainty bounds are depicted in red. To enable a better understanding of the behaviour of the convergence observed therein, the intervals where the landmark was observed are shown in vertical yellow strips. As expected, the estimates converge quite fast when the landmark is first observed and then the uncertainty grows with the motion of the vehicle away from the landmark, converging again at a new sighting. Finally, Figure 6.5(b) shows the convergence of the norm of the estimation error for all landmarks. At the right top of that figure a detail of the first 20 seconds is depicted. It can be seen that, as predicted by the theoretical results presented in Section 6.3, even when initial estimate is far off (up to 8 meters of error), the error converges quite fast, depending on how much time each landmark was observed.



(a) 3-D map.

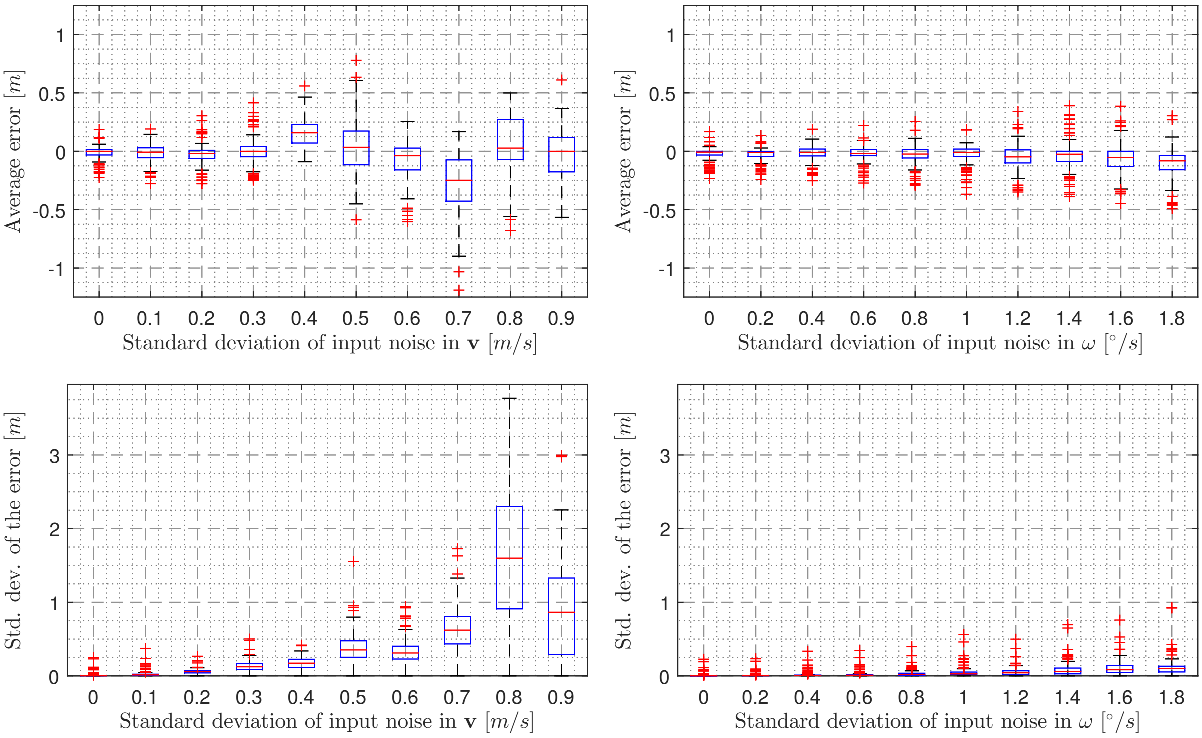
(b) Norm of the error and detail of initial convergence.

(c) Detail of one landmark.

Figure 6.5: Results of a typical simulation.

6.5.3 Different input noise results

The simulation results detailed above confirm the theoretical results of the paper and provide quantitative data for its evaluation, but do not allow a thorough understanding of the performance of the algorithm in the presence of different input noise levels. For that reason, and to illustrate how the algorithm deals with noise levels such as the ones in Figure 6.7, a series of trials were run with varying levels of noise. The results are depicted in Figure 6.6, where the distribution of the estimation error for 10 different noise levels is presented. The top figures show the distribution of the average landmark estimation error for all coordinates after initial convergence, and the bottom figures show the distribution of the standard deviation. In each column, one type of input noise was varied, while the others were kept constant. For both, the bearing measurement noise is the same as in Figure 6.5. For Figure 6.6(a) the standard deviation of the noise in the linear velocity measurements was increased from 0 to 0.9 m/s while the standard deviation of angular rate noise was constant at 0.15°/s, and for Figure 6.6(b) the standard deviation of the noise in the angular velocity measurements was increased from 0 to 1.8 °/s while the standard deviation of linear velocity noise was constant at 0.01 m/s. These figures show the impact of the noise on the estimation error, and how differently noise in the



(a) Varying noise levels in linear velocity

(b) Varying noise levels in angular velocity

Figure 6.6: Sensitivity of the estimation error with respect to noise in the linear velocity and angular velocity measurements. The red horizontal lines represent the median, the blue boxes delimits the 25th and 75th percentiles, the black whiskers are extended at 2.7σ, and the red crosses are all the individual occurrences beyond that bound. On the left, angular velocity noise standard deviation is fixed at $0.15°/s$, and on the right the linear velocity noise standard deviation is fixed at 0.01 m/s.

linear and angular velocities affects the estimates. In fact, the filter deals quite well with large levels of noise in the angular rates, and is much more sensitive to very high levels of noise in the linear velocity measurements. However, the average error for each landmark coordinate is kept within 1 meter for both cases, and the standard deviation is mostly contained below 2 meters. Note that the parameters of the filter were not adjusted to each noise level, which indicates that these results could, in theory, be better.

6.6 Experimental results

The theoretical results presented in Section 6.3 lead to the design of a BO-SLAM filter implemented as described in Section 6.4. This algorithm was tested with real data from datasets acquired by the *Rawseeds* Project [BBF+06, CFG+09], which aims "to build benchmarking tools for robotic systems"[2] while being "mainly targeted at the problems of localization, mapping and SLAM in robotics". This section covers the relevant details of the dataset used and presents the results of the experiments, while also providing a discussion of the latter.

---

[2] `http://www.rawseeds.org`

6.6.1 Rawseeds dataset

There are several available datasets in the *Rawseeds* project, including outdoor data obtained in the Politecnico di Milano campus and indoor data obtained "in a pair of building belonging to the Università di Milano-Bicocca, in Milan (Italy). The part of the location explored by the robot during the gathering of the datasets includes one floor of an office building and the corresponding floor of a nearby building. The two buildings are connected by roofed, glass-walled bridges. They sport a wide range of different of architectural features." For the purpose of validating the proposed algorithm, an indoors dataset with natural lighting and no purposefully dynamic objects, named `Bicocca_2009-02-25b`, was chosen. In this dataset, the traversed path of 774 m includes several small loops, aside from the main loop which is covered by the robot in 29 minutes. The top view of the schematic of the area and the sketched trajectory can be found in Figure 6.8 along with pictures of several regions in the area illustrating some of the difficulties in visual navigation, namely corridors without many distinguishable features, very dark places, very bright places, among other challenges.

The robot used as a data acquisition platform for the dataset, depicted in Figure 6.8(b), is equipped with a very complete sensor suite, comprising 3 *Videre Design* black and white cameras in a forward-looking trinocular configuration, a *Prosilica* camera with an hyperbolic mirror to achieve omni-directional vision, a SONAR belt, a *Unibrain* low-cost camera, a *Maxbotix* SONAR belt, an *Xsense* IMU, two double laser range finder configurations to cover the front and rear of the vehicle with both a cheaper solution by *Hokuyo* and a long range one by *Sick*, and wheel encoders for odometry. Aside from the various streams of data provided by the sensors on board the robot, the dataset also provides ground truth from a system based on industrial cameras, visual tags mounted on the robot, and ad hoc software in select parts of the trajectory. This ground truth is complemented by what is called *extended* ground truth, available for all the trajectory, which is computed using scan matching from the laser scanners in the robot.

The BO-SLAM algorithm herein proposed uses only directions, linear, and angular velocities. Therefore only one of these cameras, odometry and the inertial measurement unit are necessary. It is possible to obtain the angular velocity (or angular displacements) from both the wheel encoders and the inertial measurement unit, and therefore it is necessary to decide

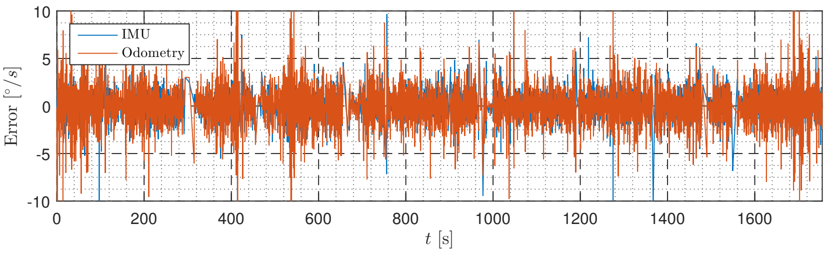
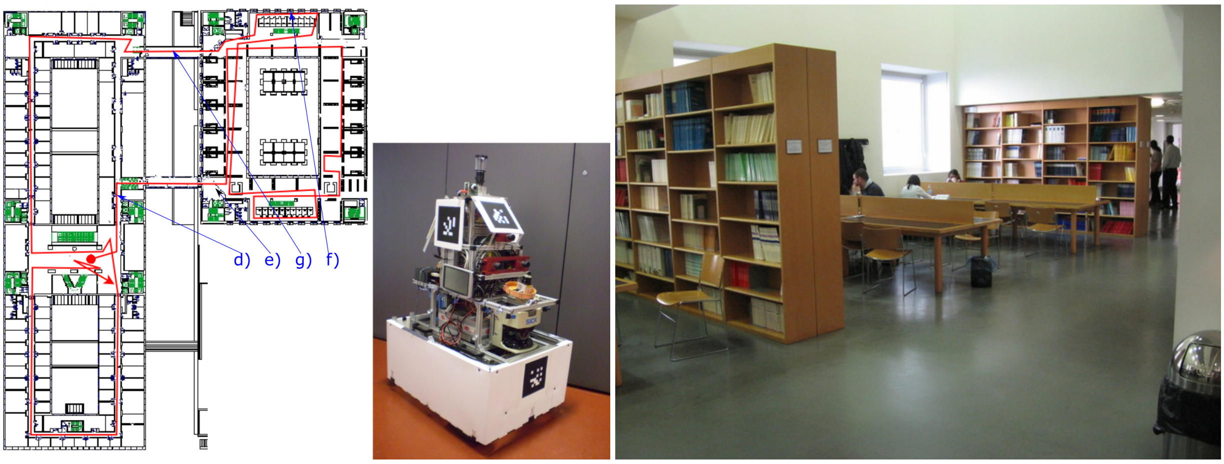


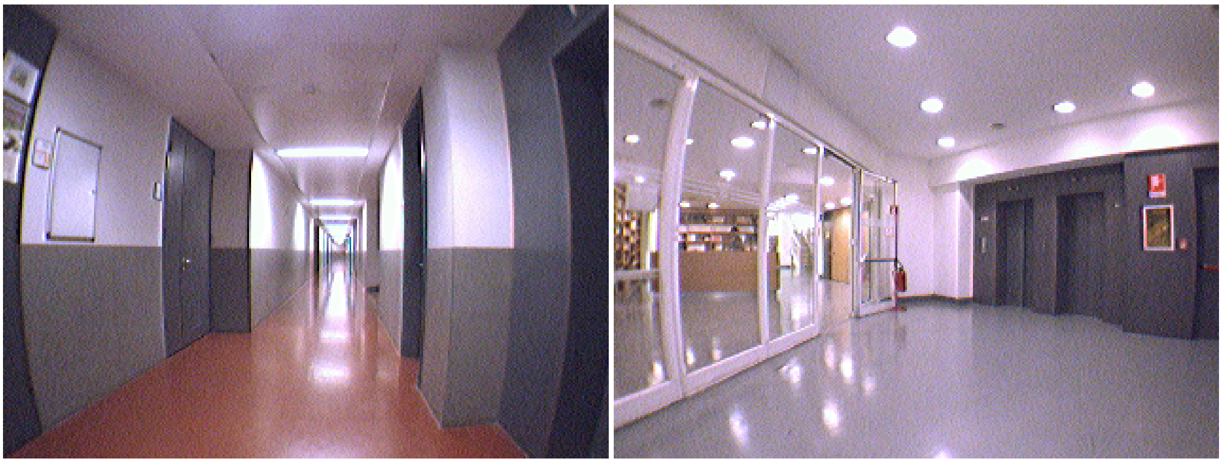
Figure 6.7: Inertial measurement unit and odometry-based angular velocities error evolution.



(a) A schematic of the top view of the area travelled by the robot in the dataset.

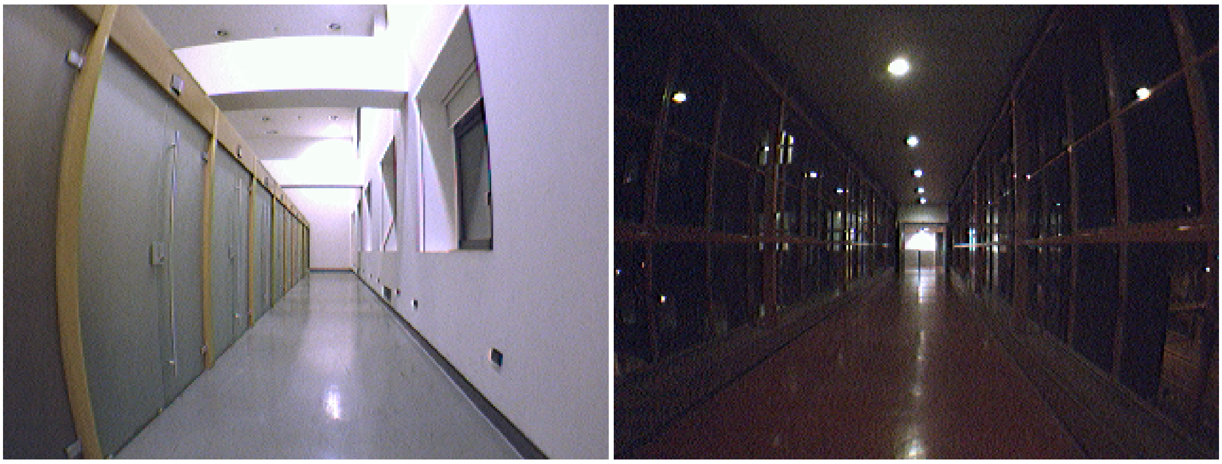
(b) The robot used in the experiments.

(c) The interior of the university library (the building on the right).



(d) A typical corridor.

(e) A very bright elevator lobby.



(f) Another typical corridor with clear repetition of features.

(g) One of the two very dark corridors between the two buildings.

Figure 6.8: Information on the dataset. Schematic of the area and the trajectory, the robot used in the experiments and examples of places visited by the robot.

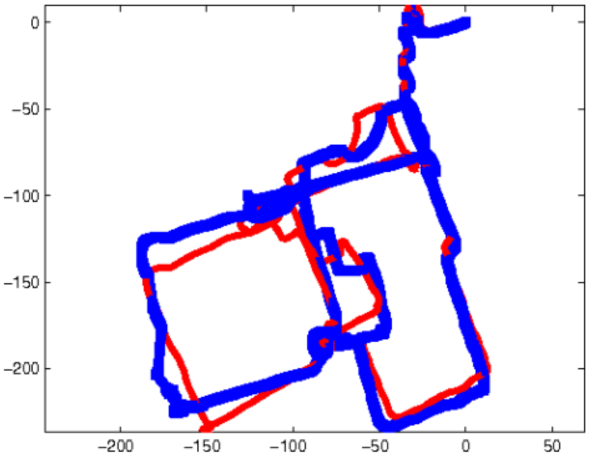


Figure 6.9: Results of dataset benchmark using Inverse Depth EKF and 1-point RANSAC. Monocular SLAM estimation from the combination of monocular camera plus wheel odometry (thin red) and ground-truth trajectory (thick blue). Reproduced from [CGM09].

which of the sources to use. For that purpose, the difference between both angular velocities and the one derived from the extended ground truth was analysed. As can be seen in Figure 6.7, even though the error has zero mean, it reaches consistently very high levels, with standard deviations of 1.398°/s (IMU) and 2.087°/s (odometry). This results in a very distorted trajectory, when dead-reckoning the body-fixed linear velocity provided by the odometry and either angular velocities. Bearing-only (and range-only) SLAM algorithms rely greatly on the quality of the ego-motion measurements that drive them, in this case the linear and angular velocities. This is particularly relevant when the field of view of the camera is limited to a region in the front of the vehicle, because, when turning, the camera will quickly lose track of features that help correct motion information. The available benchmark solution to monocular SLAM provided in the dataset (see Figure 6.9), based on the algorithms in [CDM08] and [CGDM10] shows precisely both how the monocular camera filtering results can be poor, due to the limited field-of-view and the large distance, and how the odometry is extremely deficient in turns. For these reasons, the authors decided to use an artificial measurement of angular velocity, taking the measurement provided by the scan matching procedure and adding artificial noise with a realistic standard deviation of 0.15°/s – obtained from the data in the experiments detailed in Sections 3.5 and 4.6 (see Figure 6.10). This noise level can easily be found nowadays in off-the-shelf IMUs like the *Microstrain 3DM-GX3-25*.

6.6.2 The algorithm at work

The bearing-only SLAM algorithm herein presented is a sensor-based filter, and, as such, its main product is the landmark map in the body-fixed frame. Furthermore, in the experiments to be detailed, these landmarks are naturally extracted from the environment and not placed in predefined places. Therefore, there is not a clear ground truth to validate directly the results of the algorithm, and the map can only be evaluated quantitatively through the executive

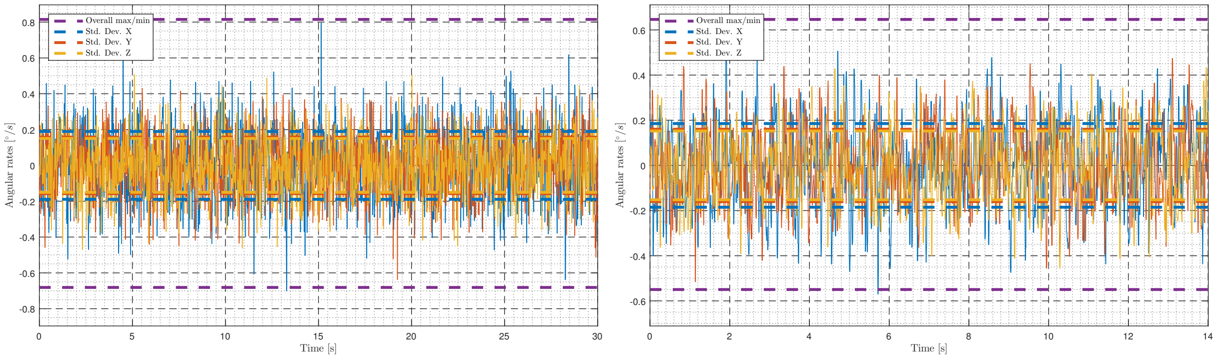


Figure 6.10: The noise in angular velocity measurements for the experiments in Section 3.5 (left) and in Section 4.6 (right).

drawings of the area. For that purpose, consider Figure 6.11. There is depicted the estimated map, in black, along with extended ground truth in red and the dead-reckoned path in blue with the executive drawings in the background. From this comparison, it can be seen that the resulting map clearly fits the executive drawings and shows the convergence of the landmark

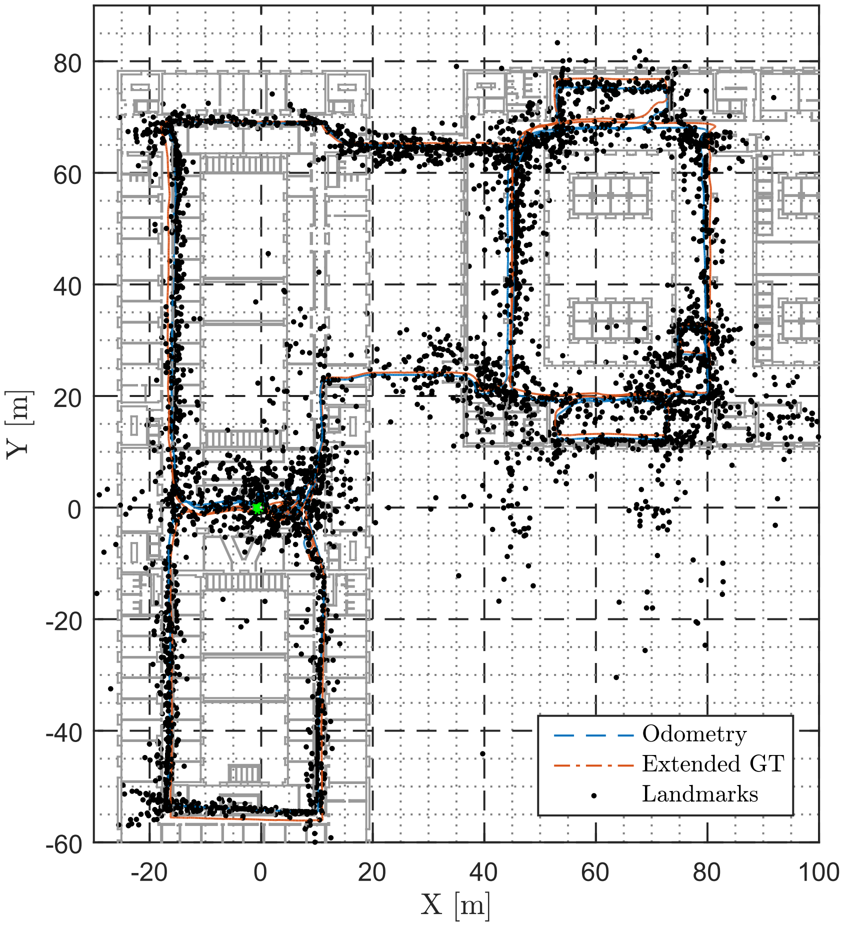


Figure 6.11: The estimated map at the end of the experiment, along with the dead-reckoned path and extended ground-truth.

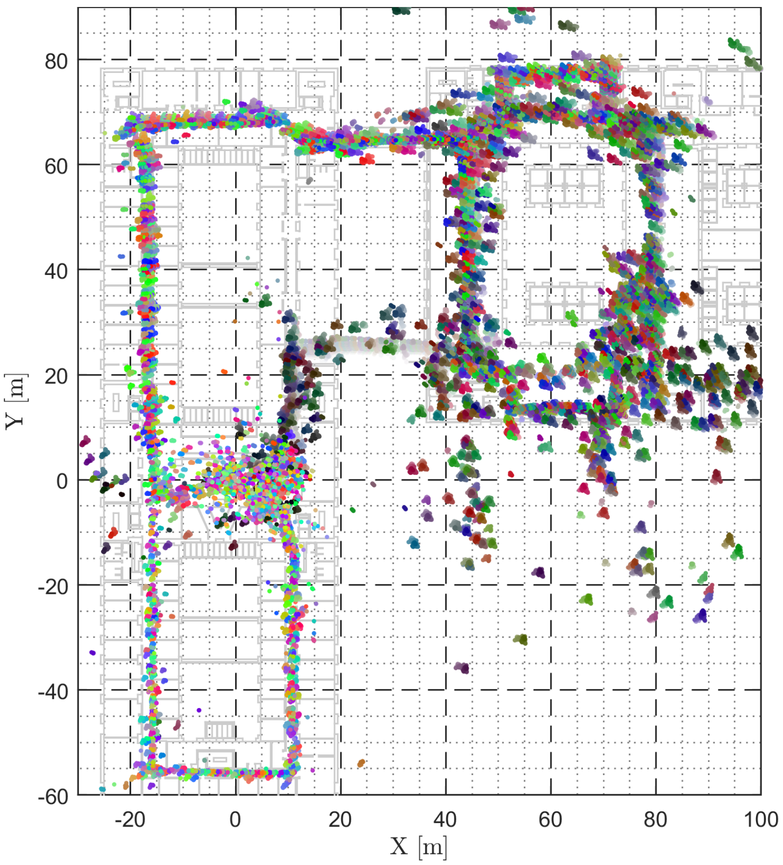
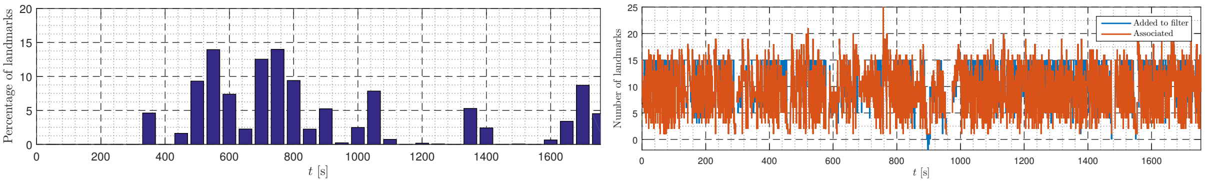


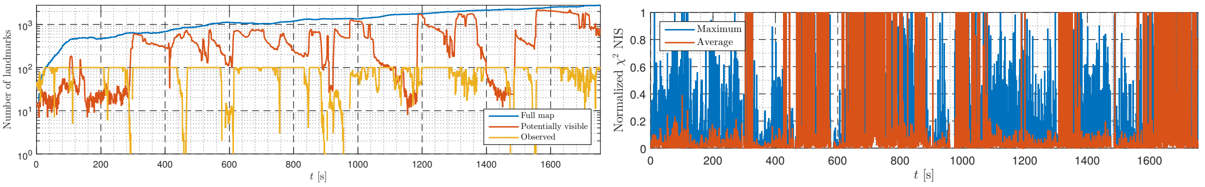
Figure 6.12: The estimated map throughout the whole of the experiment, with the spatial evolution of each landmark shown in intervals of 20 seconds.

state which was, as explained throughout the chapter, initialized without any special care. Even though some landmarks are too far away for not having had enough time to converge, either because they were only observed too few frames or because the association process did not recognize them, they help demonstrate the convergence, as they contrast with the other landmarks that fulfilled all the requirements for observability. It is also interesting to remark that both corridors that connect the two buildings have more dispersed landmarks that the rest of the building. This may be due to the fact that both corridors are very dark and that mostly only distant lights from the neighbouring buildings are detected as landmarks (see Figure 6.8(g)). Furthermore, the bottom corridor that connects the two buildings has fewer (and more dispersed) landmarks than the top one, which can be explained by the fact that the robot circulates around the bottom left part of the library facing the corridor twice, and falsely loop-closed landmarks from the area of the corridor are re-estimated as being part of the building, thus moving away from the corridor. Finally, the spatial evolution of a selected set of landmarks rotated to the Earth-fixed frame using the ground truth is shown in Figure 6.12. This enables to better visualize the convergence of the filter and effects of repeated sightings. Each small blob represents the sampled history (every 20 seconds) of a landmark, from very light colours (old) to stronger (more recent). With this figure it is possible to evaluate the dispersion



(a) Percentage of loop-closed landmarks.

(b) Number of used observations.



(c) Number of landmarks.

(d) Normalized innovation square.

Figure 6.13: Algorithm performance statistics.

of the estimates through time and space, and assess zones where convergence was slower or erroneous associations were made. For instance, the corridor at around $Y = 25$ m shows a blur of grey that is due to the landmarks therein slowly converging to their final positions.

In a sensor-based approach, all the processes exogenous to the filter, that build upon its results or upon which the filter relies, can avoid the inherent nonlinearity that stems from the transformation of the naturally body-fixed measurements to an Earth-fixed frame. For example, by performing landmark association, loop-closing, control, and decision procedures in the sensor-based frame, it is possible to reduce the effects of nonlinearities in the consistency of the filter. Consistency is a very important aspect of any estimation algorithm, and simultaneous localization and mapping is not an exception. This can be evaluated by checking how the uncertainty indicated by the algorithm compares to the actual estimation error. When this approach cannot be pursued, there are other ways to evaluate it. One of the best measures of the consistency and validity of a SLAM algorithm is the way it handles loop closures, i.e., if it does recognize previously visited places. The algorithm proposed in this filter does not include a specially tailored loop closing procedure, even though one of the many existent in the literature that avail themselves of other techniques such as place recognition from images can be applied on top of the proposed BO-SLAM filter. In this case, loop closures will occur naturally from association of new observations with old landmarks that fit the association criteria. Even though association/loop closing is based on the image descriptors, the search for associations is limited to the field-of-view predicted by the BO-SLAM filter as explained in Section 6.4.2. Furthermore, the decision to associate or add a new landmark depends on the Mahalanobis distance of the measurement. Hence, the loop closings depend both on image and filter information. Loop closures can be clearly observed when there is a sharp reduction of the uncertainty of a landmark after some time increasing. For the purpose of providing a better visualization of this process and its results, the landmarks re-observed after a long time, henceforth denoted as loop-closed landmarks, were bundled in slots of 50 seconds as shown in

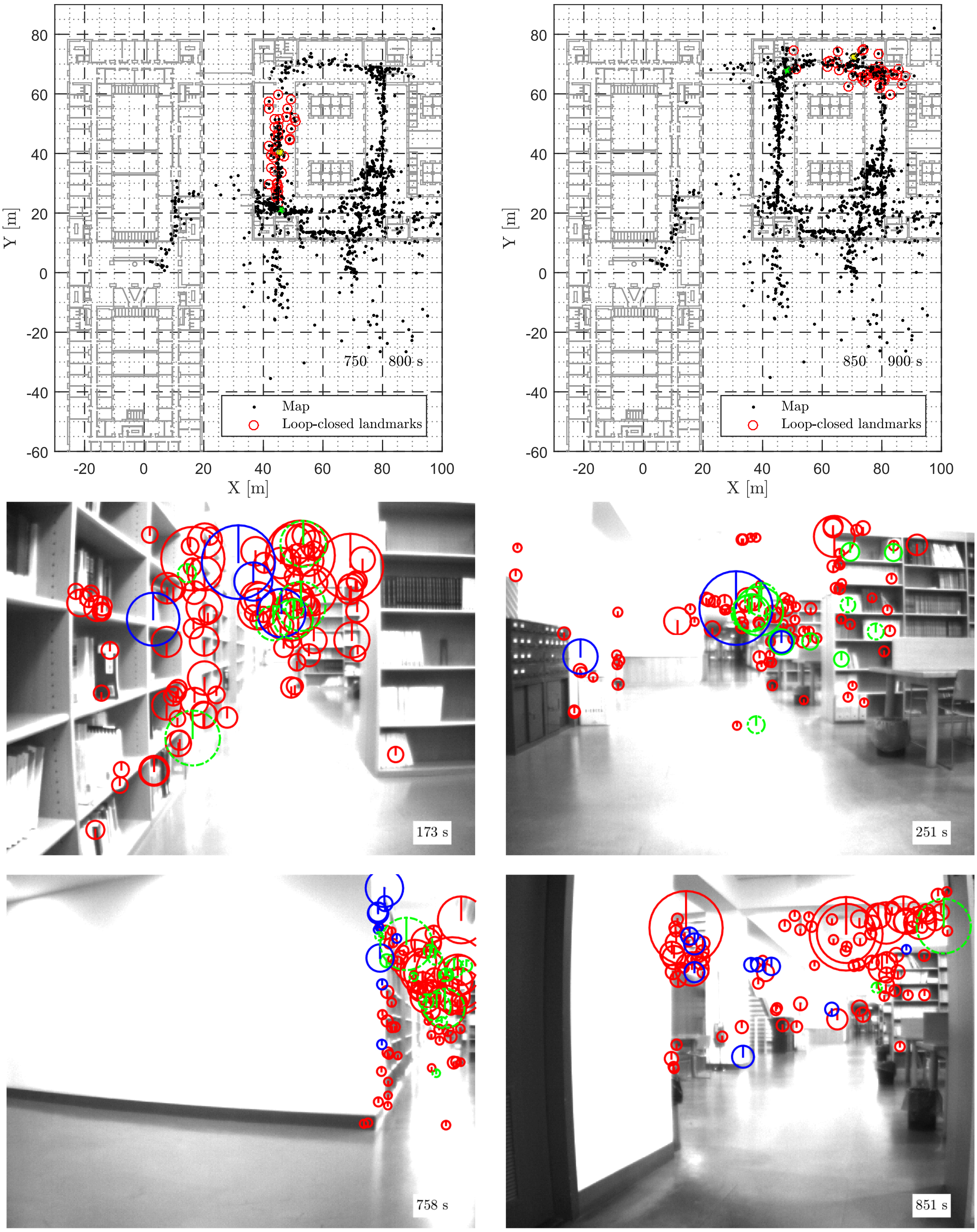


Figure 6.14: At the top: loop-closed landmarks in the first loop closing peaks of Figure 6.13(a), with loop-closed landmarks highlighted in red. The middle and bottom pictures are frames of the camera stream with detected (red), selected (green) and associated (blue) features. The bottom one contains the most loop-closed landmarks of the corresponding time slot and the top one shows the one where more loop-closed landmarks where first observed.

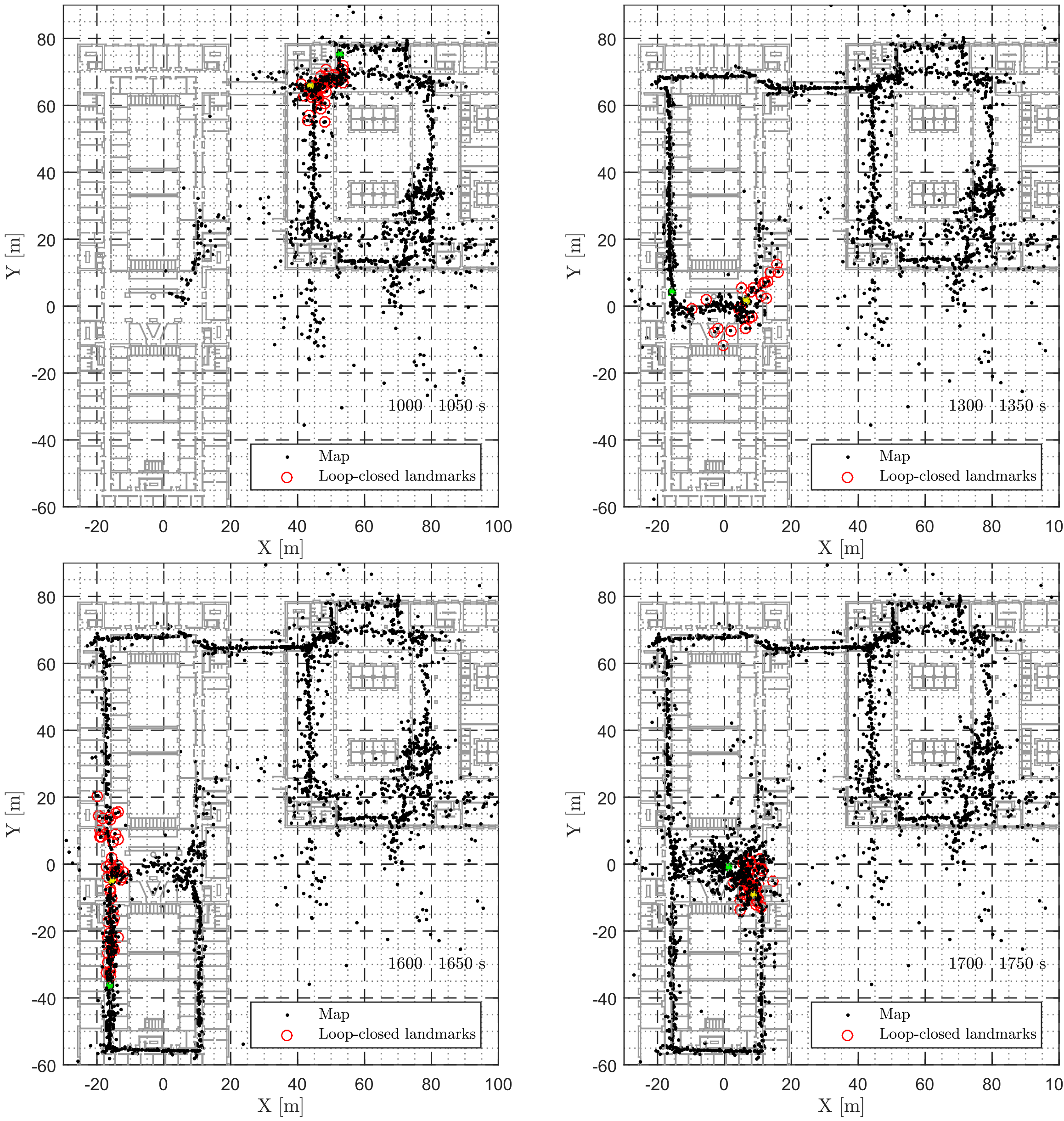


Figure 6.15: Maps corresponding to loop closing peaks in Figure 6.13(a), with loop-closed landmarks highlighted in red.

Figure 6.13(a). The number of loop-closed landmarks is very large around the 600-800 seconds mark, which corresponds to the time when the robot revisits the bottom part of the building in the right side of the map. Then there are peaks around the 1000 – the robot revisits the top left corner of the right-side building, 1400 – the robot revisits the starting place, and 1600 seconds marks – the robot revisits the leftmost corridor of the map. Details about all these can be seen in Figures 6.14 and 6.15. The maps in both figures depict the maps at the end of each of the chosen time slots, along with the loop-closed landmarks highlighted in red. The pose of the robot at the start of the time slot is represented by the green solid circle, and the yellow circle stands for the pose of the robot at the end of the time slot. In Figure 6.14 frames of the camera stream used in the filter are also presented. The bottom ones are the frames which contain the most loop-closed landmarks of the corresponding time slot, and the top ones are the frames where more of the loop-closed landmarks were observed the first time. The red circles represent detected features, the green circles mark the features selected for the filter, and the blue circles depict the features that were associated with existing landmarks. The evenly dotted circles represent loop-closed landmarks in the actual frame, and the other dotted circles show landmarks that closed a loop sometime in the run, not exactly in this frame. The vertical line in each feature indicates the sign of the Laplacian. It can be seen that, even though the loop closing is not based directly on visual place recognition, the algorithm is able to discern whether a block of features in different frames come in fact from the same place, thus demonstrating the good performance of the algorithm, as well as its consistency.

Another measure of the consistency of a SLAM algorithm is the normalized innovation square (NIS) value for each set of associations, given by the joint Mahalanobis distance of the observation as explained in Section 3.4.1. This value is employed as an association index, and is known to follow a chi-squared distribution with 3 degrees of freedom for 3-D measurements. It is compared to the 95% percentile threshold to determine if an association is valid or not. In Figure 6.13(d), the average and maximum normalized innovation square (NIS) values are also normalized with the value corresponding to the 95% percentile. It is observed that although the maximum values approach the threshold, the average association is well below it, indicating that the landmark associations and the overall sensor-based algorithm are consistent. A clear exception to this behaviour is observed in the situations where more loop closings occur, which makes sense as the uncertainty of older landmarks is larger, and the influence of the drift of the odometry sensors is also more felt than when associating with newer landmarks.

The observability results of Section 6.3 provide insight that should be taken into account when designing experiments. Even though that was not the case here, the observability conditions seem to be fulfilled for the most part of the landmarks, as they converge to recognizable positions. For that reason, these experimental results served the purpose of demonstrating the capabilities of the sensor-based BO-SLAM filter herein proposed both in terms of performance and consistency.

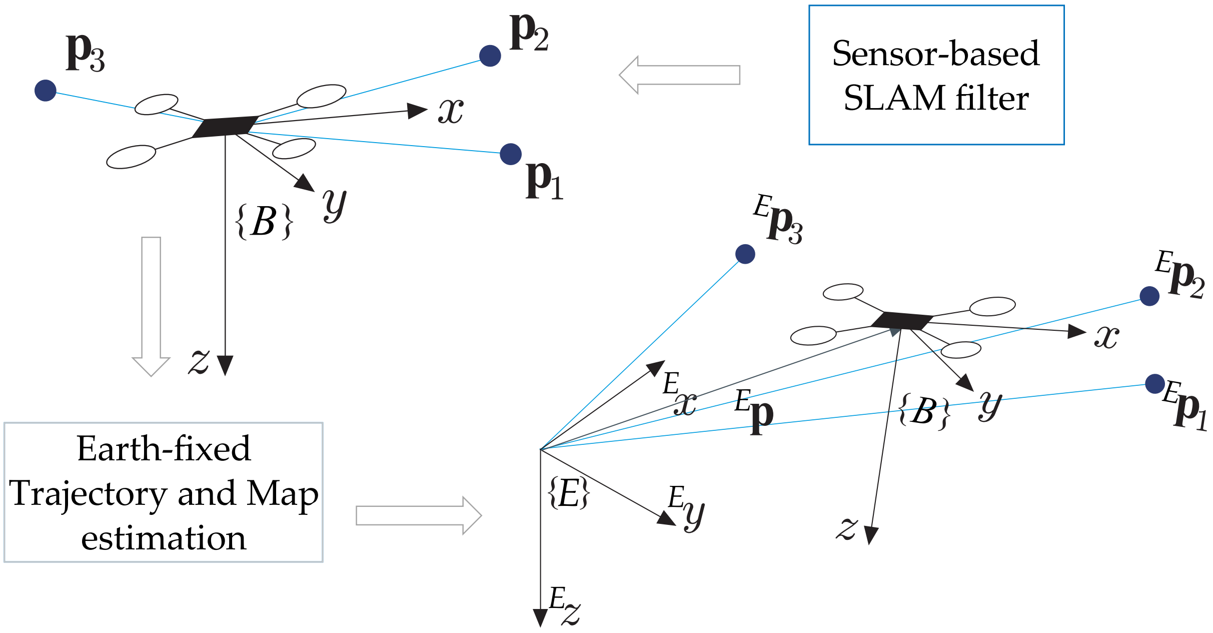
6.7 Conclusions

A novel sensor-based globally exponentially stable filter for bearing-only simultaneous localization and mapping was proposed in this chapter. The design of a linear time-varying system that mimics the dynamics of the underlying nonlinear system was facilitated by the use of a state augmentation and a simple output transformation on a nonlinear system, while disposing of non-visible landmarks. A thorough and constructive observability analysis was performed, leading to the establishment of physically-grounded necessary and sufficient conditions for observability, stability and convergence of the Kalman filter that followed, as well as conditions for the observability of the underlying nonlinear system. These conditions are interesting for trajectory design or motion planning. In opposition to the requirements for observability of the range-only problems, bearing-only observability conditions do not necessarily require 3-D motion to estimate the position of three-dimensional landmarks. In fact, a vehicle can move in only one direction and maintain observability of the whole map, as long as that direction does not coincide with the direction of one of the observed landmarks.

The sensor-based BO-SLAM filter was prepared for practical implementation with a camera as the only exteroceptor, and thus the theoretical work which is the main focus of this chapter was validated through experimental results based on a widely available dataset, that allowed to expose the good performance of the algorithm, as well as its consistency in an indoor environment with a large loop. These experimental results also helped to confirm the convergence of the error dynamics of the Kalman filter for LTV systems, as predicted by the theoretical results herein presented.

Part II

EARTH-FIXED TRAJECTORY AND MAP


$\mathbf{p}_3$
$\mathbf{p}_2$
$\mathbf{p}_1$
x
y
z
$\{B\}$
Sensor-based SLAM filter
Earth-fixed Trajectory and Map estimation
${}^E\mathbf{p}_3$
${}^E\mathbf{p}_2$
${}^E\mathbf{p}_1$
${}^E\mathbf{p}$
${}^Ex$
${}^Ey$
${}^Ez$
$\{E\}$

7

UNCERTAINTY CHARACTERIZATION OF THE ORTHOGONAL PROCRUSTES PROBLEM

THIS chapter establishes the ground work for the Earth-fixed Trajectory and Map estimation algorithm that complements the sensor-based SLAM filters of the previous chapters. As explained in Section 2.4, the algorithm revolves around the Procrustes problem (see Definition 2.5), and, as such, this chapter addresses the weighted orthogonal Procrustes problem of matching stochastically perturbed point clouds, formulated as an optimization problem with a closed-form solution. A novel uncertainty characterization of the solution of this problem is proposed resorting to perturbation theory concepts, which admits arbitrary transformations between point clouds and individual covariance and cross-covariance matrices for the points of each cloud. The method is thoroughly validated through extensive Monte Carlo simulations, and particularly interesting cases where nonlinearities may arise are further analysed. The developments of this chapter lay the foundations for estimating the Earth-fixed pose from sensor-based maps, which is detailed and experimentally validated in Chapter 8.

7.1 Introduction

The problem of finding the similarity transformation between two sets of points in n-dimensional spaces appears commonly in many applications of computer vision, robotics, statistics, and other fields of research. The study of this family of problems is usually known as the Procrustes analysis [GD04], which includes the statistical characterization of the transformation between the shape of objects [Goo91]. One particularly important problem in this family is the so-called orthogonal Procrustes problem, which can be traced back to the work presented in [Sch66], and consists in extracting the orthogonal transformation that maps one set of points into a second set of points, with known associations between them. It is closely related to Wahba's problem [Wah65] and to the Kabsch algorithm [Kab76]. The generalization for rotation, translation, and scaling has been the subject of extensive research in areas such as computer vision, and can be traced back to [Ume91], [HHN88], and [AHB87]. While initially the problem was posed without any restrictions on the transformation between the sets, i.e., rotations and reflections were allowed, a more evolved strategy appeared restricting the transformation to the special orthogonal group, as detailed in [Ume91] and [Kan94]. Furthermore, [GH95] demonstrated that the previous solutions are optimal even when both data sets are perturbed with isotropic and identical Gaussian noise.

The statistical characterization of the Procrustes analysis has also been the subject of study in works such as [Goo91], [Kan94], [Sib79], and [Sib78]. Using perturbation theory, the nonlinear problem of characterizing the uncertainty was addressed with some limiting options, such as the absence of weighting of the point sets, the use of small rotations, or the same covariance for all points. Within the field of medical imaging, the work presented in [FW01] also resorts to perturbation theory to present a statistical characterization of a target position, considering small rotations, isotropic uncertainty, and equal weights for each point. More recently, the work presented in [WLFP08] extends these results for anisotropic uncertainty in the components of the point space. This is achieved by considering the same covariance matrix for all points, which may weigh each component of the point space independently. The authors of [MA09] further expand this by considering different noise levels for each point, while keeping the linearized model for the rotation matrix. An interesting advance in the study of the uncertainty is the first order error propagation proposed in [Dor05]. The optimization problem that is considered is not weighted and therefore identical isotropic noise is assumed for all the points. The author defines a first order error model that is propagated through the solution, while assuming independent and identically distributed points (no longer necessarily isotropic). It is noted that the findings of the aforementioned works are all restricted to three-dimensional points. In [OK98] a different optimization problem is proposed that accounts for independent anisotropic noise affecting rotated-only point sets also in three dimensions. The authors determine the theoretical lower bound for the covariance of the rotation error in that case, and through an iterative solution recurring to quaternion representation reach the theoretical bound. Besides the iterative solution, some shortcomings of this work are its limitation to the tridimensional problem with rotation-only, and the fact that, although anisotropic, the input covariances are normalized and share a common normalizing factor. Regarding the stability of the solution, [S9̈3] addresses the study of this issue when the algorithm is exposed to perturbed data sets, concluding that the singular values of the matrices composed with the points in each set are closely related to the conditioning of the problem, whilst finding a bound for the perturbation on the rotation matrix when the input perturbations are bounded. In related directions of research that demonstrate the relevance of pattern point matching, and, consequently, of point registration problems such as the Procrustes problem, the authors of [TSZ14] and [CEBW14] propose algorithms that exploit different approaches to registration and matching. Furthermore, the latter is an iterative algorithm that, assuming a rotation-scaling-translation transformation between two sets of points, finds the point correspondences and a variational Bayesian approximation for the distribution of the transformation.

This chapter addresses the n-dimensional (n-D) extended orthogonal Procrustes problem considering a transformation composed of a rotation and a translation (no scaling). The problem is posed with individual scalar weights for each pair of points, and a closed-form solution is presented. Data association is assumed to be performed a priori. Founded on perturbation theory, a novel and general uncertainty description for the solution of the optimization problem is proposed. Building on the results presented in [FW01], [WLFP08], and [Dor05], and assuming a stochastic perturbation model for the point sets with individual covariance matrices for each point, as well as cross-covariances for each pair of points, the first and second moments of the resulting translation and rotation are computed. This is achieved considering arbitrary rotations and translations, individual weights, and full covariance matrices for both point sets.

The contributions of this chapter are: i) the full uncertainty characterization of the optimization problem of obtaining the transformation between corresponding n-dimensional point sets, while considering point sets perturbed by anisotropic noise, and points that are not required to be independent nor identically distributed; and ii) a thorough validation of the uncertainty characterization, using extensive Monte Carlo simulations to study the main properties of the proposed methodology.

The applications of Procrustes analysis are found in a wide variety of fields, which can benefit from the proposed approach, including rigid body motion, vibration tests of large complex structures [BS92], structural and system identification, factor analysis in n-D (e.g. checking whether two matrices are equivalent), similarity evaluation in statistical data sets [BG07, Chapter 20], medical imaging [WLFP08], photogrammetry [CB02], shape comparison (generalized Procrustes analysis) [IPSE$^+$14], and quantitative psychology [SC70] (where the problem was initially solved). In recent years several algorithms were developed in the field of computer vision that availed themselves of the Procrustes problem, from shape matching and retrieval [MV06] to similarity search in image collections [GLGP13], among others. Shape matching is in fact a more complex problem, as the problem of finding the transformations is coupled with the problem of finding the reference shape to which all the measured shapes relate. In [BPL12] the authors propose a unifying framework that has a closed-form computation for affine, similarity or Euclidean transformations between a set of shapes, while allowing to find the underlying shape and accounting for missing pairs of points. All this is performed considering noise in the measured shapes and not in the reference-space as is customary. Other applications include non-negative matrix factorization [HSS14], and phase FIR filter bank design [LTH$^+$15], whereas the work in [LKPY11] underlines the importance of addressing the problem in less common dimensionalities, such as four-dimensional shapes. Another possible application of the Procrustes problem lies in iterative closest point algorithms such as [KKM03], even tough most use quaternions to parametrize the rotation of the sets. If the registration is performed in each step with a constrained least squares approach, it can benefit from the characterization here proposed. Another interesting application of Procrustes analysis is manifold alignment [WM08] in the area of machine learning. In this n-dimensional technique, it is argued that it is possible to model the underlying structure of most datasets by manifolds, whose alignment then allows for knowledge transfer across datasets. The authors of [WM08] demonstrate the validity of this approach by applying the idea to learning transfer in reinforcement learning with Markov Decision Processes, alignment of the tertiary structure of proteins, cross lingual information alignment, and so forth. Given the noisy nature of these problems, the proposed uncertainty characterization can be useful to compute the reliability of the alignment resulting from the Procrustes procedure. Furthermore, these demonstrate the real world relevance of the n-dimensional Procrustes problem in several fields even for dimensionalities outside of the 2-D/3-D common problems. For these reasons, the current chapter is presented with a generic notation that can easily be applied to any of these situations, without explicit reference to the SLAM concepts of the previous chapters. The results of this chapter are contextualized in Chapter 8.

Chapter Structure

The chapter is organized as follows. Section 7.2 presents a brief overview of some mathematical concepts needed in the course of this chapter. Section 7.3 presents the formulation and closed-form solution of the weighted orthogonal Procrustes problem. A novel uncertainty characterization of this problem is derived in Section 7.4 and validated in Section 7.5 through extensive Monte Carlo simulations. Finally, Section 7.6 provides concluding remarks.

7.2 Preliminary definitions

This section serves the purpose of introducing a few definitions and properties needed for the mathematical derivations in the sequel.

In this chapter, except when explicitly stated, the dimension of the space $\mathbb{R}^n$ is arbitrary, i.e., all the derivations are valid for point clouds in $\mathbb{R}^n$ for all $n \geq 2$, and the rotations are expressed in the special orthogonal group $SO(n)$. Note that the term rotation, and rotation matrix, applies to all the orthogonal matrices of unitary determinant. This is related to the Lie algebra $\mathfrak{so}(n)$ comprised of skew-symmetric matrices that can be mapped to $SO(n)$ through the exponential map (see Definition 2.1 for the tridimensional case).

Definition 7.1 (The n-dimensional skew-symmetric matrix)**.** *The matrix* $\mathbf{S}(\boldsymbol{\varpi}) \in \mathfrak{so}(n)$ *is a skew-symmetric matrix parametrized by the vector* $\boldsymbol{\varpi} := \begin{bmatrix} \varpi_1 & \cdots & \varpi_{n_p} \end{bmatrix}^T \in \mathbb{R}^{n_p}$*, with* $n_p = \frac{n(n-1)}{2}$*. One* *possible parametrization is given by*

$$\mathbf{S}(\boldsymbol{\varpi}) = \begin{bmatrix} 0 & (-1)\varpi_{n_p} & (-1)^2\varpi_{n_p-1} & \cdots & (-1)^{n-2}\varpi_{2n-3} & (-1)^{n-1}\varpi_{n-1} \\ * & 0 & (-1)\varpi_{n_p-2} & \cdots & (-1)^{n-3}\varpi_{2n-4} & (-1)^{n-2}\varpi_{n-2} \\ * & * & 0 & & \vdots & \vdots \\ * & * & * & \ddots & (-1)\varpi_n & (-1)^2\varpi_2 \\ * & * & * & * & 0 & (-1)\varpi_1 \\ * & * & * & * & * & 0 \end{bmatrix}. \tag{7.1}$$

The elements under the diagonal are automatically defined by $\mathbf{S}(\boldsymbol{\varpi}) = -\mathbf{S}^T(\boldsymbol{\varpi})$, *and are therefore omitted to avoid cluttering the reading. Also note that, in general,* $n_p \geq n$, *except in the bidimensional case when* $n_p = 1$ *and the whole vector* $\boldsymbol{\varpi}$ *collapses to* ϖ_1. *The unskew operator* $\mathbf{S}^{-1} : \mathfrak{so}(n) \to \mathbb{R}^{n_p}$ *is related to this matrix as it extracts the vector that parametrizes a skew-symmetric matrix, i.e.,* $\mathbf{S}^{-1}(\mathbf{S}(\boldsymbol{\varpi})) = \boldsymbol{\varpi}$.

Notice that, given $\boldsymbol{\varpi} \in \mathbb{R}^{n_p}$, other parametrizations could be used to obtain $\mathbf{S}(\boldsymbol{\varpi})$. This form was chosen without loss of generality, as the results could be derived for any parametrization. As used throughout this dissertation, for the tridimensional case ($n = 3$), this matrix is also known to encode the cross-product, as in $\mathbf{S}(\mathbf{a})\mathbf{b} = \mathbf{a} \times \mathbf{b}$. Due to this fact, there is a possible anti-commutation between the two vectors given by $\mathbf{S}(\mathbf{a})\mathbf{b} = -\mathbf{S}(\mathbf{b})\mathbf{a} = \mathbf{S}^T(\mathbf{b})\mathbf{a}$. Even though these properties are solely applicable to the tridimensional case, a possible extension to all the dimensions is the matrix $\bar{\mathbf{S}}(.)$ defined hereafter.

Definition 7.2 (The n-dimensional generalized anti-commutation matrix)**.** *The generalized anti-commutation matrix* $\bar{\mathbf{S}}(\mathbf{a}) \in \mathbb{R}^{n_p \times n}$ *parametrized by the vector* $\mathbf{a} \in \mathbb{R}^n$, *when associated to the parametrization of* $\mathbf{S}(\boldsymbol{\varpi})$ *in* (7.1), *can be defined recursively by*

$$\bar{\mathbf{S}}(\mathbf{a}_{1:i}) = \left[\begin{array}{cccc|c} 0 & \cdots & 0 & -a_i & (-1)^0 a_{i-1} \\ \vdots & & (-1)^2 a_i & 0 & (-1)^1 a_{i-2} \\ 0 & \iddots & & \vdots & \vdots \\ (-1)^{i-1} a_i & 0 & \cdots & 0 & (-1)^{i-2} a_1 \\ \hline \multicolumn{4}{c|}{\bar{\mathbf{S}}(\mathbf{a}_{1:i-1})} & \mathbf{0}_{\frac{(i-1)(i-2)}{2} \times 1} \end{array}\right] \tag{7.2}$$

for all $i = 3, \ldots, n$, *with* $\mathbf{a}_{1:i} = \begin{bmatrix} a_1 & \cdots & a_i \end{bmatrix}^T$ *and* $\bar{\mathbf{S}}(\mathbf{a}_{1:2}) = \begin{bmatrix} -a_2 & a_1 \end{bmatrix}$.

Note that for vectors in three dimensions, this matrix degenerates in the skew-symmetric matrix, i.e., $\bar{\mathbf{S}}(\mathbf{a}_{1:3}) = \mathbf{S}(\mathbf{a}_{1:3})$. Besides this fact, the matrix (7.2) exhibits a series of properties that are used in this dissertation. These are summarized in the following lemma.

Lemma 7.1 (Properties of the generalized anti-commutation matrix)**.** *The generalized anti-commutation matrix has the following properties for every* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\boldsymbol{\varpi} \in \mathbb{R}^{n_p}$, *constants* $\alpha, \beta \in \mathbb{R}$ *and integer* $n \geq 2$.

1. *Linearity:* $\bar{\mathbf{S}}(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha\bar{\mathbf{S}}(\mathbf{a}) + \beta\bar{\mathbf{S}}(\mathbf{b})$

2. *Anti-commutativity:* $\bar{\mathbf{S}}(\mathbf{a})\,\mathbf{b} = -\bar{\mathbf{S}}(\mathbf{b})\,\mathbf{a}$
3. *Anti-commutation with a skew-symmetric matrix:* $\mathbf{S}(\boldsymbol{\varpi})\,\mathbf{a} = \bar{\mathbf{S}}^T(\mathbf{a})\,\boldsymbol{\varpi}$
4. *Relation to the unskew operator:* $\mathbf{S}^{-1}\left(\mathbf{b}\mathbf{a}^T - \mathbf{a}\mathbf{b}^T\right) = \bar{\mathbf{S}}(\mathbf{a})\,\mathbf{b}$.

Proof. Due to the recursive nature of $\bar{\mathbf{S}}(\mathbf{a})$ and $\mathbf{S}(\boldsymbol{\varpi})$, it is possible to express all the sides of these identities recursively and thus these properties can be proven using mathematical induction. The proof is presented in Appendix B.2. ☐

7.3 Procrustes optimization problem and closed-form solution

This section presents the formulation and closed-form solution of the weighted extended orthogonal Procrustes problem, featuring individual weights for each pair of points, related by a translation and rotation.

Consider the existence of two point sets in $\mathbb{R}^n$, $\mathcal{S}_A$ and $\mathcal{S}_B$, which contain, respectively, the points expressed in an arbitrary frame $\{A\}$ and the same points expressed in some other frame $\{B\}$. Each point $\mathbf{a}_i \in \mathcal{S}_A$ nominally corresponds to a point $\mathbf{b}_i \in \mathcal{S}_B$, with $i \in \mathcal{S} := \{1,\ldots,m\}$ and $m \geq n$, and that correspondence is expressed by $\mathbf{a}_i = \mathbf{R}\mathbf{b}_i + \mathbf{t}$, where the pair $(\mathbf{R},\mathbf{t}) \in \mathrm{SO}(n) \times \mathbb{R}^n$ fully defines the transformation from frame $\{B\}$ to frame $\{A\}$, as it represents the rotation and translation from $\{B\}$ to $\{A\}$. Given the relation between the two sets, it is possible to define the error function $\mathbf{e}_i = \mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}$, that represents the error between the i-th point estimate in $\mathcal{S}_A$ and its homologous in $\mathcal{S}_B$, rotated and translated into frame $\{A\}$. Obtaining the pair $(\mathbf{R},\mathbf{t})$ is the purpose of the optimization problem

$$(\mathbf{R}^*,\mathbf{t}^*) = \arg \min_{\substack{\mathbf{R}\in \mathrm{SO}(n)\\ \mathbf{t}\in\mathbb{R}^n}} G(\mathbf{R},\mathbf{t}), \tag{7.3}$$

where the function $G(\mathbf{R},\mathbf{t})$ is defined as

$$G(\mathbf{R},\mathbf{t}) := \frac{1}{m}\sum_{i=1}^{m} \sigma_i^{-2}\,\|\mathbf{a}_i - \mathbf{R}\mathbf{b}_i - \mathbf{t}\|^2 = \frac{1}{m}\left\|\left(\mathbf{Y} - \mathbf{R}\mathbf{X} - \mathbf{t}\mathbf{1}^T\right)\Sigma_e^{-1/2}\right\|^2,$$

where $\sigma_i^2 > 0$, $i \in \mathcal{S}$, accounts for the intrinsic uncertainty of each point pair, $\mathbf{Y} = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_m\end{bmatrix} \in \mathbb{R}^{n\times m}$ and $\mathbf{X} = \begin{bmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_m\end{bmatrix} \in \mathbb{R}^{n\times m}$ are, respectively, the concatenation of the point vectors expressed in frames $\{A\}$ and $\{B\}$, $\mathbf{1} = \begin{bmatrix}1 & \cdots & 1\end{bmatrix}^T \in \mathbb{R}^m$ is a vector of ones, and the weight matrix Σ_e is a diagonal matrix whose entries are the weights $\sigma_1^2,\ldots,\sigma_m^2$ that model the point uncertainty. If, for example, one of the points in a pair is missing, the corresponding σ_i^{-2} can be set to zero, thus allowing the cost function to account for missing data. Given that the true Σ_e is not known, these can be conservatively defined as $\sigma_i^2 = \lambda_{max}\left(\Sigma_{a_i}\right) + \lambda_{max}\left(\Sigma_{b_i}\right) \geq \lambda_{max}\left(\Sigma_{a_i} + \mathbf{R}\Sigma_{b_i}\mathbf{R}^T\right)$, denoting $\lambda_{max}(.)$ as the maximum eigenvalue. This weight matrix allows the use of the information regarding the different degrees of uncertainty of each point pair.

The optimization problem (7.3) has a closed-form, numerically robust, and computationally efficient solution based on the work presented in [Ume91] and [Goo91]. The weighted statistical properties of the point sets $\mathcal{S}_A$ and $\mathcal{S}_B$, in the form of their weighted centroids and covariances, can be expressed in matrix form using the symmetric weight matrix $\mathbf{W} := \Sigma_e^{-1} - \frac{1}{N_W}\Sigma_e^{-1}\mathbf{1}\mathbf{1}^T\Sigma_e^{-1} \in \mathbb{R}^{m\times m}$, where $N_W := \sum_{i=1}^{m} \sigma_i^{-2} = \mathbf{1}^T\Sigma_e^{-1}\mathbf{1}$. The resulting expressions for the weighted centroids of the sets $\mathcal{S}_A$ and $\mathcal{S}_B$, respectively $\boldsymbol{\mu}_A, \boldsymbol{\mu}_B \in \mathbb{R}^n$, are

$$\boldsymbol{\mu}_A := \frac{1}{N_W}\sum_{i=1}^{m}\sigma_i^{-2}\mathbf{a}_i = \frac{1}{N_W}\mathbf{Y}\Sigma_e^{-1}\mathbf{1}, \quad \text{and} \quad \boldsymbol{\mu}_B := \frac{1}{N_W}\mathbf{X}\Sigma_e^{-1}\mathbf{1},$$

whereas the weighted covariance $\Sigma_{AB} \in \mathbb{R}^{n\times n}$ is given by

$$\Sigma_{AB} := \frac{1}{N_W}\sum_{i=1}^{m}\sigma_i^{-2}\left(\mathbf{a}_i - \boldsymbol{\mu}_A\right)\left(\mathbf{b}_i - \boldsymbol{\mu}_B\right)^T = \frac{1}{N_W}\mathbf{Y}\mathbf{W}\mathbf{X}^T := \frac{1}{N_W}\mathbf{B}^T.$$

Consider now the singular value decomposition of $\mathbf{B}^T$

$$\mathbf{U}\mathbf{D}\mathbf{V}^T = \operatorname{svd}(\mathbf{B}^T). \tag{7.4}$$

The optimal rotation matrix, from the optimization problem (7.3), is given by

$$\mathbf{R}^* = \mathbf{U}\operatorname{diag}(1,\ldots,1,|\mathbf{U}||\mathbf{V}|)\,\mathbf{V}^T, \tag{7.5}$$

which in fact is the optimal rotation between the two sets when their weighted centroids coincide. This solution is valid as long as the covariance matrix Σ_{AB} has rank $n-1$, as reported in [Ume91]. The optimal translation vector is

$$\mathbf{t}^* = \frac{1}{N_W}\left(\mathbf{Y} - \mathbf{R}^*\mathbf{X}\right)\Sigma_e^{-1}\mathbf{1} = \boldsymbol{\mu}_A - \mathbf{R}^*\boldsymbol{\mu}_B. \tag{7.6}$$

Notice that the optimal translation is the vector that translates the weighted centroid of the points in $\mathcal{S}_B$ rotated to frame $\{A\}$ to the weighted centroid of the points in $\mathcal{S}_A$.

Some remarks on applications The derivation here presented and followed in the next section is detailed for the matching of two point clouds. Nevertheless, the problem is completely equivalent to the general problem of matching two different matrices, considering that each point is a column of the corresponding matrix. When dealing with generalized Procrustes analysis, for shape matching/registration, the problem at hand is extended to include the estimation of a reference shape. This is not the objective of the optimization problem (7.3). However, once a reference shape is chosen or found (for example using the the classical alternation approach [GD04, Chapter 9] or the more evolved approaches proposed in [BPL12] and [IPSE$^+$14]), the problem of finding the Euclidean transformation is exactly expressed by (7.3) and therefore all the derivations in this and the following sections hold.

7.4 Uncertainty characterization

The goal of this section is to compute accurate approximations of the uncertainty associated with the extended Procrustes problem solution obtained above. The study of the statistical properties of the orthogonal Procrustes problem can be traced back to [Sib79] and [Goo91], which used perturbation theory to find the approximate distribution of the cost functional $G(\mathbf{R},\mathbf{t})$, assuming $\Sigma_e^{-1} = \mathbf{I}$, small rotations, and the same covariance matrix for each landmark. These results have been extensively applied to the field of point-based medical image registration problems, as detailed in [FW01], [WLFP08], and [MA09]. The work in [Dor05] eliminates the assumptions on the rotation, but still addresses the optimization problem without weights and considers that the points are independent and identically distributed, albeit with anisotropic uncertainty. The analysis presented hereafter is based on the aforementioned works and aims at providing approximate uncertainty descriptions for the transformation parameters, $\mathbf{R}^*$ and $\mathbf{t}^*$. This uncertainty characterization is achieved while considering arbitrary rotations and translations and using individual weights as well as individual covariance matrices for each landmark and cross covariance terms.

The solution of the optimization problem now defined involves uncertainty at the input and output level: i) the point sets used to estimate the transformation between frames are not exact, having a perturbation part; and, thus, ii) the resulting estimated translation and rotation are also perturbed. Within the scope of perturbation theory, the error models of the known variables are defined as

$$\mathbf{a}_i = \mathbf{a}_i^{(0)} + \epsilon\, \mathbf{a}_i^{(1)} + \mathcal{O}(\epsilon^2) \tag{7.7a}$$

$$\mathbf{b}_i = \mathbf{b}_i^{(0)} + \epsilon\, \mathbf{b}_i^{(1)} + \mathcal{O}(\epsilon^2) \tag{7.7b}$$

for all $i \in \mathcal{S}$, where ϵ is the smallness parameter, the notation $\mathcal{O}(\epsilon^m)$ stands for the remaining terms of order m or higher, the zero order terms are the true values, hence $\langle \mathbf{a}_i^{(0)} \rangle = \mathbf{a}_i^{(0)}$ and $\langle \mathbf{b}_i^{(0)} \rangle = \mathbf{b}_i^{(0)}$, whereas the first order terms, $(.)^{(1)}$, are assumed to follow a known distribution with zero mean and covariance matrices defined by $\Sigma_{a_{ij}} := \langle \mathbf{a}_i^{(1)} \mathbf{a}_j^{(1)^T} \rangle$ and $\Sigma_{b_{ij}} := \langle \mathbf{b}_i^{(1)} \mathbf{b}_j^{(1)^T} \rangle$, respectively, for all $i, j \in \mathcal{S}$. Note that each of these clouds has cross covariance terms between their points, as suggested in the above covariance matrices expressions. The optimal translation vector is assumed to have an error model with a similar structure to that of (7.7),

$$\mathbf{t}^* = \mathbf{t}^{(0)} + \epsilon\, \mathbf{t}^{(1)} + \mathcal{O}(\epsilon^2). \tag{7.8}$$

7.4.1 Rotation uncertainty

The rotation matrix obtained through the optimization process described before is restricted to the special orthogonal group SO(n). However, nothing is known about the components of its error model. The following lemma is introduced to address these quantities.

Lemma 7.2 (Error model for orthogonal matrices)**.** *Consider the generic error model*

$$\mathbf{M} = \mathbf{M}^{(0)} + \epsilon \mathbf{M}^{(1)} + \mathcal{O}(\epsilon^2). \tag{7.9}$$

If the matrix $\mathbf{M}$ *belongs to the orthogonal group* $\mathrm{O}(n)$ *then* $\mathbf{M}^{(0)}$ *belongs to the same matrix space, its determinant is equal to the determinant of* $\mathbf{M}$*, and* $\mathbf{M}^{(1)}$ *has the special structure* $\mathbf{M}^{(1)} = \mathbf{S}(\boldsymbol{\varpi})\mathbf{M}^{(0)} = \mathbf{M}^{(0)}\mathbf{S}(\bar{\boldsymbol{\varpi}})$*, with* $\mathbf{S}(\boldsymbol{\varpi}), \mathbf{S}(\bar{\boldsymbol{\varpi}}) \in \mathfrak{so}(n)$ *and* $\boldsymbol{\varpi}, \bar{\boldsymbol{\varpi}} \in \mathbb{R}^{\frac{n(n-1)}{2}}$*. Furthermore, if* $\mathbf{M} \in \mathrm{SO}(n)$*, then* $\mathbf{M}^{(0)} \in \mathrm{SO}(n)$*.*

Proof. The proof is made by exploring the algebraic constraints imposed by the matrix spaces $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ and using error model (7.9). The identity $\mathbf{M}^T\mathbf{M} = \mathbf{I}$ is expanded, resulting in a unperturbed part that yields $\mathbf{M}^{(0)} \in \mathrm{O}(n)$ and a perturbed part that implies that $\mathbf{M}^{(1)} = \mathbf{S}(\boldsymbol{\varpi})\mathbf{M}^{(0)} = \mathbf{M}^{(0)}\mathbf{S}(\bar{\boldsymbol{\varpi}})$. Finally, the determinant of $\mathbf{M}$ noting the structure of $\mathbf{S}(\boldsymbol{\varpi})$ is expanded. The proof is presented in Appendix B.3. □

From the solution to the optimization problem in [Ume91, Lemma], it is known that $\mathbf{R}^* \in \mathrm{SO}(n)$. If (7.9) is applied to $\mathbf{R}^*$, the following error model results

$$\mathbf{R}^* = \Big[\mathbf{I} + \epsilon\, \mathbf{S}(\boldsymbol{\varpi}) + \mathcal{O}(\epsilon^2)\Big]\mathbf{R}^{(0)}. \tag{7.10}$$

To ascertain whether $\mathbf{R}^{(0)} \in \mathrm{SO}(n)$ corresponds to the true rotation, consider that matrix $\mathbf{B}$, used to compute the estimated rotation, can be described in terms of its error model, using that of matrices $\mathbf{X}$ and $\mathbf{Y}$,

$$\mathbf{X} = \mathbf{X}^{(0)} + \epsilon\, \mathbf{X}^{(1)} + \mathcal{O}(\epsilon^2)$$

and

$$\mathbf{Y} = \mathbf{Y}^{(0)} + \epsilon\, \mathbf{Y}^{(1)} + \mathcal{O}(\epsilon^2),$$

which are a generalization of (7.7b) and (7.7a), respectively. This results in

$$\mathbf{B} = \mathbf{B}^{(0)} + \epsilon\, \mathbf{B}^{(1)} + \mathcal{O}(\epsilon^2), \tag{7.11}$$

with $\mathbf{B}^{(0)} = \mathbf{X}^{(0)}\mathbf{W}\mathbf{Y}^{(0)^T}$ and $\mathbf{B}^{(1)} = \mathbf{X}^{(1)}\mathbf{W}\mathbf{Y}^{(0)^T} + \mathbf{X}^{(0)}\mathbf{W}\mathbf{Y}^{(1)^T}$. Furthermore, the singular value decomposition (7.4) of $\mathbf{B}$ can itself be expanded defining similar models for each of its components [LLM08], i.e., $\mathbf{U} = \mathbf{U}^{(0)} + \epsilon\mathbf{U}^{(1)} + \mathcal{O}(\epsilon^2)$, $\mathbf{D} = \mathbf{D}^{(0)} + \epsilon\mathbf{D}^{(1)} + \mathcal{O}(\epsilon^2)$, and $\mathbf{V} = \mathbf{V}^{(0)} + \epsilon\mathbf{V}^{(1)} + \mathcal{O}(\epsilon^2)$, which yields

$$\mathbf{B} = \mathbf{U}^{(0)}\mathbf{D}^{(0)}\mathbf{V}^{(0)^T} + \mathcal{O}(\epsilon).$$

Comparison with the error model (7.11) yields $\mathbf{X}^{(0)}\mathbf{W}\mathbf{Y}^{(0)^T} = \mathbf{U}^{(0)}\mathbf{D}^{(0)}\mathbf{V}^{(0)^T}$, thus confirming that these terms are composed of true quantities only. Combining the error model of the optimal rotation matrix with its solution (7.5) yields

$$\mathbf{R}^{(0)} + \mathcal{O}(\epsilon) = \mathbf{U}^{(0)} \operatorname{diag}\Big(1, \ldots, 1, \big|\mathbf{U}^{(0)}\big|\big|\mathbf{V}^{(0)}\big|\Big)\mathbf{V}^{(0)^T} + \mathcal{O}(\epsilon).$$

Note that, as $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, Lemma 7.2 applies, and, therefore, not only $\mathbf{U}^{(0)}$ and $\mathbf{V}^{(0)}$ are orthogonal matrices, but their determinants correspond to the determinants of the perturbed versions, a fact that was used in the last expression. From this, it follows that

$\mathbf{R}^{(0)}$ is the true rotation as it is obtained from the singular value decomposition of $\mathbf{B}^{(0)}$, itself computed using the true terms $\mathbf{a}_i^{(0)}$ and $\mathbf{b}_i^{(0)}$, for all $i \in \{1,\ldots,m\}$. As $\mathbf{R}^{(0)}$ is the true rotation and given the structure of (7.10), the vector $\boldsymbol{\varpi}$ can be seen as a rotation error around the true rotation matrix. Thus, the next steps describe the computation of this rotation error and its statistical description, starting with some properties associated with the closed-form solution of the optimization problem (7.3).

From the proof of [Ume91, Lemma], it is known that the matrix $\mathbf{B}\mathbf{R}^*$ is symmetric, and thus, using the error models (7.10) and (7.11) one can write,

$$\operatorname{skew}(\mathbf{B}\mathbf{R}^*) = \operatorname{skew}\left(\mathbf{B}^{(0)}\mathbf{R}^{(0)}\right) + \epsilon \operatorname{skew}\left(\mathbf{B}^{(1)}\mathbf{R}^{(0)} + \mathbf{B}^{(0)}\mathbf{S}(\boldsymbol{\varpi})\mathbf{R}^{(0)}\right) = \mathbf{0}.$$

This implies that each skew operator is null, meaning that $\mathbf{B}^{(0)}\mathbf{R}^{(0)}$ is also symmetric and that, after left multiplication by $\mathbf{R}^{(0)}$, right multiplication by ${\mathbf{R}^{(0)}}^T$, and using the identity $\mathbf{S}(\boldsymbol{\varpi}) = -\mathbf{S}^T(\boldsymbol{\varpi})$,

$$\operatorname{skew}\left(\mathbf{R}^{(0)}\mathbf{B}^{(0)}\mathbf{S}^T(\boldsymbol{\varpi})\right) = \operatorname{skew}\left(\mathbf{R}^{(0)}\mathbf{B}^{(1)}\right). \tag{7.12}$$

Further manipulation of this formula is needed in order to obtain a tractable expression for $\boldsymbol{\varpi}$. Consider the matrix $\mathbf{R}^{(0)}\mathbf{B}^{(0)}\mathbf{S}^T(\boldsymbol{\varpi})$ expressed as a summation of point terms, noting that $\mathbf{a}_i^{(0)} = \mathbf{r}_i^{(0)} + \mathbf{t}^{(0)}$ with $\mathbf{r}_i^{(.)} := \mathbf{R}^{(0)}\mathbf{b}_i^{(.)}$ for all $i \in \mathcal{S}$,

$$\mathbf{R}^{(0)}\mathbf{B}^{(0)}\mathbf{S}^T(\boldsymbol{\varpi}) = \sum_{i=1}^{m} \sigma_i^{-2}\left[\mathbf{r}_i^{(0)}{\mathbf{r}_i^{(0)}}^T\mathbf{S}^T(\boldsymbol{\varpi}) - \frac{1}{N_W}\sum_{j=1}^{m}\sigma_j^{-2}\mathbf{r}_i^{(0)}{\mathbf{r}_j^{(0)}}^T\mathbf{S}^T(\boldsymbol{\varpi})\right], \tag{7.13}$$

The same can be done with $\mathbf{R}^{(0)}\mathbf{B}^{(1)}$, yielding

$$\mathbf{R}^{(0)}\mathbf{B}^{(1)} = \sum_{i=1}^{m}\sigma_i^{-2}\left[\mathbf{r}_i^{(0)}{\bar{\mathbf{a}}_i^{(1)}}^T + \mathbf{r}_i^{(1)}{\bar{\mathbf{a}}_i^{(0)}}^T\right], \tag{7.14}$$

where $\bar{\mathbf{a}}_i^{(.)} := \mathbf{a}_i^{(.)} - \boldsymbol{\mu}_A^{(.)}$.

Due to the skew-symmetric nature of (7.12), it is possible to apply the unskew operator to both sides. This enables to apply Properties (4) and (2) of Lemma 7.1 combined with (7.13) and (7.14), yielding

$$\mathbf{S}^{-1}\left(2\operatorname{skew}\left(\mathbf{R}^{(0)}\mathbf{B}^{(0)}\mathbf{S}^T(\boldsymbol{\varpi})\right)\right) = \frac{1}{N_W}\sum_{i,j=1}^{m}\sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\mathbf{S}(\boldsymbol{\varpi})\mathbf{r}_j^{(0)} - \sum_{i=1}^{m}\sigma_i^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\mathbf{S}(\boldsymbol{\varpi})\mathbf{r}_i^{(0)}, \tag{7.15}$$

whereas the right-hand side becomes

$$\mathbf{c} := \mathbf{S}^{-1}\left(2\operatorname{skew}\left(\mathbf{R}^{(0)}\mathbf{B}^{(1)}\right)\right) = \sum_{i=1}^{m}\sigma_i^{-2}\bar{\mathbf{S}}\left(\bar{\mathbf{a}}_i^{(0)}\right)\mathbf{r}_i^{(1)} - \sum_{i=1}^{m}\sigma_i^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\bar{\mathbf{a}}_i^{(1)}. \tag{7.16}$$

Using Property (3) of Lemma 7.1 in (7.15), one obtains a linear matrix equation after applying the unskew operator to both sides of (7.12), given by $\mathbf{A}\boldsymbol{\varpi} = \mathbf{c}$, where the matrix $\mathbf{A} \in \mathbb{R}^{n_p \times n_p}$ is defined as

$$\mathbf{A} := -\sum_{i=1}^{m}\sigma_i^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\bar{\mathbf{S}}^T\left(\mathbf{r}_i^{(0)}\right) + \frac{1}{N_W}\sum_{i,j=1}^{m}\sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\bar{\mathbf{S}}^T\left(\mathbf{r}_j^{(0)}\right).$$

From the linear equation now derived it is possible to obtain $\boldsymbol{\varpi}$, as long as $\mathbf{A}$ is invertible. It can be shown that this result degenerates into the main result of [Dor05] when assuming independent and identically distributed points in $\mathbb{R}^3$. However, the existence and uniqueness of the solution is not addressed there. That is the focus of the forthcoming theorem, finding the conditions under which the invertibility of $\mathbf{A}$ in n-dimensional space holds.

Theorem 7.3 (Uniqueness of the rotation error)**.** *The matrix* $\mathbf{A}$ *is invertible if and only if there are at least n points in the set* $\mathcal{S}_B$ *whose connecting vectors span* $\mathbb{R}^{n-1}$*.*

Proof. The proof is made by manipulating $\boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta} = 0$ so that a clear relation between the landmarks and $\boldsymbol{\eta}$ appears, allowing to explore the nature of $\mathbf{S}(.)$ to analyse the solutions of $\boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta} = 0$ for $\|\boldsymbol{\eta}\| = 1$ when $\mathbf{A}$ is singular. Suppose that $\mathbf{A}$ is singular. Then, there exists a unit vector $\boldsymbol{\eta} \in \mathbb{R}^{n_p}$ such that

$$\boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta} = 0. \tag{7.17}$$

The proof of the sufficiency part follows by contraposition, i.e., by showing that if $\mathbf{A}$ is not-invertible, then the condition of the theorem cannot hold. It can be shown after some algebraic manipulation that (7.17) is the same as

$$\sum_{i=1}^{m}\left\|\sigma_i^{-1}\bar{\mathbf{S}}^T\left(\mathbf{r}_i^{(0)}\right)\boldsymbol{\eta}\right\|^2 - \frac{1}{N_W}\left\|\sum_{i=1}^{m}\sigma_i^{-2}\bar{\mathbf{S}}^T\left(\mathbf{r}_i^{(0)}\right)\boldsymbol{\eta}\right\|^2 = 0, \tag{7.18}$$

where $\boldsymbol{\eta}^T\bar{\mathbf{S}}(\mathbf{d}_i)\bar{\mathbf{S}}^T(\mathbf{d}_i)\boldsymbol{\eta} = \left\|\bar{\mathbf{S}}^T(\mathbf{d}_i)\boldsymbol{\eta}\right\|^2$ was used to obtain the first parcel and $\sum_{i,j}\mathbf{d}_j^T\mathbf{d}_i = \|\sum_i \mathbf{d}_i\|^2$ was employed in the second one.

Consider now Lemma A.2 in Appendix A. Let $k_i := \sigma_i^{-1}$ and $\boldsymbol{\alpha}_i := \bar{\mathbf{S}}^T\left(\mathbf{r}_i^{(0)}\right)\boldsymbol{\eta}$. This implies that $\sum_{j=1}^{N} k_j^2 = N_W$. Then, it can be seen that Lemma A.2 applies, showing that the left hand member of (7.18) is always greater than or equal to zero, and zero if and only if

$$\bar{\mathbf{S}}^T\left(\mathbf{r}_i^{(0)}\right)\boldsymbol{\eta} = \bar{\mathbf{S}}^T\left(\mathbf{r}_j^{(0)}\right)\boldsymbol{\eta},\ \forall i,j \in \mathcal{S}, i \neq j.$$

This can be rewritten using Property (3) of Lemma 7.1 to obtain

$$\mathbf{S}(\boldsymbol{\eta})\left[\mathbf{r}_1^{(0)}\ \mathbf{r}_2^{(0)}\ \cdots\ \mathbf{r}_m^{(0)}\right] = \mathbf{e}\mathbf{1}^T, \tag{7.19}$$

where $\mathbf{e} = \left[e_1\ \cdots\ e_n\right] = \mathbf{S}(\boldsymbol{\eta})\mathbf{r}_i^{(0)}$ for any $i \in \mathcal{S}$. As $\mathbf{S}(\boldsymbol{\eta})$ is a skew-symmetric matrix, if $\boldsymbol{\eta} \neq \mathbf{0}$, then the rank of $\mathbf{S}(\boldsymbol{\eta})$, $p := \operatorname{rank}\mathbf{S}(\boldsymbol{\eta})$, is always greater than or equal to 2. Thus, it is possible to write

$$\begin{cases} \mathbf{d}_1^T\left[\mathbf{r}_1^{(0)}\ \mathbf{r}_2^{(0)}\ \cdots\ \mathbf{r}_m^{(0)}\right] = e_1\mathbf{1}^T \\ \vdots \\ \mathbf{d}_p^T\left[\mathbf{r}_1^{(0)}\ \mathbf{r}_2^{(0)}\ \cdots\ \mathbf{r}_m^{(0)}\right] = e_p\mathbf{1}^T \end{cases}, \tag{7.20}$$

where all the $\mathbf{d}_i \in \mathbb{R}^n$, $i = 1,\ldots,p$, are linearly independent vectors. Each of these equations defines distinct hyperplanes of dimension $n-1$, and, as such, (7.20) restricts the points to an hyperplane of dimensions $n-p$. Hence, the solution of (7.19) implies that either $\eta = \mathbf{0}$, contradicting the hypothesis of the proof ($\|\eta\| = 1$), or that the points in $\mathcal{S}_B$ are restricted to an hyperplane of dimension $n-2$, at most. Thus, if the matrix $\mathbf{A}$ is singular, the condition of the theorem cannot hold, and, by contraposition, if the conditions hold, i.e., if the points are not restricted to an hyperplane of dimension $n-2$ at most, or, equivalently, if from all the combinations of $\mathbf{r}_i^{(0)} - \mathbf{r}_j^{(0)}$ for all $i \neq j \in \mathcal{S}$ one can find $n-1$ that span $\mathbb{R}^{n-1}$, then $\mathbf{A}$ is invertible, concluding the sufficiency part of the proof.

For the proof of the necessity part, suppose that the matrix $\mathbf{A}$ is indeed invertible. In that case, the only solution of (7.17) is the trivial solution $\eta = \mathbf{0}$. Following the same reasoning of the sufficiency proof, this implies that $\eta = \mathbf{0}$ is also the only solution of (7.19). If the points in $\mathcal{S}_B$ are at most restricted to an hyperplane of dimension $n-2$, (7.19) will be satisfied for $\eta \neq \mathbf{0}$, as seen before in the proof of sufficiency. Therefore, if $\eta = \mathbf{0}$ is indeed the only solution of (7.19), the points in $\mathcal{S}_B$ cannot be restricted to $\mathbb{R}^{n-2}$, i.e., their connecting vectors must span $\mathbb{R}^{n-1}$. Hence, if the matrix $\mathbf{A}$ is invertible, the conditions of the theorem must hold, which concludes the proof. □

By observation of (7.16), it can be seen that $\langle \mathbf{c} \rangle = 0$, as its uncertain elements are only $\mathbf{r}_i^{(0)}$ and $\bar{\mathbf{a}}_i^{(0)}$, which have zero mean. Therefore $\langle \varpi \rangle = \mathbf{A}^{-1}\langle \mathbf{c} \rangle$ is also zero, and the corresponding covariance matrix is given by

$$\begin{aligned}
\Sigma_{\varpi} &= \langle \varpi\varpi^T \rangle = \mathbf{A}^{-1}\langle \mathbf{c}\mathbf{c}^T \rangle \mathbf{A}^{-1^T}\\
&= \mathbf{A}^{-1}\Bigg[\sum_{i,j=1}^{m} \sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\bar{\mathbf{a}}_i^{(0)}\right)\mathbf{R}^{(0)}\Sigma_{b_{ij}}\mathbf{R}^{(0)^T}\bar{\mathbf{S}}^T\left(\bar{\mathbf{a}}_j^{(0)}\right)\\
&\qquad + \sum_{i,j=1}^{m} \sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\langle \bar{\mathbf{a}}_i^{(1)}\bar{\mathbf{a}}_j^{(1)^T}\rangle \bar{\mathbf{S}}^T\left(\mathbf{r}_j^{(0)}\right)\\
&\qquad - \sum_{i,j=1}^{m} \sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\bar{\mathbf{a}}_i^{(0)}\right)\langle \mathbf{r}_i^{(1)}\bar{\mathbf{a}}_j^{(1)^T}\rangle \bar{\mathbf{S}}^T\left(\mathbf{r}_j^{(0)}\right)\\
&\qquad - \sum_{i,j=1}^{m} \sigma_i^{-2}\sigma_j^{-2}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\langle \bar{\mathbf{a}}_i^{(1)}\mathbf{r}_j^{(1)^T}\rangle \bar{\mathbf{S}}^T\left(\bar{\mathbf{a}}_j^{(0)}\right)\Bigg]\mathbf{A}^{-1^T}.
\end{aligned} \tag{7.21}$$

The resulting individual covariances in the summations are either already defined or can be obtained from similar computations. The covariances related to points in $\mathcal{S}_A$ are given by

$$\langle \bar{\mathbf{a}}_i^{(1)}\bar{\mathbf{a}}_j^{(1)^T} \rangle = \Sigma_{a_{ij}} - \frac{1}{N_W}\sum_{k=1}^{m}\sigma_k^{-2}\left(\Sigma_{a_{ik}} + \Sigma_{a_{kj}}\right) + \frac{1}{N_W^2}\sum_{k,l=1}^{m}\sigma_k^{-2}\sigma_l^{-2}\Sigma_{a_{kl}},$$

and

$$\langle \bar{\mathbf{a}}_i^{(1)}\mathbf{a}_j^{(1)^T} \rangle = \Sigma_{a_{ij}} - \frac{1}{N_W}\sum_{k=1}^{m}\sigma_k^{-2}\Sigma_{a_{kj}}.$$

while the cross term is

$$\langle \bar{\mathbf{a}}_i^{(1)} {\mathbf{r}_j^{(1)}}^T \rangle = \Sigma_{a_i b_j} {\mathbf{R}^{(0)}}^T - \frac{1}{N_W} \sum_{k=1}^{m} \sigma_k^{-2} \Sigma_{a_k b_j} {\mathbf{R}^{(0)}}^T.$$

The cross-covariance between points of both sets, i.e., $\Sigma_{a_i b_j}$, is specific to the application, and may in most cases be zero due to the possible independence of the two sets.

7.4.2 Translation uncertainty

The optimal translation between frames is given by (7.6) and the associated error model is assumed to be (7.8). Using this information along with the error models for the points in both sets, defined in (7.7), it is possible to expand (7.6) to obtain

$$\mathbf{t}^* = \frac{1}{N_W} \sum_{i=1}^{m} \sigma_i^{-2} \left[\left(\mathbf{a}_i^{(0)} + \epsilon \mathbf{a}_i^{(1)} + \mathcal{O}(\epsilon^2) \right) - \left(\mathbf{I}_n + \epsilon \mathbf{S}(\boldsymbol{\varpi}) + \mathcal{O}(\epsilon^2) \right) \mathbf{R}^{(0)} \left(\mathbf{b}_i^{(0)} + \epsilon \mathbf{b}_i^{(1)} + \mathcal{O}(\epsilon^2) \right) \right].$$

from which expressions for the error model components can be found, yielding, respectively, the true translation

$$\mathbf{t}^{(0)} = \frac{1}{N_W} \sum_{i=1}^{m} \sigma_i^{-2} \left[\mathbf{a}_i^{(0)} - \mathbf{R}^{(0)} \mathbf{b}_i^{(0)} \right],$$

and the first order perturbed part

$$\mathbf{t}^{(1)} = \frac{1}{N_W} \sum_{i=1}^{m} \sigma_i^{-2} \left(\mathbf{a}_i^{(1)} - \mathbf{S}(\boldsymbol{\varpi}) \mathbf{r}_i^{(0)} - \mathbf{r}_i^{(1)} \right). \tag{7.22}$$

It is confirmed that $\mathbf{t}^{(1)}$ has zero mean, noting that all the perturbation quantities involved have zero mean. Therefore, the covariance matrix of the position estimate Σ_t is simply given by $\Sigma_t = \langle \mathbf{t}^{(1)} {\mathbf{t}^{(1)}}^T \rangle$, and it can be computed by expanding $\mathbf{t}^{(1)}$ according to (7.22), using Property (3) of Lemma 7.1 to extract $\boldsymbol{\varpi}$ from the skew-symmetric matrix. This yields

$$\begin{aligned} \Sigma_t = \sum_{i,j=1}^{m} \frac{\sigma_i^{-2} \sigma_j^{-2}}{N_W^2} \Big(& \Sigma_{a_{ij}} + \mathbf{R}^{(0)} \Sigma_{b_{ij}} {\mathbf{R}^{(0)}}^T + \bar{\mathbf{S}}^T \left(\mathbf{r}_i^{(0)} \right) \Sigma_{\varpi} \bar{\mathbf{S}} \left(\mathbf{r}_j^{(0)} \right) \\ & - \bar{\mathbf{S}}^T \left(\mathbf{r}_i^{(0)} \right) \langle \boldsymbol{\varpi} {\mathbf{a}_j^{(1)}}^T \rangle - \langle \mathbf{a}_i^{(1)} \boldsymbol{\varpi}^T \rangle \bar{\mathbf{S}} \left(\mathbf{r}_j^{(0)} \right) - \mathbf{R}^{(0)} \Sigma_{b_i a_j} \\ & + \bar{\mathbf{S}}^T \left(\mathbf{r}_i^{(0)} \right) \langle \boldsymbol{\varpi} {\mathbf{r}_j^{(1)}}^T \rangle + \langle \mathbf{r}_i^{(1)} \boldsymbol{\varpi}^T \rangle \bar{\mathbf{S}} \left(\mathbf{r}_j^{(0)} \right) - \Sigma_{a_i b_j} {\mathbf{R}^{(0)}}^T \Big), \end{aligned} \tag{7.23}$$

The computation of the cross covariances between points in $\mathcal{S}_A$ and the rotation error leads to

$$\langle \boldsymbol{\varpi} {\mathbf{a}_j^{(1)}}^T \rangle = \mathbf{A}^{-1} \sum_{i=1}^{m} \sigma_i^{-2} \left[\bar{\mathbf{S}} \left(\bar{\mathbf{a}}_i^{(0)} \right) \mathbf{R}^{(0)} \Sigma_{b_i a_j} - \bar{\mathbf{S}} \left(\mathbf{r}_i^{(0)} \right) \langle \bar{\mathbf{a}}_i^{(1)} {\mathbf{a}_j^{(1)}}^T \rangle \right], \tag{7.24}$$

and

$$\langle \boldsymbol{\varpi} {\bar{\mathbf{a}}_j^{(1)}}^T \rangle = \mathbf{A}^{-1} \sum_{i=1}^{m} \sigma_i^{-2} \left[\bar{\mathbf{S}} \left(\bar{\mathbf{a}}_i^{(0)} \right) \langle \mathbf{r}_i^{(1)} {\bar{\mathbf{a}}_j^{(1)}}^T \rangle - \bar{\mathbf{S}} \left(\mathbf{r}_i^{(0)} \right) \langle \bar{\mathbf{a}}_i^{(1)} {\bar{\mathbf{a}}_j^{(1)}}^T \rangle \right],$$

while the cross covariances between points in $\mathcal{S}_B$ and the rotation error are given by

$$\langle \boldsymbol{\varpi} {\mathbf{r}_j^{(1)}}^T \rangle = \mathbf{A}^{-1} \sum_{i=1}^{m} \sigma_i^{-2} \left[\bar{\mathbf{S}} \left(\bar{\mathbf{a}}_i^{(0)} \right) \mathbf{R}^{(0)} \Sigma_{b_{ij}} {\mathbf{R}^{(0)}}^T - \bar{\mathbf{S}} \left(\mathbf{r}_i^{(0)} \right) \langle \bar{\mathbf{a}}_i^{(1)} {\mathbf{r}_j^{(1)}}^T \rangle \right]. \tag{7.25}$$

7.4.3 Cross translation-rotation uncertainty

The optimal translation estimate is obtained using the optimal rotation, and for that reason there is a correlation between the two. The corresponding cross covariance is given by

$$\begin{aligned}\Sigma_{\varpi t} &= \mathbf{A}^{-1}\sum_{i,j=1}^{m}\frac{\sigma_i^{-2}\sigma_j^{-2}}{N_W}\bar{\mathbf{S}}\left(\bar{\mathbf{a}}_i^{(0)}\right)\left(\mathbf{R}^{(0)}\Sigma_{b_i a_j}-\langle\mathbf{r}_i^{(1)}\varpi^T\rangle\bar{\mathbf{S}}\left(\mathbf{r}_j^{(0)}\right)-\mathbf{R}^{(0)}\Sigma_{b_i j}\mathbf{R}^{(0)^T}\right)\\ &+\mathbf{A}^{-1}\sum_{i,j=1}^{m}\frac{\sigma_i^{-2}\sigma_j^{-2}}{N_W}\bar{\mathbf{S}}\left(\mathbf{r}_i^{(0)}\right)\left(\Sigma_{a_i b_j}\mathbf{R}^{(0)^T}+\langle\bar{\mathbf{a}}_i^{(1)}\varpi^T\rangle\bar{\mathbf{S}}\left(\mathbf{r}_j^{(0)}\right)-\langle\bar{\mathbf{a}}_i^{(1)}\mathbf{a}_j^{(1)^T}\rangle\right),\end{aligned} \tag{7.26}$$

where the cross covariances between points in either sets and the rotation error are computed in the same manner, and are detailed in the sequel. The cross-covariances between points in $\mathcal{S}_B$ and $\mathcal{S}_A$ with the translation can be computed following the reasoning used to obtain (7.24), i.e.,

$$\langle\mathbf{t}^{(1)}\mathbf{r}_j^{(1)^T}\rangle=\frac{1}{N_W}\sum_{i=1}^{m}\sigma_i^{-2}\left(\Sigma_{a_i b_j}\mathbf{R}^{(0)^T}-\bar{\mathbf{S}}^T\left(\mathbf{R}^{(0)}\mathbf{b}_i^{(0)}\right)\langle\varpi\mathbf{r}_j^{(1)^T}\rangle-\mathbf{R}^{(0)}\Sigma_{b_{ij}}\mathbf{R}^{(0)^T}\right), \tag{7.27}$$

and the cross-covariance between points in $\mathcal{S}_A$ and the translation is given by

$$\langle\mathbf{t}^{(1)}\mathbf{a}_j^{(1)^T}\rangle=\frac{1}{N_W}\sum_{i=1}^{m}\sigma_i^{-2}\left(\Sigma_{a_{ij}}-\bar{\mathbf{S}}^T\left(\mathbf{R}^{(0)}\mathbf{b}_i^{(0)}\right)\langle\varpi\mathbf{a}_j^{(1)^T}\rangle-\mathbf{R}^{(0)}\Sigma_{b_i a_j}\right).$$

In these expressions, Property (3) from Lemma 7.1 was once more used to extract ϖ from the skew-symmetric matrix. The resulting individual covariances in the summations are either already defined or can be obtained from similar computations.

Remark 7.1. *As the true quantities* $\mathbf{R}^{(0)}$, $\mathbf{a}_i^{(0)}$, *and* $\mathbf{b}_i^{(0)}$ *are unknown for all* $i\in\mathcal{S}$, *a possible approximation is to use the optimal values for the rotation and translation and the original perturbed points* $\mathbf{a}_i$ *and* $\mathbf{b}_j$ *instead. This applies to all the covariance expressions derived in this section.*

7.5 Algorithm validation

In this section, the proposed uncertainty characterization is validated through extensive Monte Carlo simulations. Despite the general derivation presented above, the numerical validation requires the particularization for one or several dimensions. Although bidimensional and tridimensional problems are the most common in the literature, as in [GBSO13] and in the previous chapters of this dissertation, the study of rotations and other quantities in seven-dimensional space has also been addressed in [Mas83]. For these reasons and to represent meaningful examples, this section will focus on those dimensionalities.

As the optimization and uncertainty description presented in this chapter are intended to work for any rotation and translation, and taking into account that some approximations were introduced, a Monte Carlo simulation with $N = 1000$ samples was performed for each of $M = 500$ random configurations, to validate the proposed technique. In each of these configurations, the transformation is uniformly distributed, with each component of the translation vector ranging from -10 to 10 m, the locations of the points from 0 and 10 m in each component, and the rotation matrix is built using the exponential map $\mathbf{R} = \mathrm{e}^{\mathbf{S}(\theta \mathbf{u})}$, where θ is uniformly distributed between $-\pi$ and π radians and $\mathbf{u}$ is an uniformly distributed random unit vector. In order to comprehensively test the properties of the proposed method, the covariance matrices used to generate the normally distributed location errors of the points in each set are complete covariances with cross-correlations between all the points and different individual covariances. In the two-dimensional case, the eigenvalues of the covariance matrices are $(0.01^2, 0.05^2)$. For the three-dimensional case, the eigenvalues are $(0.01^2, 0.03^2, 0.05^2)$, and the seven-dimensional eigenvalues are $(0.01^2, 0.017^2, 0.023^2, 0.03^2, 0.037^2, 0.043^2, 0.05^2)$. In each of these cases, all the permutations of eigenvalues are used an equal number of times. The full covariance for each point cloud is then $\Sigma^{nom} = \mathbf{U}\,\mathrm{diag}(\lambda_1, \ldots, \lambda_{nm})\,\mathbf{U}^T$, where the eigenvector matrix $\mathbf{U}$ is an $nm \times nm$ rotation matrix built following the same reasoning used to build $\mathbf{R}$.

The performance of the algorithm was tested for a different number of landmarks, and the summary of the results can be found in Tables 7.1, 7.2, 7.3, and 7.4. The mean and standard deviation of the differences between the results obtained and the true values is shown in Table 7.1. The absolute angular difference between the rotation matrices $\mathbf{R}^{true}$ and $\mathbf{R}^{opt}$ is defined as

$$\mathrm{d}(\mathbf{R}) := \left\| \mathbf{S}^{-1}\left(\log\left(\mathbf{R}^{opt}\mathbf{R}^{true^T}\right)\right) \right\|$$

where the superscript ^{opt} is used to denote the quantities obtained using the solution to the optimization problem and the uncertainty computations derived above, whereas the superscript ^{true} denotes the respective true values. The relative difference for the translation can be defined as

$$\mathrm{d}(\mathbf{t}) := 100 \frac{\|\mathbf{t}^{opt} - \mathbf{t}^{true}\|}{\|\mathbf{t}^{true}\|}.$$

The norm of the rotation error is greater in the tridimensional case for every number of points tested, but that may be comprehensible given that the error has thrice the dimension. The relative translation difference for $n = 3$ is also larger for smaller numbers of points, but quickly converges to similar values. In any case, the distances to the true values are quite small, and decrease with the number of available points as it is expected.

The meaningful comparison between two different models, usually denoted as the null model and the alternate model, is a difficult problem which is frequently performed using likelihood ratio tests. In short, the likelihood ratio expresses how many times more likely the data is explained under one model than the other, by obtaining the probability distribution of the test, assuming the null model to be true (null hypothesis). Thus, the null hypothesis is rejected if its probability is below a desired significance value, typically 0.01 or 0.05. However, a central limitation for the usage of these tests is the assumption that the stochastic variable follows a specific distribution, usually Gaussian. Indeed, and as a consequence of the considered arbitrary rotations, the translation tends to be Gaussian distributed as the number of points increases, but may fail to be so for a reduced number of landmarks, which is in accordance with the central limit theorem.

Table 7.1: Performance of the algorithm in terms of the distance to the true values (mean±std)

	$n=2$		$n=3$	
m	d(**R**) [deg]	d(**t**) [%]	d(**R**) [deg]	d(**t**) [%]
2	0.948 ± 2.36	2.59 ± 9.57	–	–
3	0.493 ± 0.52	1.40 ± 3.42	1.419 ± 1.70	2.26 ± 3.56
4	0.369 ± 0.32	1.04 ± 1.83	0.803 ± 0.51	1.31 ± 1.20
5	0.308 ± 0.25	0.88 ± 1.59	0.634 ± 0.34	1.05 ± 0.90
7	0.246 ± 0.19	0.72 ± 1.28	0.480 ± 0.23	0.81 ± 0.63
10	0.195 ± 0.15	0.58 ± 1.05	0.378 ± 0.17	0.64 ± 0.48
20	0.132 ± 0.10	0.40 ± 0.67	0.255 ± 0.11	0.43 ± 0.32
50	0.082 ± 0.06	0.25 ± 0.43	0.156 ± 0.07	0.27 ± 0.20
100	0.058 ± 0.04	0.18 ± 0.31	0.109 ± 0.05	0.19 ± 0.14

Figures 7.1-7.3 and Tables 7.2, 7.3 and 7.4 summarize the likelihood ratio tests for the comparison between the sample covariance matrices resulting from the Monte Carlo simulations, Σ_{ϖ}^{sim}, Σ_{t}^{sim}, and the joint error covariance Σ_{ξ}^{sim} where $\boldsymbol{\xi} = \left[\boldsymbol{\varpi}^T \mathbf{t}^T\right]^T$, and the covariance matrices resulting from the proposed methodology and respective uncertainty approximation, Σ_{ϖ}^{opt}, Σ_{t}^{opt}, and Σ_{ξ}^{opt}. The definition of each ratio depends on whether the quantity is a scalar or a matrix. For $n=2$, the rotation error is a scalar, and therefore its covariance will be too. Then, the ratio for the null hypothesis $H_{\varpi}^{n=2} : \Sigma_{\varpi}^{sim} = \Sigma_{\varpi}^{opt}$ is defined as

$$\lambda_{\varpi}^{*} := N^{*} \frac{\Sigma_{\varpi}^{sim}}{\Sigma_{\varpi}^{opt}} \sim \chi_{N^{*}}^{2}$$

which is asymptotically χ^2 distributed with $N^{*} := N-1$ degrees of freedom and, thus, can be compared with a predefined significance level, which in the presented results is considered to be 0.01. For the multivariable cases of the null hypothesis of the general case for rotation $H_{\varpi}^{n\geq 3} : \Sigma_{\varpi}^{sim} = \Sigma_{\varpi}^{opt}$, the translation $H_{t} : \Sigma_{t}^{sim} = \Sigma_{t}^{opt}$ and the joint error covariance $H_{\xi} : \Sigma_{\xi}^{sim} = \Sigma_{\xi}^{opt}$, the specific likelihood ratio test can be found in [And84, Section 10.8]. The likelihood ratio for these cases is defined as

$$\lambda_{1}^{*} = \left(\frac{\mathrm{e}}{N^{*}}\right)^{\frac{1}{2}pN^{*}} \left|\mathbf{B}^{sim}\Sigma^{opt-1}\right|^{\frac{1}{2}N^{*}} \mathrm{e}^{-\frac{1}{2}\operatorname{tr}\left(\mathbf{B}^{sim}\Sigma^{opt-1}\right)},$$

where $\mathbf{B}^{sim} := N^{*}\Sigma^{sim}$, and the comparison with the significance level is achieved by noting that $-2\log\lambda_{1}^{*}$ is asymptotically χ^2 distributed with $\frac{1}{2}p(p+1)$ degrees of freedom where p is the dimension of the vector (n for the translation, n_p for the rotation error, and $n+n_p$ for the joint error vector $\boldsymbol{\xi}$). Hence, for the likelihood ratio tests, λ_{ϖ}^{*} is compared with a threshold $t_{\varpi}^{n=2} := F_{\chi_{N^{*}}^{2}}^{-1}(1-\alpha)$ in the bidimensional case, and the remaining likelihood tests are made by comparing $-2\log\lambda_{1}^{*}$ with similar thresholds $t_{\varpi}^{n\geq 3} := F_{\chi_{n_p}^{2}}^{-1}(1-\alpha)$, $t_{t} := F_{\chi_{n}^{2}}^{-1}(1-\alpha)$ and $t_{\xi} := F_{\chi_{n_p+n}^{2}}^{-1}(1-\alpha)$, all with $\alpha = 0.01$.

As the true values of the coordinates of the points and of the rotation matrix are not available in real-world scenarios, the covariances resulting from the uncertainty description of the

Table 7.2: Worst-case covariance likelihood ratio tests for $n = 2$.

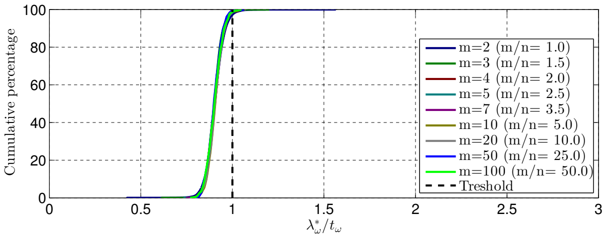
m	H_ω [%]	H_t [%]	H_ξ [%]	H_ω^{100} [%]	H_t^{100} [%]	H_ξ^{100} [%]
2	77.2	27.8	29.2	100	71.4	73.2
3	94.4	51.8	55.4	100	90.2	90.4
4	96.6	65.4	70.8	100	96.2	95.4
5	96.8	79.4	82.8	100	97.6	97.8
7	96.4	87.6	89.2	100	97.0	97.0
10	98.8	93.2	94.8	100	98.2	98.4
20	97.4	95.8	96.6	100	99.0	98.8
50	98.2	98.6	98.4	100	99.6	99.2
100	99.0	98.4	98.6	100	99.0	99.6

Table 7.3: Worst-case covariance likelihood ratio tests for $n = 3$.

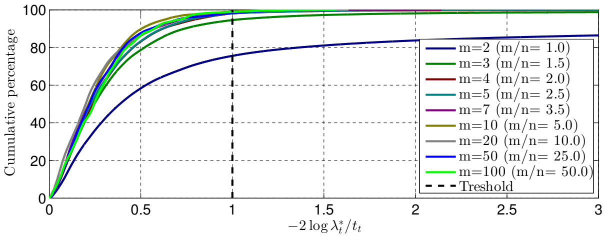
m	H_ω [%]	H_t [%]	H_ξ [%]	H_ω^{100} [%]	H_t^{100} [%]	H_ξ^{100} [%]
3	67.6	13.4	9.8	95.4	60.0	60.2
4	88.4	32.8	27.8	98.0	84.8	85.8
5	93.8	50.0	47.4	98.8	93.8	93.6
7	96.0	76.6	75.2	99.0	97.8	98.8
10	98.2	92.0	91.2	97.8	96.8	97.2
20	97.8	96.6	96.0	99.0	97.8	97.6
50	98.8	98.0	97.8	99.2	99.0	98.2
100	98.4	98.0	97.6	98.6	99.4	99.4

Table 7.4: Worst-case covariance likelihood ratio tests for $n = 7$.

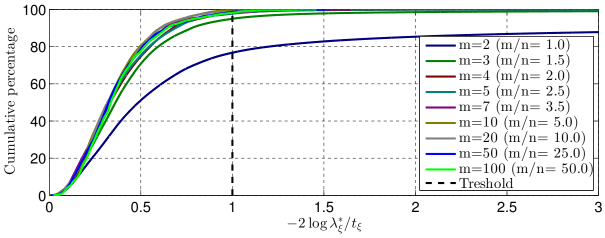
m	H_ω [%]	H_t [%]	H_ξ [%]	H_ω^{100} [%]	H_t^{100} [%]	H_ξ^{100} [%]
7	61.4	1.6	0.0	84.4	32.2	7.4
10	95.2	28.6	9.0	89.0	89.8	56.4
15	98.0	74.6	65.2	93.0	97.2	70.2
20	98.2	90.2	86.0	90.8	98.2	73.4
30	98.4	96.6	94.0	93.0	98.6	75.8
50	98.6	97.6	97.4	93.4	98.4	76.0
100	98.4	97.8	98.2	93.0	99.0	79.6



(a) Rotation error $\boldsymbol{\varpi}$



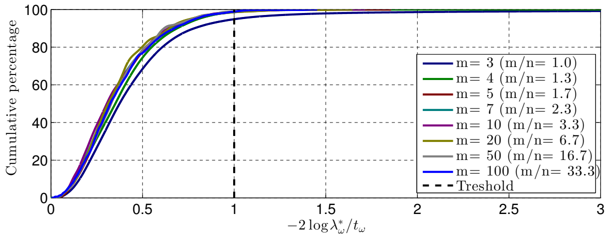
(b) Translation error $\mathbf{t}$



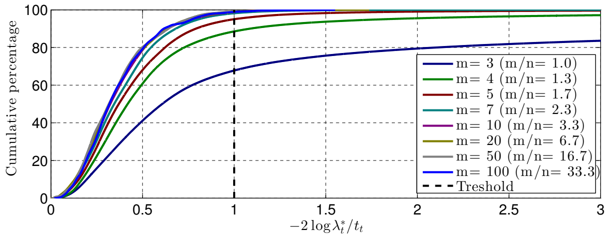
(c) Joint error vector $\boldsymbol{\xi}$

Figure 7.1: The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and nominal input covariances.

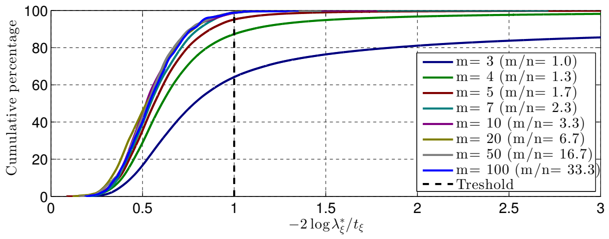
previous section must be computed using the perturbed quantities instead, therefore depending on the actual values of the simulation. The covariances and respective likelihood ratios are computed for each sample of the Monte Carlo simulations, that is, for each configuration, one Σ^{sim} and N instances of Σ^{opt} are computed, resulting in $N \times M$ covariances, and consequently, the same quantity of ratios and tests. The highest ratio in each configuration is used for the hypothesis tests presented in Tables 7.2, 7.3, and 7.4, thus representing the worst-case scenarios (in terms of likelihood ratios) for each configurations and number of landmarks. The values shown in the tables for these tests denote the percentage of successful tests among the M different configurations for $n = 2$, $n = 3$, and $n = 7$, respectively. As mentioned before, the Gaussian assumption may not be valid when dealing with an arbitrary rotation $\mathbf{R}$, which, together with the fact that the ratio expressions and the degrees of freedom of their approximation depend on the number of Monte Carlo samples, the tests are expected to fail more pronouncedly when the number of Monte Carlo samples increases. For this reason, an additional set of likelihood ratio tests are presented using a smaller number of samples, $N = 100$, and denoted as H^{100}. From the values shown in these tables, it can be concluded that, for more than 5 points in 2-D, 7 points in 3-D, and 15 points in seven-dimensional (7-D), the obtained covariance matrices



(a) Rotation error $\boldsymbol{\varpi}$



(b) Translation error $\mathbf{t}$

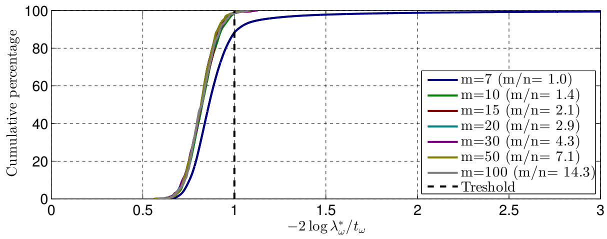


(c) Joint error vector $\boldsymbol{\xi}$

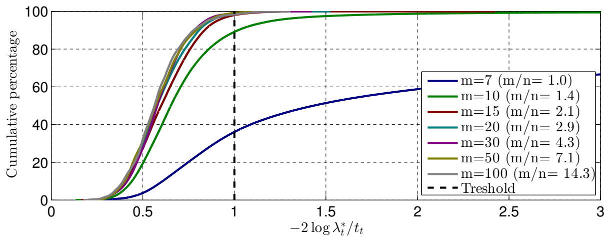
Figure 7.2: The cumulative distribution of the likelihood ratios for the 3-D case including all the covariances built with each particular simulation values and nominal input covariances.

pass more than 70% of the tests in the worst-case, and can therefore be considered a good approximation.

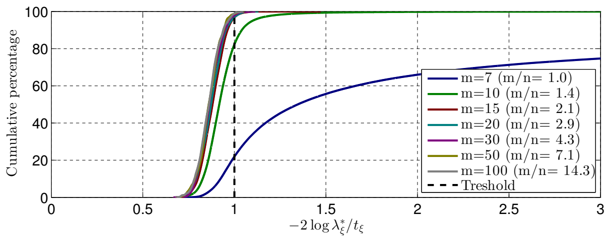
The cumulative distributions of the ratios depicted in Figures 7.1-7.3, however, show that the worst-case scenario in terms of likelihood ratio is not representative of the majority of the computed covariance matrices. These figures present the cumulative distribution of the total $N \times M$ likelihood ratios, normalized by the respective thresholds, and therefore depict the quantity of samples that is explained by the model according to the null hypothesis (area to the left of the vertical dashed line) as well as its variation with the number of available points and space dimensions. Comparing with the worst cases provided in the previous tables, it can be seen that when using the covariances built with every sample in each configuration the results are significantly better even when a low ratio of points per dimension is present. In particular, even with the minimum amount of points ($m/n = 1$), it is possible to achieve more than 70% of positive rotation and translation tests for all the dimensions except for $n = 7$, where 10 points are needed to surpass that threshold. In fact, in most cases the growth of the cumulative function is quite fast, reaching high values of cumulative percentage for relatively low values of the ratio. It is also confirmed that adding a few points to the $m/n = 1$ case leads



(a) Rotation error $\boldsymbol{\varpi}$



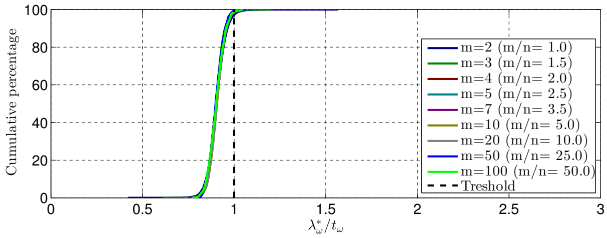
(b) Translation error $\mathbf{t}$



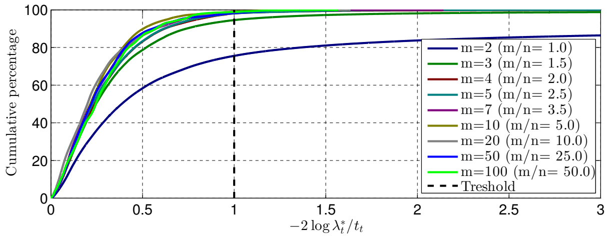
(c) Joint error vector $\boldsymbol{\xi}$

Figure 7.3: The cumulative distribution of the likelihood ratios for the 7-D case including all the covariances built with each particular simulation values and nominal input covariances.

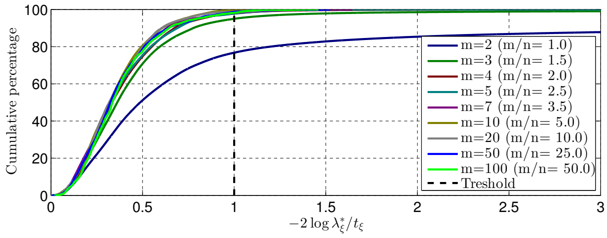
to a relevant improvement of the test results for all the three n, noting that even the joint error ratio displays success values above 90% for each dimensionality, respectively with $m = 3$, $m = 5$, and $m = 10$. Nonetheless, the $m/n = 1$ test results are degraded with the increase of n (from 79% in 2-D to 64% in 3-D and 22% in 7-D for the joint error), a tendency less apparent for the tests with $m/n \geq 2$. To better explore the various aspects of the uncertainty characterization, and achieve a broader validation, additional simulations for $n = 2$ were performed with several noise scales, namely: 1) nominal noise; 2) $\mathcal{S}_B$ noise scaled by 10^2; 3) $\mathcal{S}_A$ noise scaled by 10^2; and 4) both multiplied by 10^2, which are depicted in Figures 7.1, 7.4, 7.5, and 7.6. One important fact to note prior to this analysis is that, as the covariances are built with the noise-perturbed point sets and the estimated rotation, increasing the noise will inevitably lead to worse test results in general. Notwithstanding this fact, for a reasonable number of points ($m = 10$) more than 70% of the samples pass the test for all the quantities tested in all 4 profiles. Regardless of the noise profile, the rotation error covariances perform much better than the translation in these tests, as 3 points are enough to have more than 70% of positive tests for the rotation, while for the translation 7 points are needed for profiles 2 and 3, and 10 points for the last profile. Furthermore, with 2 points the performance may be below 20% in the translation for



(a) Rotation error $\boldsymbol{\varpi}$



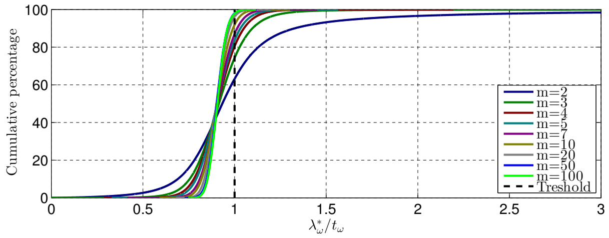
(b) Translation error $\mathbf{t}$



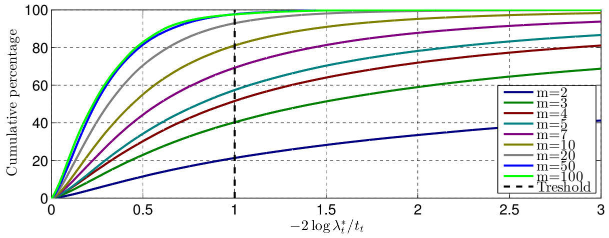
(c) joint error vector $\boldsymbol{\xi}$

Figure 7.4: The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 2 input covariances.

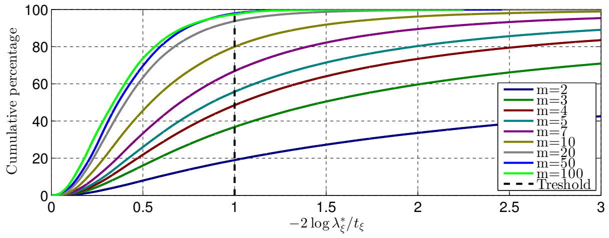
profile 4, while the rotation is never below 60% in any profile. Even though Figures 7.4 and 7.5 are very similar, the translation tests are worse in the former (around 1.4% less for each m), when the higher noise is on $\mathcal{S}_B$, while the rotation tests are better (around 0.7% higher for each m). This may be explained by the fact that, in the mathematical expressions derived in Section 7.4.2, the actual values of the points in that set are always multiplied by the rotation, hence amplifying the influence of the noise in that set. When both input covariances are scaled, the results worsen as every distribution crosses the threshold with around 10% less samples than in profiles 2 and 3. Note however that in realistic situations, these noise levels would imply that 99% of the noisy points could be in a ellipse with a major axis of 1.5 meters around the true value (3σ) which is not a usual situation. Nevertheless, to further assess the influence of the noise in the tests, new simulations were made where both nominal input covariances are multiplied by 100^2. In these tests, the rotation still performs adequately (60% for 20 points, 80% for 50 points), but the translation is far from that level (with 100 points only 3.37% of the samples pass the test). Nevertheless, when the true values are used to compute the covariances in these harsh noise conditions the performance is much better (55% for rotation, 44% for translation, and 12% for the joint error vector, all for $m = 2$), confirming that it is also the fact



(a) Rotation error $\boldsymbol{\varpi}$



(b) Translation error $\mathbf{t}$

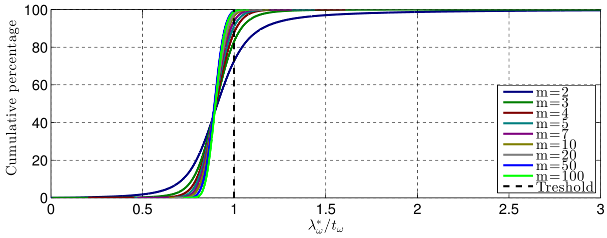


(c) joint error vector $\boldsymbol{\xi}$

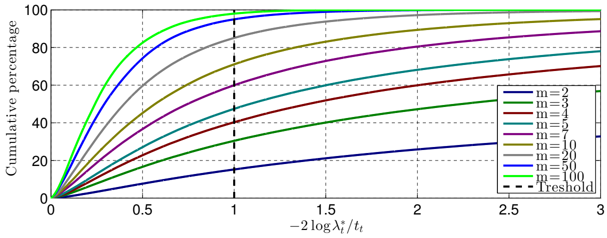
Figure 7.5: The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 3 input covariances.

that the covariances must be built with the perturbed values that leads to the worsening of the results. This is intrinsic to the nonlinear problem at hand, and not specific to the proposed characterization.

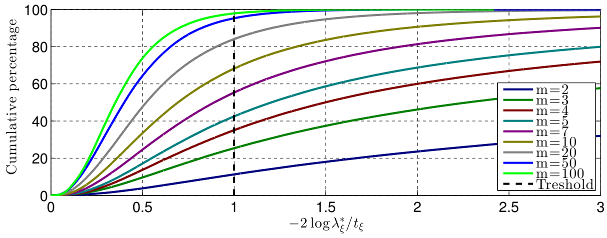
To provide a better understanding of the underlying problem, Figures 7.7 and 7.8 are provided, depicting the spatial distribution of the translation error (red dots) with the corresponding 99% bound in dashed black, along with 3σ bounds given by the median ratio covariance matrix in solid blue, and the simulation covariance in dashed green. In Figure 7.7, a difficult case is presented, where all the points in $\mathcal{S}_B$ are evenly distributed along the x-axis, there is a translation of 2 units in each axis and a 180° rotation between sets, with narrow input covariances on both sets, $\Sigma_b = \Sigma_a = \text{diag}\left(10^{-5}, 9\times10^{-3}, 4\times10^{-5}, 16\times10^{-3}\right)$ for each two points. The baseline (Δ), i.e., the maximum distance between points in $\mathcal{S}_B$, increases from the leftmost figure to the rightmost. When moving downwards, the number of equidistant points increases, maintaining the baseline. A careful analysis of the information contained in Figure 7.7 indicates that the baseline is the dominant factor that can lead to highly nonlinear uncertainty distributions. As these nonlinearities are not captured by the proposed uncertainty computation, its performance will be degraded in difficult cases. Although a short baseline leads to



(a) Rotation error $\boldsymbol{\varpi}$



(b) Translation error $\mathbf{t}$



(c) joint error vector $\boldsymbol{\xi}$

Figure 7.6: The cumulative distribution of the likelihood ratios for the 2-D case including all the covariances built with each particular simulation values and noise profile 4 input covariances.

highly nonlinear error distributions, as the number of points increases these distributions become more conventional, although not as notoriously as with the increase of the baseline. In fact, even with only 2 points, a baseline of 4 meters leads to a median covariance quite similar to the experimental covariance, and conservative to the actual shape of the 99% simulation bound.

In a further simulation, which results are presented in Figure 7.8, the rotation, translation and the shape of the input noise covariance (narrow versus round) were varied for only 2 points with a fixed baseline of 2 meters. The translation is chosen to be either zero or equal to 2 meters in both directions, the rotation is either zero or 180°, and the input covariance is either $\Sigma_b = \Sigma_a = \operatorname{diag}\left(10^{-5}, 9\times10^{-3}, 4\times10^{-5}, 16\times10^{-3}\right)$ (narrow case) or $\Sigma_b = \Sigma_a = \operatorname{diag}\left(10^{-3}, 9\times10^{-3}, 4\times10^{-3}, 16\times10^{-3}\right)$ (round case). It can be observed in this figure that the influence of the rotation and translation is not evidently visible in the results, further supporting that the baseline, the number of points, and the shape and size of the input noise distributions are the most significant parameters in both the results of the optimization problem and the validity of the uncertainty characterization. In fact, these two aspects of the whole problem are closely related: the uncertainty characterization here proposed is not adequate when the

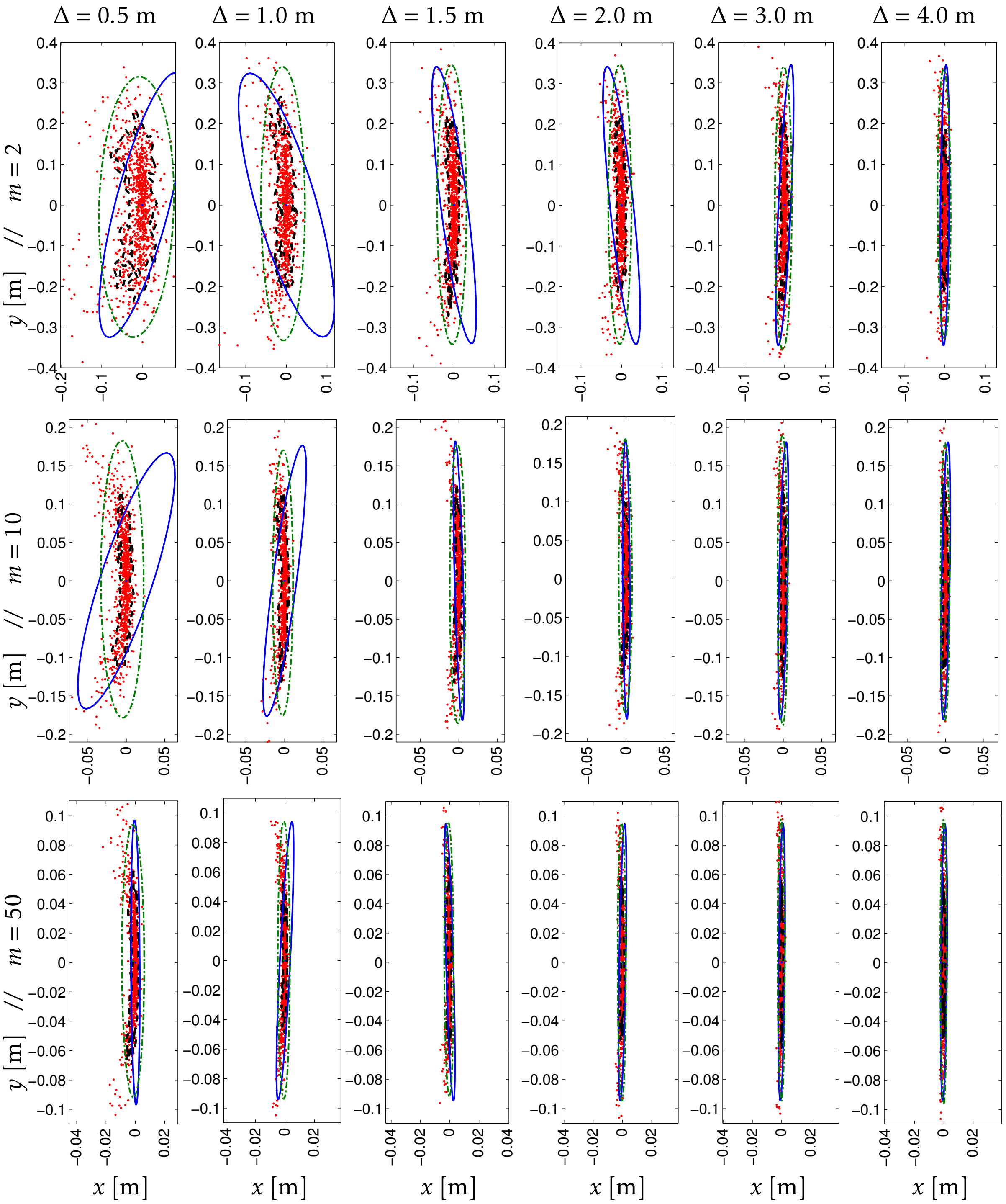
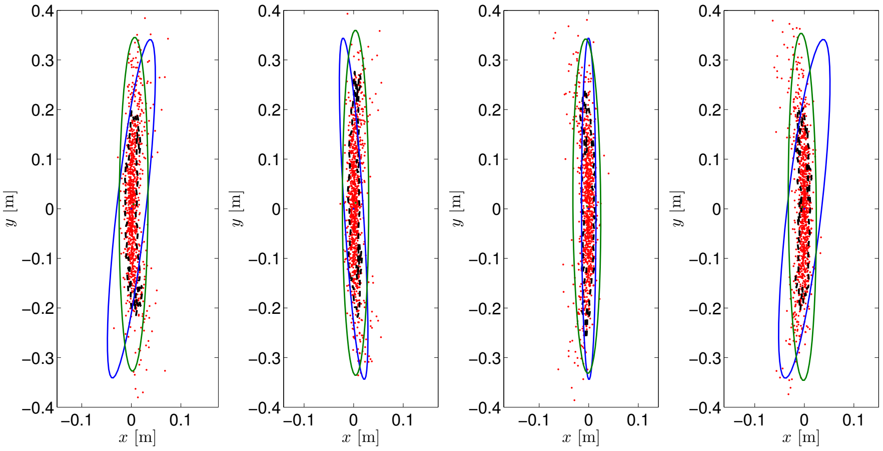


Figure 7.7: Translation error and uncertainty for a series of combinations of baseline-number of points. Translation error Monte Carlo samples in red, 99% bound in dashed black, 3σ simulation covariance ellipse in dashed green, and the 3σ covariance ellipse resulting from the uncertainty characterization with the median likelihood ratio. The baseline ($\Delta = \max \|\mathbf{b}_i - \mathbf{b}_j\|$ for all $i \neq j$, $i, j \in \mathcal{S}$) increases in each figure from left to right, whereas the number of points used in the optimization increases from top to down. Note that the scaling changes from line to line as the number of points increases (ellipses in the bottom are smaller than those at the top).

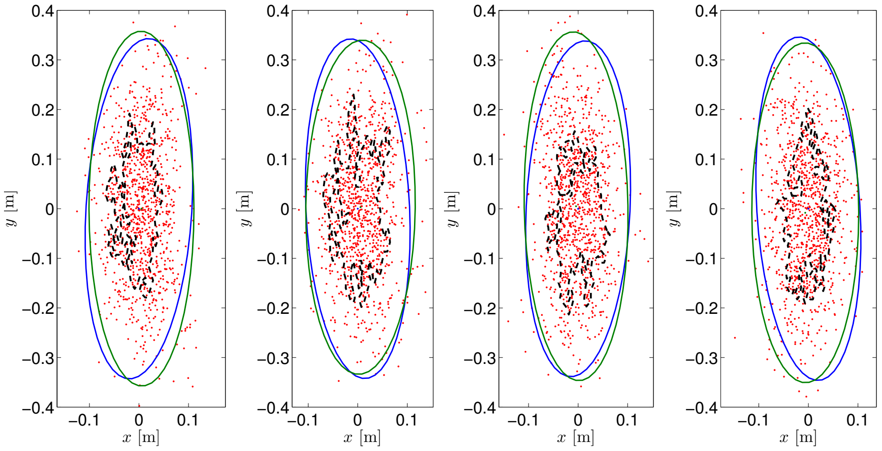


(a) Narrow Σ, $\mathbf{R}(0°)$, $\mathbf{t} = (0, 0)$.

(b) Narrow Σ, $\mathbf{R}(0°)$, $\mathbf{t} = (2, 2)$.

(c) Narrow Σ, $\mathbf{R}(180°)$, $\mathbf{t} = (0, 0)$.

(d) Narrow Σ, $\mathbf{R}(180°)$, $\mathbf{t} = (2, 2)$.



(e) Round Σ, $\mathbf{R}(0°)$, $\mathbf{t} = (0, 0)$.

(f) Round Σ, $\mathbf{R}(0°)$, $\mathbf{t} = (2, 2)$.

(g) Round Σ, $\mathbf{R}(180°)$, $\mathbf{t} = (0, 0)$.

(h) Round Σ, $\mathbf{R}(180°)$, $\mathbf{t} = (2, 2)$.

Figure 7.8: Translation error and uncertainty for a series of combinations of (rotation, translation, covariance shape). Translation error Monte Carlo samples in red, 99% bound in dashed black, 3σ simulation covariance ellipse in dashed green, and the 3σ covariance ellipse resulting from the uncertainty characterization with the median likelihood ratio. The first four figures have narrow covariances, and the last ones have rounder covariances. The translation and rotation combinations are, in each group, respectively, $\{[0, 0]^T, \mathbf{I}\}$, $\{[2, 2]^T, \mathbf{I}\}$, $\{[0, 0]^T, \mathbf{R}(180°)\}$, and$\{[2, 2]^T, \mathbf{R}(180°)\}$.

distribution of the error cannot be sufficiently characterized by its first two moments, mean and covariance, even though it demonstrates great accuracy when the error distribution is more conventional. In sum, narrow shapes of input noise, low baselines, and low number of points lead to non-conventional shapes of error distributions, and bettering any of these three parameters leads to more natural error distributions, especially for the case of the baseline. Even though this places restrictions on the baseline and number of points available for the optimization, it is by no means a very harsh limitation, as in the vast majority of applications it is common practice to use a relatively large number of points and reasonable baselines when compared to noise levels.

7.6 Conclusions

This chapter presented an uncertainty formulation of the weighted orthogonal Procrustes problem for stochastically perturbed n-dimensional point clouds that yields a closed form, computationally efficient, and numerically robust optimization-based algorithm that considers arbitrary rotations and translations, as well as individual weights for each pair of points. In this chapter, the uncertainty involved in the weighted orthogonal Procrustes problem for stochastically perturbed n-dimensional point clouds was studied thoroughly and analytical expressions were derived for the first and second moments of the stochastic outputs, the translation and rotation, as well as cross terms that characterize the anisotropic uncertainty of the problem, not imposing assumptions on the actual rotation and translation. These were obtained after assuming an error model based on perturbation theory for each point in the point clouds, and advancing similar error models both for the translation and the rotation matrix. The resulting uncertainty characterization is anisotropic and does not impose assumptions on the actual rotation and translation nor does it require the points in each set to be independent and identically distributed. This novel uncertainty characterization for the Procrustes problem was validated through extensive Monte Carlo simulations exploring the general framework of the algorithm, using arbitrary rotations and translations, as well as full covariance matrices for each point. For a relatively low number of points, and reasonable noise conditions, the proposed uncertainty characterization performed very well in likelihood ratio tests that encompass a wide variety of configurations varying in translation, rotation, and spatial distribution of points. A thorough analysis of the influence of several parameters on the error distribution and the validity of the results was performed, resulting in the conclusion that with a reasonable baseline the uncertainty characterization performs accurately. The perturbation error models imposed in the input point clouds assume that the error distribution is characterizable using only its first and second moments. These results are consistent with the assumptions, and further validate this approach.

8

EARTH-FIXED TRAJECTORY AND MAP ONLINE ESTIMATION

THIS chapter addresses the problem of obtaining an Earth-fixed trajectory and map, with the associated uncertainty, using the sensor-based map provided by a globally exponentially stable (GES) SLAM filter. The algorithm builds on an optimization problem with a closed-form solution, and its uncertainty description is derived resorting to perturbation theory. The combination of the algorithm proposed in this chapter with sensor-based SLAM filtering results in a complete SLAM methodology, which is directly applied to the three main different formulations: range-and-bearing, range-only, and bearing-only. Simulation and experimental results for all these formulations are included in this work to illustrate the performance of the proposed algorithm under realistic conditions.

8.1 Introduction

The filters presented in Part I represent all landmark positions in the vehicle/sensor coordinate frame, making the positioning of the vehicle in the map trivial by construction (i.e. it is at the origin and aligned with the sensor frame) and consequently, not included in the state vector. The framework of the sensor-based SLAM filter is then completely independent of the inertial frame, as every input and state are expressed in the body-fixed frame. Nevertheless, most SLAM algorithms perform the mapping and localization in an inertial reference frame, as many applications require the inertial (or Earth-fixed) map and the trajectory of the vehicle. For that reason, recovering the pose of the vehicle is important, and it is possible to do by finding the transformation that best relates the Earth-fixed map and the sensor-based map provided by the SLAM filter. This contextualizes the relevance of the work in Chapter 7, since the Procrustes problem studied therein addresses precisely this process.

This chapter formulates an optimization problem with a solution that corresponds to an estimate of the transformation between the body-fixed frame $\{B\}$ and the Earth-fixed reference frame $\{E\}$ in each time instant, allowing the update of the Earth-fixed map through the algorithm here proposed. An error function is defined and then used to construct a cost function for the optimization problem. The algorithm builds on the derivation in Chapter 7 and the uncertainty characterization proposed therein, and aims at estimating, in real time, both the vehicle trajectory and the Earth-fixed map described in the same frame, with the respective uncertainty characterization.

Chapter Structure

The remainder of the chapter is organized as follows. In Section 8.2, the optimization problem that drives the algorithm is formulated and an overview of the Earth-fixed trajectory and map (ETM) algorithm is presented, followed by the uncertainty characterization in Section 8.3. Simulation results using a sensor-based range-only simultaneous localization and mapping (RO-SLAM) filter are presented in Section 8.4, and experimental results of two different implementations of sensor-based filters (range-and-bearing simultaneous localization and mapping (RB-SLAM) and bearing-only simultaneous localization and mapping (BO-SLAM)) using different exteroception sensors are detailed in Section 8.5. Finally, concluding remarks and some directions for the future are presented in Section 8.6.

8.2 Overview

The algorithm proposed in this chapter was designed to complement a sensor-based SLAM filter, as part of the SLAM methodology that provides estimates in both the body-fixed frame and the local inertial/Earth-fixed frame. The fundamental idea is to use the estimated sensor-based landmarks $\mathbf{p}_i$ to formulate an optimization problem with a closed-form solution to enable the real-time estimation of the transformation from $\{B\}$ to $\{E\}$ at each time instant k, which is defined by the position and orientation of the vehicle in $\{E\}$, respectively denoted as ${}^{E}\mathbf{p}_k \in \mathbb{R}^n$ and $\mathbf{R}_k \in \mathrm{SO}(n)$. This can be accomplished in several ways, however, the computation of the resulting transformation uncertainty depends highly on the choice of error that is minimized.

The proposed algorithm resorts to the solution of the orthogonal Procrustes problem presented in Chapter 7, to obtain the transformation between the sensor and Earth-fixed frames at each time instant, thus yielding the vehicle trajectory in the Earth-fixed frame. However, this implies that there is already a known Earth-fixed map of landmarks ${}^{E}\mathbf{p}_i$, $i \in \mathcal{M}$.

Assuming that the Earth-fixed landmarks considered for the ETM algorithm are static, as the vehicle navigates through the environment maintaining visibility with some of the previously visible sensor-based landmarks, it is always possible to estimate the pose of the vehicle, using the Procrustes solution, and to initialize the Earth-fixed map by using the current vehicle pose estimate. Furthermore, it is also possible to update the current estimate of each landmark of the Earth-fixed map. In general, at a given time instant k, if ${}^{E}\hat{\mathbf{p}}_k$ and $\hat{\mathbf{R}}_k$ are known, the update for each Earth-fixed landmark estimate can be computed using

$${}^{E}\hat{\mathbf{p}}_{i_k} = \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} + {}^{E}\hat{\mathbf{p}}_k, \tag{8.1}$$

for all $i \in \{1,\ldots,N_M\}$. This algebraic loop may be averted if it is noticed that, as the landmarks in $\{E\}$ are static, there is also a correspondence between ${}^E\mathbf{p}_{i_{k-1}}$ and ${}^E\mathbf{p}_{i_k}$. This step is of the utmost importance in the design of the algorithm, and yields the following error function

$$ {}^E\mathbf{e}_{i_k} = {}^E\hat{\mathbf{p}}_{i_{k-1}} - \hat{\mathbf{R}}_k \hat{\mathbf{p}}_{i_k} - {}^E\hat{\mathbf{p}}_k, $$

for $k > 0$, that represents the error between the previous estimate of the Earth-fixed landmark and its sensor-based homologous at time k, rotated and translated with the estimated transformation. The pair $\left(\mathbf{R}_k, {}^E\mathbf{p}_k\right)$ can be obtained using the optimization problem

$$ \left(\mathbf{R}_k^*, {}^E\mathbf{p}_k^*\right) = \arg \min_{\substack{\mathbf{R}_k \in \mathrm{SO}(n) \\ {}^E\mathbf{p}_k \in \mathbb{R}^n}} G\left(\mathbf{R}_k, {}^E\mathbf{p}_k\right), \tag{8.2} $$

considering the cost function

$$ G\left(\mathbf{R}_k, {}^E\mathbf{p}_k\right) = \frac{1}{N_T} \sum_{i=1}^{N_T} \sigma_{i_k}^{-2} \|{}^E\mathbf{e}_{i_k}\|^2, $$

where $\sigma_{i_k}^2$ accounts for the uncertainty in each pair $\left({}^E\mathbf{p}_{i_{k-1}}, \mathbf{p}_{i_k}\right)$, and is conservatively defined as

$$ \begin{aligned} \sigma_{i_k}^2 &= \lambda_{max}(\Sigma_{{}^E p_{i_k}}) + \lambda_{max}(\Sigma_{p_{i_k}}) \\ &\geq \lambda_{max}(\Sigma_{{}^E p_{i_k}} + \mathbf{R}_k \Sigma_{p_{i_k}} \mathbf{R}_k^T). \end{aligned} $$

Note that the number of landmark pairs used in the ETM algorithm, N_T, may be different from the available number of landmarks, N, depending on the set of pairs of landmarks used in the optimization problem. It may be beneficial to use only a subset containing, for instance, the most recently visible (or least uncertain) landmarks in the sensor-based frame and their Earth-fixed analogous, provided that the dimension of the resulting sets is greater, if possible, than a predefined threshold. Other possibility is to choose the subset of sensor-based landmarks whose filtered estimates have already converged. Reducing the number of landmarks makes the algorithm more computationally efficient, and the threshold may be imposed to guarantee the numerical robustness and statistical consistency of the algorithm, as shown by the Monte Carlo simulation of Section 7.5. As mentioned before, the optimization problem (8.2) is in fact the orthogonal Procrustes problem (7.3) as solved in [Ume91] and [Goo91], whose uncertainty characterization is proposed in Chapter 7. From the original optimization problem it is known that the optimal translation is given by (7.6), reproduced here for convenience,

$$ \begin{aligned} {}^E\mathbf{p}_k^* &= \frac{1}{\sum_{j=1}^{N_T} \sigma_{j_k}^{-2}} \sum_{i=1}^{N_T} \sigma_{j_k}^{-2} \left({}^E\hat{\mathbf{p}}_{i_{k-1}} - \mathbf{R}_k^* \hat{\mathbf{p}}_{i_k}\right) \\ &= \boldsymbol{\mu}_{E_k} - \mathbf{R}_k^* \boldsymbol{\mu}_{B_k} \end{aligned} \tag{8.3} $$

which is the vector that translates the weighted centroid of the sensor-based landmarks $\boldsymbol{\mu}_{B_k}$ rotated to $\{E\}$ to the weighted centroid of the Earth-fixed landmarks $\boldsymbol{\mu}_{E_k}$. The optimal rotation is given by (7.5), or

$$ \mathbf{R}_k^* = \mathbf{U}_k \operatorname{diag}(1, \cdots, 1, |\mathbf{U}_k||\mathbf{V}_k|)\, \mathbf{V}_k^T, \tag{8.4} $$

where $\mathbf{U}_k$ and $\mathbf{V}_k$ come from the singular value decomposition

$$\mathbf{U}_k\mathbf{D}_k\mathbf{V}_k^T = \operatorname{svd}\sum_{i=1}^{N_T}\sigma_{j_k}^{-2}\left({}^E\hat{\mathbf{p}}_{i_{k-1}} - \boldsymbol{\mu}_{E_k}\right)\left(\hat{\mathbf{p}}_{i_{k-1}} - \boldsymbol{\mu}_{B_k}\right)^T.$$

The Earth-fixed map estimate at instant k is computed using the update equation (8.1), following the computation of the optimal translation and rotation using the sensor-based map estimate of instant k and the Earth-fixed estimate of the previous iteration. The traditional SLAM approach assumes that the transformation between Earth-fixed and sensor-based frames is known at time k_0, i.e., that the initial pose of the vehicle is deterministic. Noting this, and since the sensor-based SLAM filter provides a map in the body-fixed frame at all times, it is possible to compute the landmark map in $\{E\}$ from the sensor-based map in $\{B\}$,

$${}^E\hat{\mathbf{p}}_{i_0} = \mathbf{R}_0\hat{\mathbf{p}}_{i_0} + {}^E\mathbf{p}_0,$$

with

$$\Sigma_{{}^E p_{ij_0}} = \mathbf{R}_0\Sigma_{p_{ij_0}}\mathbf{R}_0^T.$$

Therefore, all the estimated Earth-fixed quantities are computed with respect to the initial pose. In a range-and-bearing formulation, where relative positions of landmarks are readily available on first observation, the use of this strategy is straightforward. However, for its use in conjunction with range-only or bearing-only filters, special care must be taken. Due to the fact that the algorithm uses the first observation as an Earth-fixed estimate, the algorithm will produce better results if initialized when the uncertainty is low enough. Therefore, the first Earth-fixed pose estimate may not correspond to the beginning of the run. However, it is possible to use a GES smoothing filter [AM79, BH97] to improve the initial sensor-based map estimates and then compute the missing trajectory.

The work presented in this section, including the estimates for the vehicle pose, given by (8.3) and (8.4), and the update equation (8.1), allows the real-time computation of the vehicle trajectory and the Earth-fixed map. However, the ETM algorithm here described assumes the knowledge of the uncertainty of both the Earth-fixed and sensor-based landmark estimates. The latter is directly provided by the SLAM filter, but the former is yet to be described. The scope of the next section is to provide uncertainty descriptions of the estimates yielded by this algorithm, using perturbation theory and building on the work in Chapter 7.

8.3 Uncertainty characterization

The solution of the optimization problem (8.2) is based on the weighted orthogonal Procrustes problem. When dealing with sets of points stochastically perturbed as inputs, the result is itself perturbed and therefore it is important to look at the pair rotation-translation that is the solution along with a suitable uncertainty description. In Chapter 7 the underlying uncertainty was thoroughly studied and analytical expressions were derived for the first and second moments of the stochastic outputs, the translation and rotation, as well as cross terms that characterize the anisotropic uncertainty of the problem, not imposing assumptions on the actual rotation and translation. Since the optimization problem underlying the algorithm proposed in this chapter is the same as the one presented in the previous chapter, if similar error models are chosen for the body- and Earth-fixed sets, the uncertainty characterization is also valid. Therefore, consider the following error models for the input landmarks sets

$$\mathbf{p}_{i_k} = \mathbf{p}_{i_k}^{(0)} + \epsilon\, \mathbf{p}_{i_k}^{(1)} + \mathcal{O}(\epsilon^2) \tag{8.5}$$

and

$${}^E\mathbf{p}_{i_k} = {}^E\mathbf{p}_{i_k}^{(0)} + \epsilon\, {}^E\mathbf{p}_{i_k}^{(1)} + \mathcal{O}(\epsilon^2), \tag{8.6}$$

while the translation and rotation follow the models (7.8) and (7.10), reproduced here for convenience,

$${}^E\mathbf{p}_k^* = {}^E\mathbf{p}_k^{(0)} + \epsilon\, {}^E\mathbf{p}_k^{(1)} + \mathcal{O}(\epsilon^2) \tag{8.7}$$

and

$$\mathbf{R}_k^* = \left(\mathbf{I}_n + \epsilon\, \mathbf{S}(\varpi_k) + \mathcal{O}(\epsilon^2)\right)\mathbf{R}^{(0)}, \tag{8.8}$$

where ϵ is the smallness parameter, the notation $\mathcal{O}(\epsilon^m)$ stands for the remaining terms of order m or higher, $(.)^{(0)}$ are the zero order terms or true values, and the first order terms, $(.)^{(1)}$, are assumed to follow a known distribution with zero mean and covariance matrices defined by $\mathbf{\Sigma}_{(.)} = \langle (.)^{(1)} (.)^{(1)^T} \rangle$. The resulting covariances for the rotation error $\varpi_k \in \mathbb{R}^{\frac{n(n-1)}{2}}$ and for ${}^E\mathbf{p}_k^*$ are thoroughly defined in Section 7.4, where they are denoted by Σ_ω (in (7.21)) and Σ_t (in (7.23)). In this chapter, however, they are henceforth presented as Σ_{ϖ_k} and $\Sigma_{{}^E p_k}$.

8.3.1 Earth-fixed map uncertainty

The Earth-fixed map estimation depends on the translation and rotation estimates, as well as on a sensor-based map. Then, it is expected that the associated uncertainty will include terms related to the uncertainty of each one of these components, as well as cross-covariance terms. Recall that the Earth-fixed map estimate is calculated with the update equation (8.1). Using this expression and the error models (8.5), (8.6), (8.7), and (8.8) it is possible to write

$${}^E\mathbf{p}_{i_k}^{(0)} = {}^E\mathbf{p}_k^{(0)} + \mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(0)} \tag{8.9}$$

and

$${}^E\mathbf{p}_{i_k}^{(1)} = {}^E\mathbf{p}_k^{(1)} + \mathbf{S}(\varpi_k)\mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(0)} + \mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(1)}. \tag{8.10}$$

Looking at (8.9), it is confirmed that ${}^E\mathbf{p}_{i_k}^{(0)}$ is the true quantity. Furthermore, from (8.10), it can be seen that ${}^E\mathbf{p}_{i_k}^{(1)}$ has zero mean, since all the quantities that compose it are themselves true quantities or have zero mean. The covariance matrix of the landmark position estimate is given by

$$\begin{aligned}
\Sigma_{{}^E p_{ij_k}} =& \langle {}^E\mathbf{p}_{i_k}^{(1)}\, {}^E\mathbf{p}_{k_j}^{(1)^T} \rangle \\
=& \Sigma_{{}^E p_k} + \mathbf{R}^{(0)} \Sigma_{p_{ij_k}} \mathbf{R}^{(0)^T} + \bar{\mathbf{S}}^T\left(\mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(0)}\right) \Sigma_{\varpi_k} \bar{\mathbf{S}}\left(\mathbf{R}^{(0)}\mathbf{p}_{j_k}^{(0)}\right) \\
& + \langle {}^E\mathbf{p}_k^{(1)} \mathbf{p}_{j_k}^{(1)^T} \rangle \mathbf{R}^{(0)^T} + \mathbf{R}^{(0)} \langle \mathbf{p}_{i_k}^{(1)}\, {}^E\mathbf{p}_k^{(1)^T} \rangle \\
& + \bar{\mathbf{S}}^T\left(\mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(0)}\right) \Sigma_{\varpi_k {}^E p_k} + \Sigma_{\varpi_k {}^E p_k}^T \bar{\mathbf{S}}\left(\mathbf{R}^{(0)}\mathbf{p}_{j_k}^{(0)}\right) \\
& + \bar{\mathbf{S}}^T\left(\mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(0)}\right) \langle \varpi_k \mathbf{p}_{j_k}^{(1)^T} \rangle \mathbf{R}^{(0)^T} + \mathbf{R}^{(0)} \langle \mathbf{p}_{j_k}^{(1)} \varpi_k^T \rangle \bar{\mathbf{S}}\left(\mathbf{R}^{(0)}\mathbf{p}_{j_k}^{(0)}\right),
\end{aligned}$$

where the covariances $\Sigma_{{}^E p_k}$ and Σ_{ϖ_k}, and the cross-covariance $\Sigma_{\varpi_k {}^E p_k}$ are denoted as Σ_t (7.21), Σ_ϖ (7.23), and $\Sigma_{\varpi t}$ (7.26) in Section 7.4. The cross-covariances between the translation and the body-fixed map and between the rotation and the body-fixed map can be computed using $\langle \mathbf{t}^{(1)} \mathbf{r}_j^{(1)^T} \rangle \mathbf{R}^{(0)}$ in (7.27) and $\langle \boldsymbol{\omega}^{(1)} \mathbf{r}_j^{(1)^T} \rangle \mathbf{R}^{(0)}$in (7.25). Finally, the cross-covariances between sensor-based landmarks $\mathbf{p}_{i_k}^{(1)}$ at time k and older Earth-fixed landmarks ${}^E\mathbf{p}_{j_{k-1}}^{(1)}$ for all i and j are assumed to be zero.

It is important to notice that, in this procedure, an inertial landmark is only updated if the associated uncertainty decreases in that iteration. Thus, in each iteration, the covariance matrix for the candidate inertial landmarks is computed, and a measure of the uncertainty in each $\Sigma_{{}^E p_{i_{k+1}}}$ is compared to its previous value. If the uncertainty is raised, then the old covariance is kept and the cross-covariances between an updated landmark and a non-updated one are set to zero, i.e., $\Sigma_{{}^E p_{ij_{k+1}}} = \mathbf{0}$ for all $j \neq i$. There are many possible measures of the covariance, from the volume of the covariance ellipsoid (proportional to the determinant of $\Sigma_{{}^E p_{i_k}}$) to the size of its largest axis (the maximum eigenvalue). In this work, the trace of $\Sigma_{{}^E p_{i_k}}$ is used.

In the event that the attitude of the vehicle is provided by an external source, such as an AHRS [CMC07], then only the Earth-fixed position and map need to be estimated using (8.3) and (8.1). In that case all the cross-terms involving ϖ_k are zero and Σ_{ϖ_k} must be either estimated or provided by the AHRS. Furthermore, if a relative map is obtained in a world-centric framework, as in [JB16], the position and Earth-fixed map can still be found as before, and then all the occurrences of $\mathbf{R}^{(0)}\mathbf{p}_{i_k}^{(\cdot)}$ are substituted by the correspondent quantity in the world-centric relative map and all quantities involving ϖ_k disappear.

8.4 Simulation results: Range-only SLAM

The ETM strategy proposed in this chapter is completely independent of the source of the sensor-based map that serves as input. This means that, as long as it is possible to extract a coherent map using range and/or bearing measurements, the ETM algorithm will be able to match that evolving map to a frame fixed to the initial pose. However, depending on the source there are some particularities that influence the result. For example, in range-and-bearing or

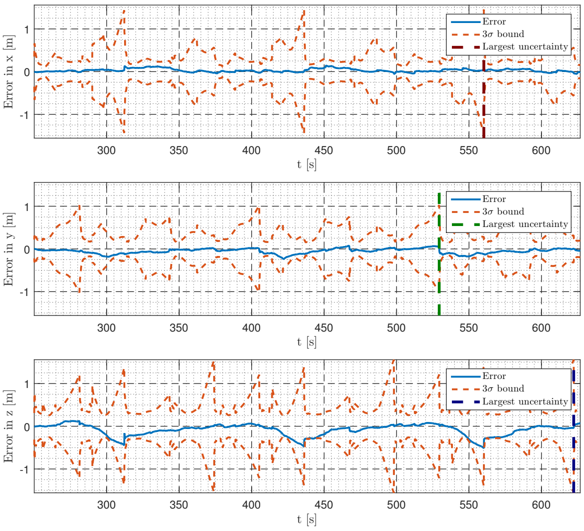


Figure 8.1: RO-SLAM: Evolution of the vehicle position estimation error with uncertainty bounds.

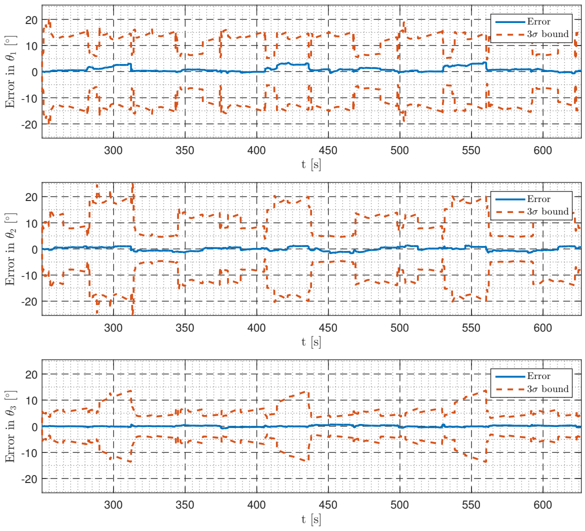


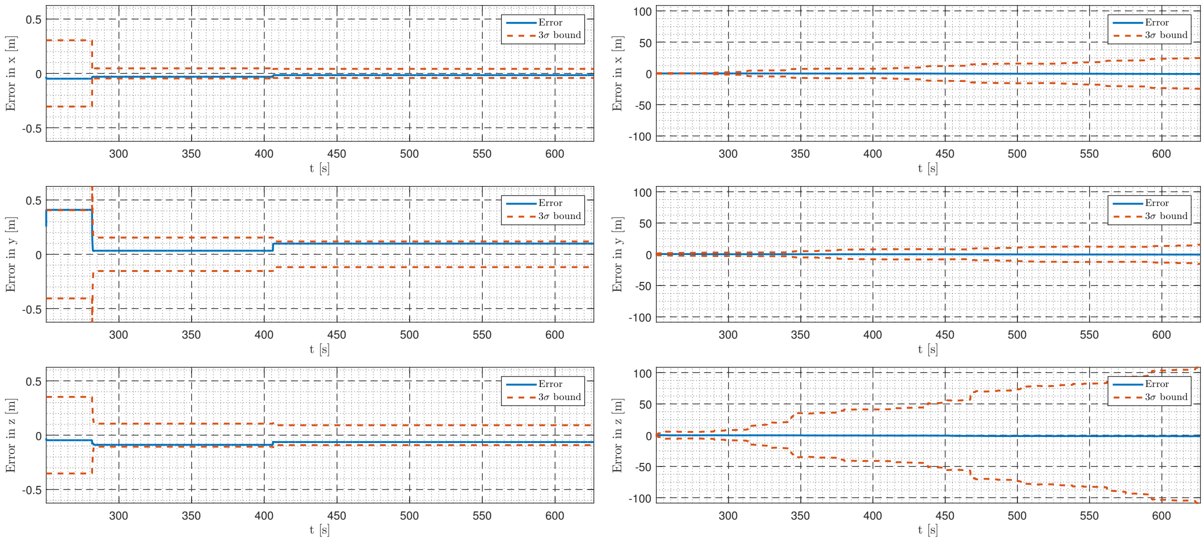
Figure 8.2: RO-SLAM: Evolution of the vehicle attitude estimation error with uncertainty bounds.

bearing-only procedures where landmark association is not perfect, the Earth-fixed pose and map obtained may be impaired by erroneous associations.

To show a situation where this does not occur, consider the sensor-based range-only SLAM filter proposed by the authors in Chapter 4. This filter uses distances to landmarks and measurements of linear and angular velocities, yielding as output the body-fixed map (along with other auxiliary quantities), noting that the association of landmarks is known for range-only frameworks, since the acoustic/electromagnetic signals are usually tagged. The simulations herein reported were devised to better explore several aspects of the algorithm and showcase its performance. The environment, as detailed in previous chapters and in Appendix C, tries to emulate the fifth floor of the North Tower at IST. It consists of a 16 by 16 by 3 m corridor, where 36 landmarks were put in notable places such as corners and doors, with random heights. The aerial vehicle is initialized on the ground and, after take-off, makes several laps around the corridor while trying to guarantee observability of the map with its trajectory (see

Table 8.1: Error norms for the RO-SLAM simulation results (mean±standard deviation)

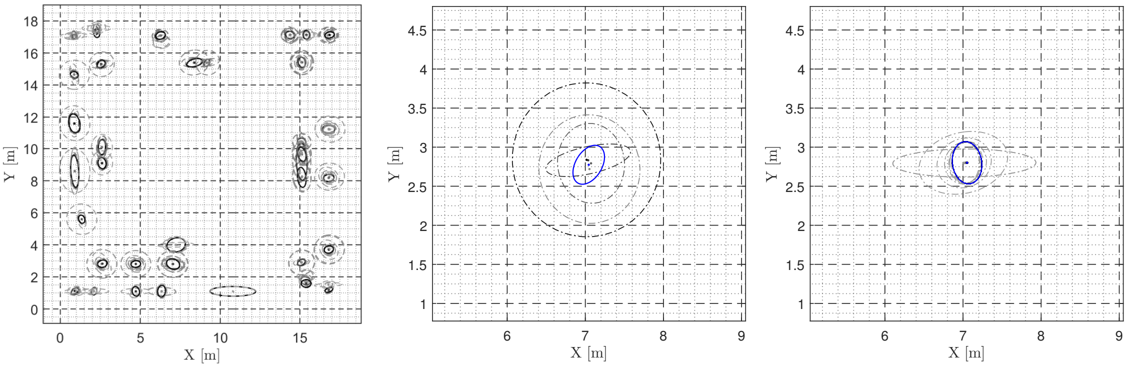
Algorithm	Position [m]	Attitude [deg]	1 Landmark [m]	Map [m]
Sensor-based RO-SLAM	-	-	0.296±0.457	1.145±6.172
ETM with update on trace decrease	0.142±0.115	1.273±0.806	0.049±0.012	0.171±0.137
ETM with update on determinant decrease	0.133±0.091	1.183±0.619	0.058±0.001	0.170±0.132
ETM with update on largest eigenvalue decrease	0.151±0.115	1.211±0.844	0.053±0.009	0.179±0.141
ETM with no restrictions on updates	0.511±0.334	4.679±2.549	0.563±0.368	0.583±0.391



(a) With covariance trace as decision factor.

(b) Without restrictions.

Figure 8.3: RO-SLAM: One particular landmark position error and uncertainty, with two different runs of the algorithm.



(a) Full map.

(b) Detail of one sensor-based landmark.

(c) Detail of one Earth-fixed landmark.

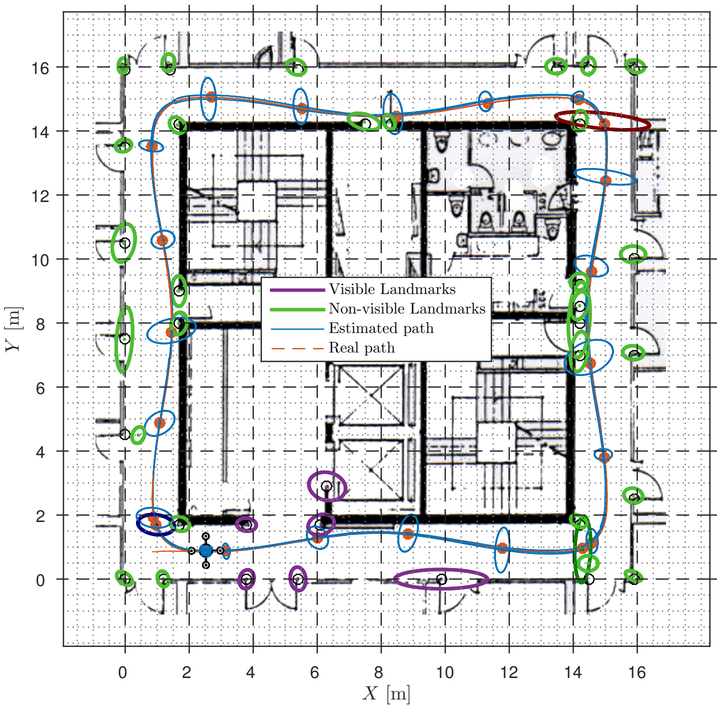
Figure 8.4: RO-SLAM: The spatial evolution of the map covariances. On the left, overlapped snapshots of the Earth-fixed map taken every second. On the center, details of the evolution of the input sensor-based covariance for one of the landmarks in the map, taken every 100 seconds and rotated to the Earth-fixed frame. On the right, the corresponding output Earth-fixed covariance every second, respectively. In this figure, the darker the colour of the ellipse the later in the run it appears.

Chapter 4 for further details). The total distance travelled is 294 meters in 627 seconds, at an average speed of 0.469 m/s, while one single loop of 58 m along the corridor takes 124 s. All the measurements are assumed to be perturbed by zero-mean Gaussian white noise, with standard deviations of $\sigma_\omega = 0.05$ °/s for the angular rates, $\sigma_v = 0.03$ m/s for the linear velocity, and $\sigma_r = 0.03$ m for the ranges. Taking advantage of the simulated environment, several runs are performed with different decision factors for the update step of the algorithm: the trace of the covariance of the individual landmark, the determinant, the largest eigenvalue (all with comparable accuracy in the resulting estimates). An additional run with updates every time-step is also provided. Statistics for the results of estimating position, attitude and maps for these four variants are presented in Table 8.1. There the mean and standard deviations of the error norms are presented for the three products of the algorithm: position, attitude, and map. The statistics of a single landmark (see Figure 8.3) are also shown in the *"1 Landmark"* column. To allow a better understanding of the performance of the ETM algorithm, the results for the sensor-based quantities fed to it are also presented: average landmark error norm and its standard deviation for the same single landmark and the whole map.

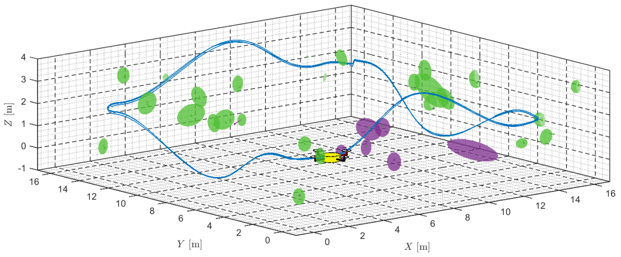
Figure 8.1 shows the evolution of the vehicle position estimation error for a typical simulation in solid blue and the accompanying 3σ uncertainty bound. It can be seen that the error is kept at a low level and that the algorithm manages to keep consistent estimates throughout the run. This assertion is confirmed by Figure 8.2, which shows the attitude estimation error in solid blue and the 3σ uncertainty bounds. This error is computed using the exponential map, $\mathbf{e}_\theta = \mathbf{S}^{-1}\Big(\mathrm{logm}(\hat{\mathbf{R}}\mathbf{R}^T)\Big)$, and, as the position error, it is always low and inside the confidence interval provided by the uncertainty characterization. Note that the uncertainty characterization depends solely on the sensor-based map and the vehicle attitude. Since the trajectory contains periodic repetitions of the pose of the vehicle, and consequent loop closings occur constantly throughout the run, the computed uncertainty shows also a periodic nature. However, there are clear peaks in the position uncertainty, identified for each direction with a vertical dashed line. The poses corresponding to these peaks are depicted in Figure 8.5(a) following the same colour code.

To understand the impact of restricting updates to uncertainty decreases, consider Figure 8.3. On the left side (Figure 8.3(a)), the evolution of the error and corresponding uncertainty when the decision to update is based on the decrease of the trace of the covariance, and on the right side (Figure 8.3(b)) the results for unrestricted updates. It can be seen that not requiring an uncertainty decrease renders the estimates unstable since the uncertainty has a rising trend with local decreases. This is confirmed in the average results shown in Table 8.1, where the errors are considerably larger and with higher variance than in the restricted versions. On the other hand, using the covariance trace as the factor provides a fast convergence in all directions, as do the other restricted versions. Furthermore, the overall map error throughout the run is decreased from the sensor-based estimates to the Earth-fixed ones, which is an interesting and expected effect. Indeed, the inertial model is exact, as the landmarks remain constant, which allows for its uncertainty to decrease over time as more measurements are obtained.

To provide a general overview of the results of the ETM algorithm, Figure 8.5 shows the estimated trajectory and map at the end of the run. The small quadrotor indicates the pose at the current moment, the blue line is the estimated path, the red dashed line is the ground truth, the purple (green) ellipses/ellipsoids are the landmarks visible (occluded) at the current time. The ellipses along the path represent the 3σ bounds for the pose estimates at that point. The black circumferences and red circles are the real landmark and vehicle positions, respectively. These figures show that the map and the estimated trajectory are quite accurate and the



(a) Top-view of the map.



(b) 3-D map.

Figure 8.5: RO-SLAM: Earth-fixed map and trajectory with uncertainty bounds for both landmarks and equally time-spaced poses. Top-view of the map with the floor blueprint in the background.

uncertainty characterization is consistent with the real level of error displayed.

Finally, Figure 8.4 shows the spatial evolution of the uncertainty of the whole Earth-fixed map and a detail comparing the varying uncertainty of a sensor-based landmark, which rotates, increases and decreases throughout the run, and the corresponding uncertainty of the Earth-fixed estimate. The ellipses from covariances estimated by the ETM procedure (Figures 8.4(a) and 8.4(c)) are taken every second and help visualize the convergence of the uncertainty and the shapes it assumes throughout time. To better understand the behaviour of the procedure, the uncertainty ellipses for the same sensor-based landmark are also presented in Figure 8.4(b), where the various ellipses were obtained with intervals of 100 seconds and transformed to the Earth-fixed frame for a fair comparison. In all of these sub-figures of Figure 8.4, the darker the colour of the ellipse the later in the run it appears.

8.5 Experimental results

This section provides experimental results for the performance and consistency evaluation of the ETM algorithm coupled with the sensor-based SLAM filters in Chapters 3 and 6. Since the ETM algorithm works in cascade after the sensor-based filter, its performance is obviously dependent on the individual performance of the latter. The separation of the estimation process in these two stages aims at providing a less uncertain and consistent Earth-fixed trajectory and landmark map. This assertion is grounded on the notion of performing control, decision, and loop-closing procedures in the sensor-based frame, lessening possible effects of nonlinearities in the filter. However, the final Earth-fixed result is still affected by the nonlinearity inherent to the problem of arbitrarily transforming maps between coordinate frames, which is common in EKF-based SLAM algorithms. It should be noted, that, while the uncertainty characterization is approximate, it is usually possible to have always more than 10 landmark pairs in ETM iterations, hence guaranteeing that the uncertainty characterization is valid, as shown by the extensive tests in Section 7.5. In summary, the results reported in this section serve as practical validation of the approach proposed in this thesis as a whole and, while still an indication of the performance of its integral parts, these have to be viewed together. Those that include position and attitude ground truth are summarized in Table 8.2, and Figure 8.6 provides a graphic overview of each experimental setup. An in-depth report is provided in the remainder of the section.

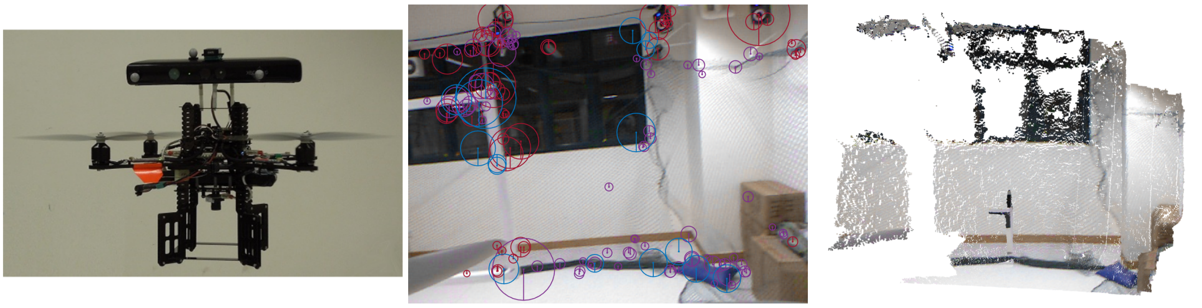
Table 8.2: Error norms for the experimental results (mean±standard deviation)

Algorithm	Position [m]	Attitude [deg]
RB-SLAM (Run #1)	0.097±0.076	3.695±3.673
BO-SLAM	1.220±0.494	0.429±1.074

8.5.1 Range-and-bearing SLAM

The first of the experiments in this chapter combines the sensor-based SLAM filter for range-and-bearing measurements of Chapter 3 with the ETM algorithm. Two runs are detailed: (i) a 20 m long 3-D experiment using an RGB-D camera with ground truth; and (ii) an 80 m long 3-D experiment using an RGB-D camera with partial ground truth (both detailed in Chapter 3). As explained in Chapter 2, the filters output the map, the linear velocity and the rate-gyro biases, all expressed in body-fixed coordinates.

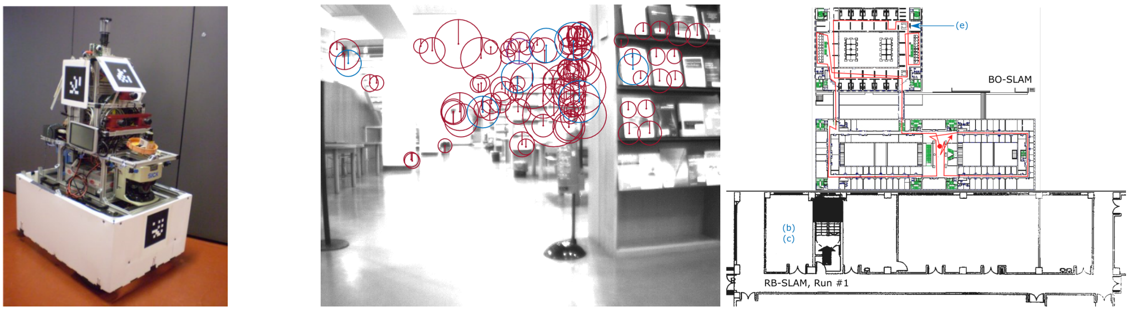
These results come from a series of experiments conducted in the Sensor-based Cooperative Robotics Research Laboratory - SCORE Lab of the Faculty of Science and Technology of the University of Macau. The setup consists of an *Asctec Pelican* quadrotor equipped with a *Microstrain 3DM-GX3-25* inertial measurement unit working at 50 Hz and a *Microsoft Kinect* camera, at 10 Hz. Ground truth is available using of a *VICON* motion capture system, which



(a) RB-SLAM: Run 3-D – Instrumented *AscTec Pelican* quadrotor for data acquisition.

(b) RB-SLAM: Run 3-D – *Kinect* RGB picture with detected features.

(c) RB-SLAM: Run 3-D – *Kinect* pointcloud corresponding to Figure 8.6(b).



(d) BO-SLAM: Instrumented ground robot used in the experimentis in the *Rawseeds* dataset ([BBF$^+$06] and [CFG$^+$09]).

(e) BO-SLAM: Picture of a typical space with detected features.

(f) The blueprints of the regions covered in all trials.

Figure 8.6: Information on the experimental setup of each trial described in this section. The top column depicts RB-SLAM and the bottom one depicts BO-SLAM. On the left, the data acquisition platforms for each algorithm. On the bottom right, the locations of the acquisition of each picture/LiDAR scan.

provides accurate estimates of the position, attitude, linear and angular velocities of any vehicle placed inside the working area with the correct markers. Technical details on the actual implementation can be found in Chapter 3.

In the first experiment, depicted on Figures 8.7-8.9, the vehicle does not leave the area covered by the *VICON* system and, as such, ground truth is always available. The second, longer, experiment consists of a small lap inside the lab followed by a larger exploration of the outside corridor as shown in Figure 8.10. In both runs, the vehicle is hand-driven at an average speed of around 0.4 m/s. Figures 8.7 and 8.8 depict the position and orientation estimates against the ground truth and 95% uncertainty bounds for the first run. It can be seen that the estimates are rather close to the ground truth, and, although there are some moments where the algorithm is somewhat optimistic, the overall performance shows consistency. However, the vertical performance is worse than the horizontal one, which is quite accurate. This is further apparent in Figure 8.8, where the pitch and roll are shown to be estimated with higher relative error. This is most likely due to the lack of vertical motion by the vehicle, and the reduced angle-of-view of the *Kinect* camera which limits the vertical separation of landmarks (smaller baseline). Both aspects contribute to fewer information to extract from the measurements and

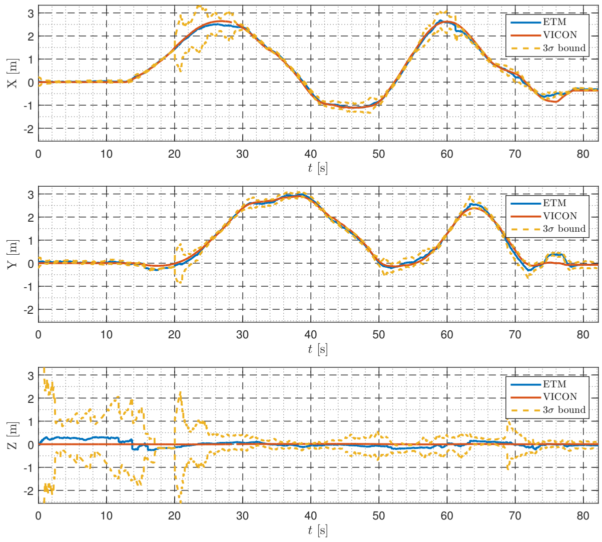


Figure 8.7: RB-SLAM: Run #1 – Evolution of the position estimates with ground truth and uncertainty bounds. Horizontal trajectory (top two figures), Vertical trajectory (bottom).

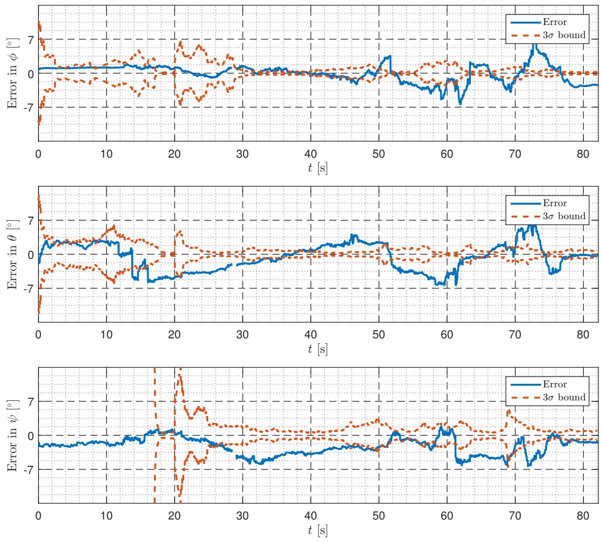


Figure 8.8: RB-SLAM: Run #1 – Evolution of the estimation error of the Euler angles with uncertainty bounds. From top to bottom: roll, pitch, and yaw.

consequently worsen the estimation performance. Furthermore, there exists inconsistency in the attitude, which can be explained by the intrinsic nonlinearity of the problem, made more extreme by erroneous associations and movements with large rotations of the camera. Note that in the second run (Figure 8.12) this effect is much less noticeable. The ending result of the run is shown in Figure 8.9 where a top-view of the Earth-fixed map is depicted along with the estimated trajectory and ground truth, demonstrating the good performance of the overall algorithm.

Figures 8.10, 8.11, and 8.12 depict the results of the second, larger, run. In this run, ground truth is only available initially. In Figure 8.10, a top view of the Earth-fixed map along the estimated trajectory (solid blue), shows the floor blueprint which allows a qualitative valida-

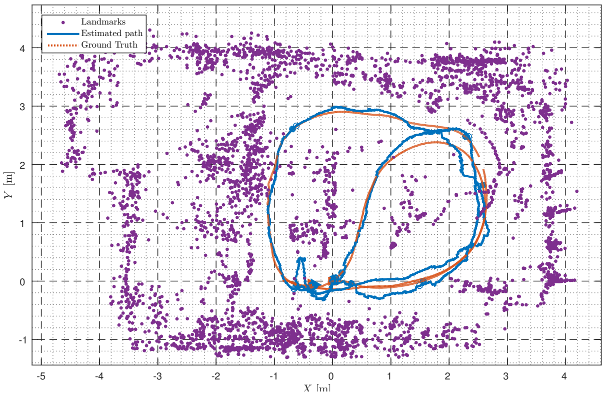


Figure 8.9: RB-SLAM: Run #1 – Top-view of the Earth-fixed map and trajectory with ground truth and uncertainty bounds.

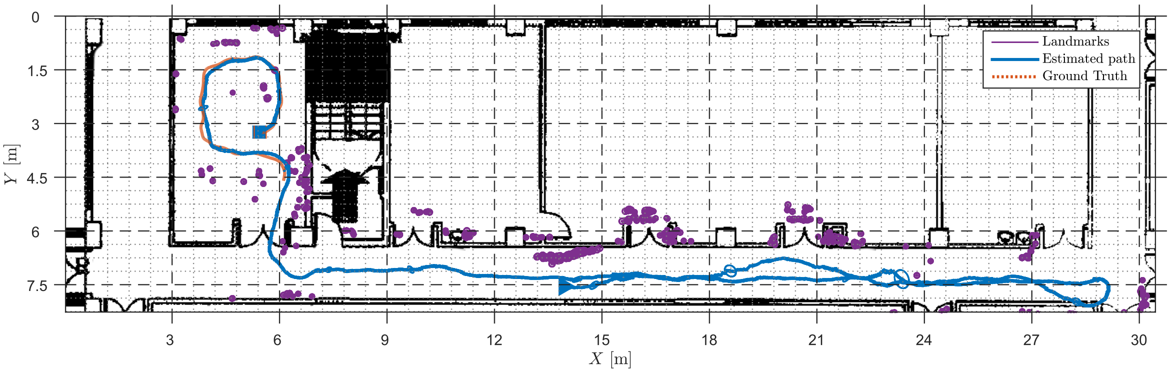


Figure 8.10: RB-SLAM: Run #2 – Top-view of the Earth-fixed map and trajectory with ground truth, uncertainty bounds, and floor blueprint.

tion. In that figure, the ground truth trajectory (dashed red) obtained from the *VICON* is also shown. The trajectory starts inside the lab and ends in the middle of the corridor, marked by the blue triangle. The small ellipses are the 2-D projection of the 3σ uncertainty ellipsoids. The landmark map can be clearly related to objects in the blueprint which once more indicates the good performance of the algorithm.

The vehicle has full 3-D motion, as the attitude in Figure 8.12 and the vertical position in the bottom of Figure 8.11 reveal. In the latter, the Euler angles of the vehicle are depicted along with ground truth, when available, and the 3σ uncertainty bounds provided by the uncertainty description derived in this chapter. As for the position estimates, once again the horizontal performance is better, as there is a drift in the vertical estimate that was not present in reality. It can be seen that the algorithm performs very well while the ground truth is available, with low and consistent uncertainty.

8.5.2 Bearing-only SLAM

The final part of this experimental section reports the use of the ETM algorithm in conjunction with another sensor-based SLAM filter, this time tailored for bearing-only measurements aided by linear and angular velocity measurements. The filter is proposed in Chapter 6 and is tested

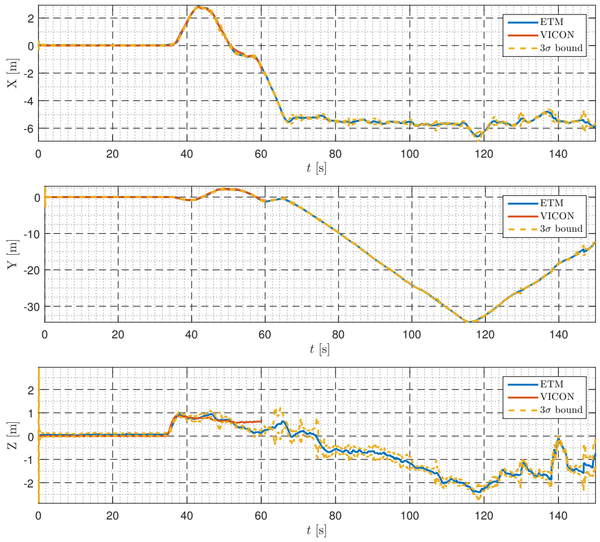


Figure 8.11: RB-SLAM: Run #2 – Evolution of the position estimates with ground truth (available during the first 60 seconds) and uncertainty bounds. Horizontal trajectory (top two figures), Vertical trajectory (bottom).

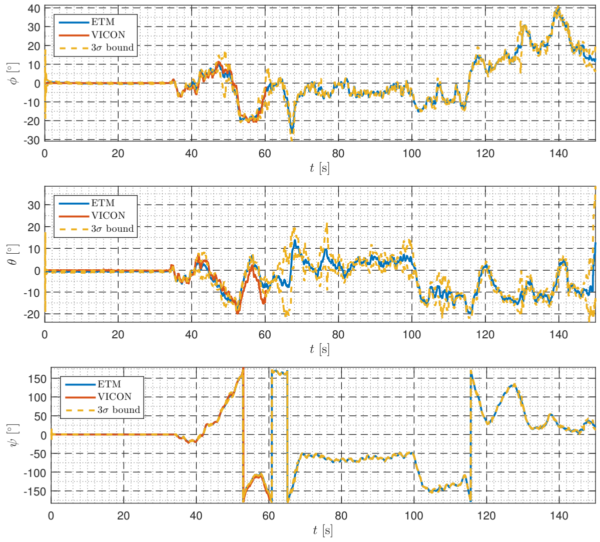
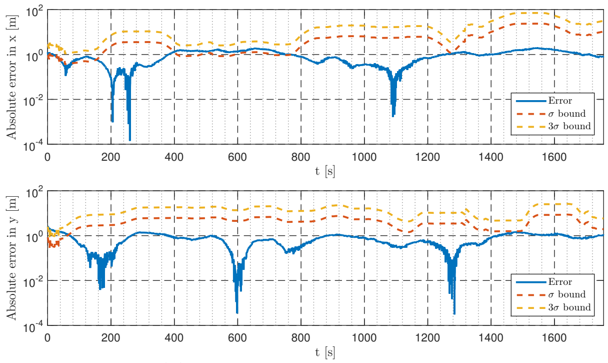
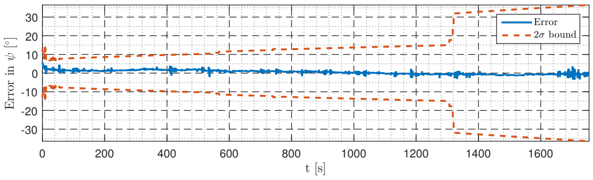


Figure 8.12: RB-SLAM: Run #2 – Evolution of the Euler angle estimates with uncertainty bounds against the ground truth (available during the first 60 seconds). From top to bottom: roll, pitch, and yaw.

with real data from datasets acquired by the *Rawseeds* Project (see [BBF$^+$06] and [CFG$^+$09]). A black and white camera is used to survey the environment and, as in the 3-D RB-SLAM experiments above, a SURF [BETG08] implementation allows the extraction of natural features from the obtained images. The remaining data sources are the wheel encoders which provide body-fixed linear velocity and simulated angular velocity (since the measured angular velocity is largely corrupted by noise, as explained in Chapter 6 and demonstrated by the dataset benchmark results). More details of the experimental setup are given in Chapter 6. The travelled distance in this run is 774 m in 29 minutes, with an average speed of 0.4 m/s, and, since the trajectory is purely two-dimensional, only horizontal results are shown here.



(a) Absolute position error.



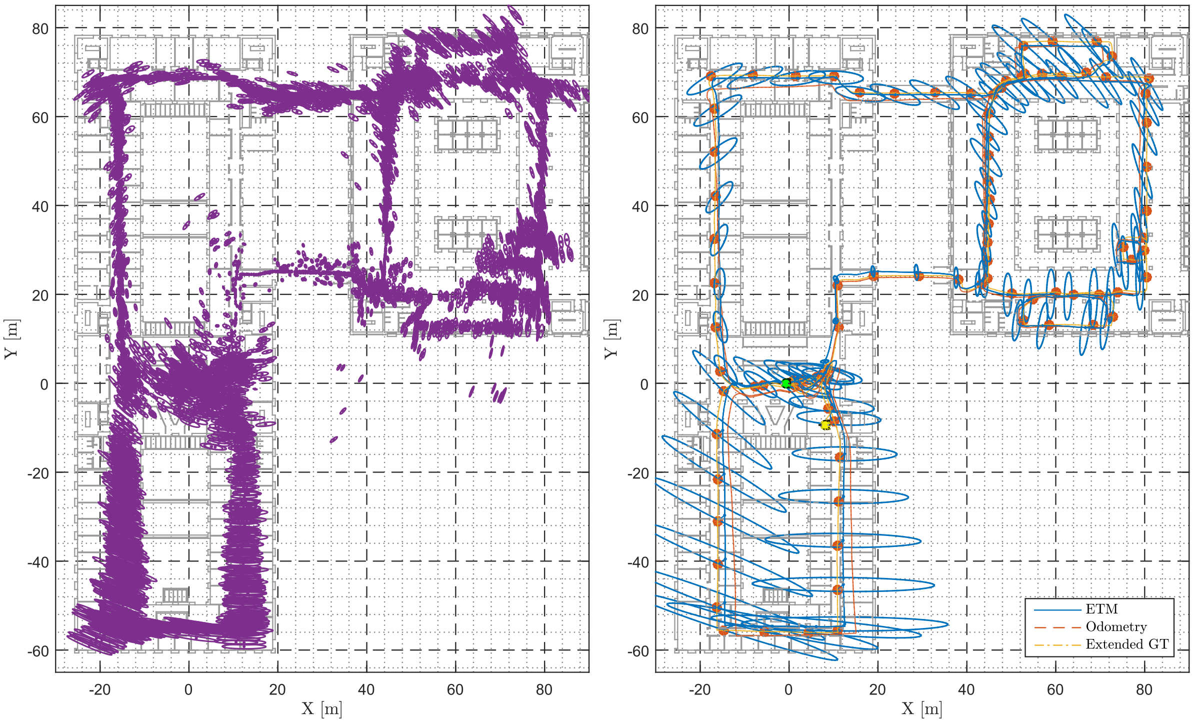
(b) Yaw error.

Figure 8.13: BO-SLAM: Evolution of the estimation error of the vehicle pose. On top, the two horizontal coordinates of the position estimation error, and on the bottom, the attitude estimation error in the form of the yaw error.

Figures 8.13 and 8.14 depict, respectively, the estimation error of the pose of the vehicle and the estimated map and trajectory. In Figure 8.13(a), the absolute estimation error (in solid blue) is accompanied by the uncertainty bounds at σ level (dashed red) and 3σ (dashed yellow). It can be seen that the uncertainty characterization is, in some cases, very conservative, but the consistency is maintained throughout the long run. Furthermore, the error is low, with its norm averaging at 1.2 metres. The attitude in this problem is reduced to the yaw angle, whose estimation error is presented in Figure 8.13(b) in blue with the uncertainty bounds, clearly demonstrating consistency, even if conservative, and a low level of error (average norm of 0.4°).

Finally, in Figure 8.14, the estimated Earth-fixed map and the vehicle trajectory are depicted, along with the ground truth for the latter and the executive drawings of the area. To avoid overcrowding the figures with too much information, the landmark map and corresponding uncertainty ellipses are presented alone in Figure 8.14(a) and the trajectory, ground truth, dead-reckoned odometry, and uncertainty ellipses at fixed intervals are shown in Figure 8.14(b). Even though there is no ground truth for the estimated map, since the landmarks are obtained through feature detection, the executive drawings allow a good qualitative evaluation of the result. As to the trajectory, it can be observed that the complete SLAM algorithm corrects the odometry with low error, even though the last part of the map has much higher uncertainty than the rest. It should be noted that in the sensor-based BO-SLAM filter no dedicated loop closure procedures are used, which may explain this effect.

These experiments were designed to practically validate the optimization-based Earth-fixed Trajectory and Map estimation algorithm presented in this chapter, as part of an integrated



(a) Landmark map with uncertainty ellipses.

(b) Trajectory with ground truth and uncertainty.

Figure 8.14: BO-SLAM: The two-dimensional Earth-fixed trajectory and map, with ground truth, executive drawings, and uncertainty bounds.

two-step SLAM filter, with the demonstration of consistency and good performance of the proposed uncertainty characterization. Furthermore, they allow the performance assessment of the ETM algorithm under real-world conditions, using the proposed formulation of the Procrustes problem and implementing it in each time-step to address a relevant robotics issue, SLAM.

8.6 Conclusions

This chapter proposed an optimization-based algorithm that is part of a novel methodology for simultaneous localization and mapping. The algorithm, fully characterized uncertainty-wise, provides estimates of the landmark map and of the attitude and position of the vehicle in an Earth-fixed frame, using only the body-fixed map provided by existing globally convergent filters. Building on the body-fixed map provided by the sensor-based filter, the problem of obtaining the Earth-fixed trajectory and map was formulated using an orthogonal Procrustes problem approach. The resulting optimization problem has a closed-form solution and a statistical description of the obtained Earth-fixed map is also proposed building on the work presented in Chapter 7. Furthermore, the performance and consistency of the algorithm were validated in simulation for a range-only SLAM formulation and experimentally for range-and-bearing and bearing-only formulations. These results, with ground truth data, showed also the good performance of the various SLAM algorithms detailed throughout this thesis as a whole.

9

CONCLUSIONS AND FUTURE WORK

This dissertation presented novel algorithms for simultaneous localization and mapping in various formulations, reporting the design, analysis, implementation and experimental validation stages. During the course of the thesis, concluding remarks were provided for each chapter. This chapter summarizes the conclusions pertaining each of the chapters and provides a global perspective.

9.1 Conclusions

The general objective of the thesis was to find previously absent global convergence guarantees for this nonlinear problem, regardless of the formulation chosen. This was achieved successfully for the first time in 3-D. The strategy proposed was to divide the problem in two, noting that, since all the measurements associated with SLAM are relative to the vehicle, removing the unobservable pose from the main filter was beneficial. Following this, a class of systems rooted in sensor-based considerations was designed, and a strategy to convert sensor-based to Earth-fixed maps by computation of the vehicle pose was proposed, its statistical coherence validated and its experimental performance evaluated.

In the introduction of this dissertation, several relevant effects of global convergence in SLAM algorithms were enumerated. One of these was the inherent solution to the kidnapped robot problem. It was argued that a kidnapping or a failure with similar effect could simply be seen as a wrong initial guess that would, after a transient period, converge to the true value as long as the conditions for convergence were met – from the point of view of the Kalman filter an instantaneous failure like this poses the same problem as a highly erroneous initialization. However, both of these problems, i.e., initialization and failures, highly depend on how accurate the landmark association process is. In the naïve implementations of Chapters 3 and 6, chosen for their ease of implementation and computational cost but that can be replaced by any other more involved strategy, the association depends on the previous estimate and the search is restricted to the predicted field-of-view of the vehicle. This is not a problem for initialization or tracking, since the prediction stage accounts for inaccuracies in the prediction model. However, in the case of a kidnapping, it is likely that the association process would not recognize that features in the new surroundings were already in the filter. In a situation where association only depends on information obtained with the measurement (tagged radio/acoustic signals, unique or highly reliable visual marks) then the algorithm will take care of the transient process to converge back to the true position and recover from the failure. The algorithms of Chapter 4 and 5 fit in this description. This caveat does not lower the relevance of having a global exponentially stable algorithm, since this property ensures without question that the algorithm will perform as intended. Furthermore, it provides a natural solution to the initialization problem, something that has fuelled many research efforts in RO-SLAM and BO-SLAM. This argument does, however, point towards the relevance of implementing informed strategies for association.

In Chapter 2, an overarching methodology to tackle the design of globally convergent simultaneous localization and mapping filters was introduced. This was inspired in the previous two-dimensional range-and-bearing sensor-based filter in [GBSO13], and the source-localization algorithms in [BSO11b] and [BSO13b]. This chapter proposed a new class of sensor-based filters for simultaneous localization and mapping, established the necessary results and provided the foundations over which the two parts of the thesis were built.

In the first part of the thesis there was a common thread uniting every chapter, the quest for filters with globally exponentially stable error dynamics for the main SLAM problems. With that purpose in mind, Chapter 3 details the first application to 3-D of the methodology explained in Chapter 2. A nonlinear system was designed to encompass range-and-bearing SLAM capabilities, whereas a small transformation of the dynamics allowed the automatic calibration of the rate-gyros by incorporating the angular velocity measurement bias on the filter state. While this system is indeed nonlinear, it can be considered as LTV for observability purposes. The (uniform complete) observability of this new system was analysed, and its results related back to the original nonlinear system. This analysis was constructive, in the sense that it enabled the use of a Kalman filter for linear time-varying systems with globally exponentially stable error dynamics. It was shown that the convergence properties of the filter will hold whenever three or more non-collinear landmarks are visible at the same instant, which is a very common occurrence. This, along with the fact that it does not require linear velocity measurements, makes range-and-bearing SLAM the most versatile of the formulations addressed in this thesis. It is, however, more computationally demanding than, for example, the RO-SLAM filter. The details for a practical implementation of the filter were presented, paving the way to real experiments performed in the University of Macau with a quadrotor instrumented with a *Microsoft Kinect* RGB-D camera and a *Microstrain 3DM-GX3-25* IMU. The experimental results showed the good performance of the filter in a real environment and contribute to the validation of both the pure sensor-based filter and an early application of the Earth-fixed methodology without uncertainty characterization.

In Chapter 4, the methodology that permeates Part I of the thesis was employed towards the successful design of a novel GES filter for RO-SLAM. The basis of this work was a nonlinear dynamic system containing the map and the linear velocity of the vehicle, which is also measured. The angular velocity was assumed to be calibrated prior to its use in the filter. Following the ideas advanced in Chapter 2, this system was augmented and the measured quantities were directly employed in the dynamics to pave the way to the design of an LTV system. As in the previous chapter, a constructive observability analysis of this system was performed, showing that it did mimic the original nonlinear system, which validated this approach of tackling nonlinearities without any linearizations whatsoever. The conditions resulting from this analysis are the dual of the well-known conditions for range-based localization: each landmark must be observed from four non-coplanar points, or, in other words, the vehicle must move in all the directions of the mapped space. Even though range-only SLAM has the advantage of avoiding computationally heavy and error-prone vision systems, it is the formulation which imposes the strongest restrictions on the overall motion of the vehicle (not counting its extension to moving object tracking). The mentioned analysis was once again the stepping stone for the design of a Kalman filter for the LTV system with globally exponentially stable (GES) error dynamics. A discrete version of this filter was implemented in practice, and simulation results were provided that allowed to understand the inner workings of the purely sensor-based filter. These were then expanded upon with experiments in the University of Macau, where a constellation of *Crossbow Cricket* beacons was installed and a quadrotor instrumented with a receiver, an IMU and a down-facing *Microsoft Kinect* camera. While the performance of the algorithm in these experiments was adequate, the global convergence influence was clear in the initialization, as well as the need to have trajectories that consistently cover all the dimensions.

The work of Chapter 4 was extended to the field of sensor networks in Chapter 5. An long baseline (LBL)-like framework was blended with the RO-SLAM filter and adapted to the problem of simultaneous localization and mapping for sensor networks with moving object tracking. This was based upon the range-only systems derived in the previous chapter, inter-beacon communication and an assumed constant velocity motion model for the moving object. The systematic methodology of the thesis is applied to the resulting systems, once again moving from a nonlinear system to an LTV system that is shown to mimic the former. Unsurprisingly, the observability analysis provided the most demanding conditions on the motion of the vehicle yet – the algorithm must now estimate another three-dimensional quantity from the same distance information. The end result of the chapter is a globally convergent filter that localizes a sensor network including a moving node with respect to the main vehicle. Simulation results were presented to validate and assess the concept. Even though the constant Earth-fixed velocity model is somewhat restrictive, simulations where the target followed trajectories that do not comply with the model showed that the stochastic setting of the filter accommodates those differences rather well.

To close the first part of the thesis, in Chapter 6 the bearing-only formulation of SLAM was pursued. As in the range-only formulation, the unbiased angular velocity was assumed to be available along with the linear velocity of the vehicle. With this in mind, a nonlinear system considering direction measurements to landmarks was designed. An output transformation inspired by a similar one in [BSO13b] that was not used elsewhere in the thesis and the state augmentation used in Chapter 4 enabled the design of a new LTV system. Following the common methodology of this part of the thesis, the two systems were related through a thorough observability analysis that was constructive in the sense that it lead to a Kalman filter with the convergence properties sought for in the whole thesis. This convergence and observability results depend on the motion of the vehicle relative to the landmarks, as they are clearly related to the triangulation problem in the sense that a landmark becomes observable when it is observed from two different viewpoints. This shows that the bearing-only formulation is less demanding than its range-only counterpart in terms of acceptable scenarios, since it is only required that the vehicle does not consistently move in the direction of a landmark for it to be observable. Situations as simple as circulating in 2-D throughout a corridor with features detected on the side walls fit on this criteria. While there is a focus on bearing measurements throughout the chapter and historically bearings are associated with radio direction finding equipment, the most obvious practical implementation of such an algorithm is with monocular vision. Following that line of reasoning, this chapter also presented a practical implementation of the algorithm applied to monocular vision. As a by-product of that implementation, it was shown that, without modifying the Kalman filter and losing its properties, it is possible to choose the input covariances of the filter in order to reduce the computational complexity of the filter. This version of the filter was then experimentally validated with data made available by the European project *Rawseeds* [BBF+06]. These experiments use a wheeled ground vehicle equipped with a forward-facing black-and-white monocular camera and wheel encoders for the odometry. As in the chapter tackling the RO-SLAM formulation, the results presented here are purely sensor-based (the map of the environment), which only allows for a qualitative evaluation of the filter. In any case, the convergence effects on initialization were apparent.

The second part of the dissertation was focused on complementing the sensor-based idea underlying all the filters from Part II with the objective of designing an algorithm to estimate the pose of the vehicle and the map in the Earth-fixed frame, effectively providing the prospective user a complete algorithm for simultaneous localization and mapping, from sensor information to Earth-fixed, human-readable, maps. With that purpose in mind, the ground work for the estimation of Earth-fixed quantities from a sensor-based map was laid in Chapter 7. The n-dimensional weighted orthogonal Procrustes optimization problem was studied and its uncertainty characterization with arbitrary covariance matrices was derived. The anisotropic nature of this characterization is a novelty in the field. A thorough analysis of the applicability of the method was also presented based on extensive Monte Carlo simulations for several dimensionalities.

Finally, the full Earth-fixed Trajectory and Map estimation algorithm was proposed in Chapter 8. There the relevance of the orthogonal Procrustes problem was exposed in the SLAM context, and the approach of the previous chapter was employed to derive a characterization for the resulting Earth-fixed map. The algorithm was coupled with each of the main SLAM filters proposed in the first part of the thesis in order to show not only its behaviour but also that of the complete strategy presented in this thesis. In simulation, the Earth-fixed range-only approach was studied in depth, showing the good performance of the algorithm and allowing the exploration of some of the decisions behind it. The range-and-bearing and bearing-only experimental data of Chapters 3 and 6 were then used to experimentally validate and evaluate the algorithm, both with the instrumented quadrotor of the University of Macau and the wheeled ground robot of the *Rawseeds* project. These experiments serve as a conclusion for the extensive work of the thesis.

9.2 Directions for Future Work

Simultaneous localization and mapping as a field of research exists since the 1980's [SC86], and even though a huge body of work has been done within its scope, there are still open problems being solved today. The relevance of SLAM will not fade in the foreseeable future, especially with the advent of autonomous driving and automation in industry. For this reason, one course of future action of particular interest is the optimization of the various algorithms for real-time operation, which is of paramount importance for achieving a truly online filter, thus paving the way for its use in practice, for example for autonomous/unmanned ground or aerial vehicles. The refinement of the associated algorithms for improved performance, robustness and reliability is an important point as well: the naïve feature detection/data association combo can be replaced by more state-of-the-art procedures such as key-frame based algorithms [LFR+13] or higher level image processing for more robust association and loop closing.

Other interesting direction of future research would be further additions to the sensor suites to improve performance, such as altitude, accelerometer, or magnetometer measurements. While adding these measurements will provide more, otherwise unavailable, information, it may introduce new calibration problems that would need to be addressed. In that line of reasoning, another relevant advancement would be the application of the framework proposed in this thesis to automatic calibration of a camera in RB-SLAM or BO-SLAM by the computation of the projection matrix.

With respect to the general work of Chapter 7 not directly related to SLAM, possible directions of future work include applying the methodology therein to similar optimization problems (e.g. LIDAR calibration from 3-D point clouds [GSO14]), and research on the use of the uncertainty characterization in iterative Procrustes or ICP algorithms.

In general, estimation and control are usually not considered together, with control problems assuming that system states are available and with estimation problems assuming that the state evolves with some external control. One of the first blaring exceptions comes from linear system theory where the linear quadratic Gaussian (LQG) control combines the linear quadratic regulator (LQR) with the Kalman filter [AM08]. In recent times, several fields of research have been spawned with the objective of either choosing the motion or the location of landmarks that minimize the uncertainty of the estimation problem. The first problem is called active estimation and the latter sensor/landmark placement [TI10]. As detailed in the motivation for this dissertation, localization or navigation as a problem in itself is an ages old problem. It is associated with known landmark locations. For that reason, in order to optimize the navigation an obvious course of action is optimal landmark placement, which is usually based in measures such as dilution of precision (DOP) – a common measure in GPS algorithms. When focusing in mapping, which is associated with known trajectories it makes sense to optimize them to obtain the most accurate map possible. The main paradigm in SLAM is to move to gain new knowledge and improve what is known. For that reason, both situations could be of interest. Since the idea is to map uncharted terrain (and the majority of applications use naturally extracted landmarks), active estimation is nevertheless the most important problem. In the formative years of SLAM, the question of how to move was completely separated from the estimation problem. However, in recent years, several works addressed the issue of intelligent moving in the context of SLAM, thus introducing Active SLAM [FLS99, SR05, LHD06, KKG09, CDKN$^+$14]. The objective of Active SLAM is to plan ahead the motion of the vehicle in order to maximize the explored areas and minimize the uncertainty associated with the estimation. These two objectives are, in a sense, complementary: exploration involves moving in previously unvisited terrain with the objective of increasing the overall knowledge of the environment, while the latter is exploitation, i.e., it involves revisiting areas to maximize the information gain. A complete Active SLAM algorithm would decide between which of these actions to take at each moment and how to perform them. Bridging the gap between estimation and control is then a more ambitious view of directions for research based on the current work on simultaneous localization and mapping. In the whole dissertation there was some concern on how the vehicle should move to ensure that observability conditions were satisfied, but actual trajectories were never discussed. Finding strategies to optimize the trajectory so that, for example, the uncertainty of the filter is kept at a minimum is, thus, a very relevant field of future research.

A

SUPPLEMENTARY DEFINITIONS AND RESULTS

This appendix presents a collection of definitions and results that are helpful to the theoretical work of the dissertation.

A.1 Results and definitions necessary for observability analysis

Definition A.1 (Lyapunov transformation, [Bro70, Chapter 1, Section 8]). *A transformation* $\mathbf{z}(t) = \mathbf{T}(t)\mathbf{x}(t)$, $\mathbf{T}(t) \in \mathbb{R}^{N\times N}$ *is called a Lyapunov transformation if*

i. $\mathbf{T}$ *has a continuous derivative on the interval* $(-\infty, \infty)$*,*

ii. $\mathbf{T}$ *and* $\dot{\mathbf{T}}$ *are bounded on the interval* $(-\infty, \infty)$*, in the sense that* $\left|t_{ij}(t)\right| \leq M \ \forall i, j \in [1, N]$ *and* $M < \infty$ *for all* $-\infty < t < \infty$*, where* $t_{ij}(t)$ *is the entry in the i-th row and j-th of* $\mathbf{T}(t)$*,*

iii. *there exists a constant* m *such that* $0 < m \leq |\det(\mathbf{T}(t))|$ *for all* t*.*

Proposition A.1 ([BSO11a, Proposition 4.2]). *Let* $\mathbf{f}(t) : [t_0, t_f] \subset \mathbb{R} \to \mathbb{R}^n$ *be a continuous and* i*-times differentiable function on* $\mathcal{T} := [t_0, t_f]$*,* $T := t_f - t_0 > 0$*, and such that*

$$\mathbf{f}(t_0) = \dot{\mathbf{f}}(t_0) = \cdots = \mathbf{f}^{(i-1)}(t_0) = \mathbf{0}.$$

Further assume that $\max\limits_{t\in\mathcal{T}} \left\|\mathbf{f}^{(i-1)}(t)\right\| \leq C$*. If there exists a positive constant* α *and an instant* $t_1 \in \mathcal{T}$ *such that* $\left\|\mathbf{f}^{(i)}(t_1)\right\| \geq \alpha$*, then there exists a positive constant* β *and* $0 < \delta \leq T$ *such that* $\|\mathbf{f}(t_0 + \delta)\| \geq \beta$*.*

Proof. The proof can be found in [BSO11a, Appendix]. □

A.2 Lemmas

Lemma A.2 (Weighted triangle inequality). *Let* V *be an Euclidean normed vector space of dimension* n*,* $\boldsymbol{\alpha}_i \in V$*, and* $k_i \in \mathbb{R}$*, with* $i = 1, \ldots, N$*. Then,*

$$\sum_{j=1}^{N} k_j^2 \sum_{i=1}^{N} \|k_i \boldsymbol{\alpha}_i\|^2 \geq \left\| \sum_{i=1}^{N} k_i^2 \boldsymbol{\alpha}_i \right\|^2,$$

where the equality applies if and only if $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_j\ \forall i, j = 1, \ldots, N$.

Proof. Consider the case $N = 2$. The triangle inequality, in which equality is only valid if $\boldsymbol{\alpha}_1 = \beta\ \boldsymbol{\alpha}_2$, for any $\beta > 0$, states that

$$\left\|k_1^2\boldsymbol{\alpha}_1 + k_2^2\boldsymbol{\alpha}_2\right\|^2 \leq k_1^4\left\|\boldsymbol{\alpha}_1\right\|^2 + k_2^4\left\|\boldsymbol{\alpha}_2\right\|^2 + 2k_1^2k_2^2\left\|\boldsymbol{\alpha}_1\right\|\left\|\boldsymbol{\alpha}_2\right\|.$$

Its generalization for $N > 2$ is given by

$$\left\|\sum_{i=1}^{N} k_i^2\boldsymbol{\alpha}_i\right\|^2 \leq \sum_{i=1}^{N}\sum_{j=1}^{N} k_i^2k_j^2\left\|\boldsymbol{\alpha}_i\right\|\left\|\boldsymbol{\alpha}_j\right\|. \tag{A.1}$$

This can be easily obtained if every pair $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ is accumulated in another $\boldsymbol{\alpha}_k$ (as in $\boldsymbol{\alpha}_k = k_i^2\boldsymbol{\alpha}_i + k_j^2\boldsymbol{\alpha}_j$) repeatedly until only two points exist, and then the right hand side of the inequality is expanded.

Consider Young's inequality, which states that, for $a, b > 0$ and $p, q > 0$, with $\frac{1}{p} + \frac{1}{q} = 1$,

$$ab \leq a^p/p + b^q/q, \tag{A.2}$$

with the equality valid only if $a^p = b^q$. Then, using (A.2) in (A.1) yields

$$\begin{aligned}
\left\|\sum_{i=1}^{N} k_i^2\mathbf{a}_i\right\|^2 &\leq \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} k_i^2k_j^2\left(\left\|\boldsymbol{\alpha}_i\right\|^2 + \left\|\boldsymbol{\alpha}_j\right\|^2\right)\\
&\leq \sum_{i=1}^{N}\sum_{j=1}^{N} k_i^2k_j^2\left\|\boldsymbol{\alpha}_i\right\|^2\\
&\leq \sum_{j=1}^{N} k_j^2\sum_{i=1}^{N} k_i^2\left\|\boldsymbol{\alpha}_i\right\|^2,
\end{aligned}$$

which concludes the proof of the inequality part of the lemma.

For the sufficiency part of the lemma, consider that the condition of the lemma holds, i.e., $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_j = \boldsymbol{\alpha}$ for all $i, j = 1, \ldots, N$. Then, it is a simple matter of computation to see that both sides are equal to $\left(k_1^2 + \ldots + k_N^2\right)^2\|\alpha\|^2$, and hence if the condition holds, the equality is true.

Suppose now that the condition does not hold. In that case, Young's inequality is a strict inequality, unless $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_j = -\boldsymbol{\alpha}$ for $i \in \mathcal{I}$ and $j \in \mathcal{J}$, with $\mathcal{I} \cup \mathcal{J} = \{1, \ldots, N\}$ and $\mathcal{I} \bigcap \mathcal{J} = \emptyset$. However, this case transforms (A.1) in a strict inequality. Therefore, if the conditions of the lemma do not hold, the equality is not true, and by contraposition, if the equality is true, the conditions of the lemma must hold, thus concluding the proof of the lemma. $\square$

B

PROOFS OF MINOR RESULTS

This appendix provides proofs for minor results whose statements are presented in the dissertation, but are not followed by the proofs to avoid disrupting the flow of the text.

B.1 Proof of Lemma 6.5

Consider the update equation [Gel74] for the covariance of the Kalman filter, given by

$$\Sigma_{k|k} = \left(\mathbf{I} - \Sigma_{k|k-1}\mathbf{H}_{F_k}^T \Sigma {\nu_k}^{-1}\right)\Sigma_{k|k-1}.$$

Recall that $\mathbf{H}_{F_k}$ is introduced in (2.9), and that the innovation covariance $\Sigma\nu_k$ is given by

$$\Sigma\nu_k = \mathbf{H}_{F_k}\Sigma_{k|k-1}\mathbf{H}_{F_k}^T + \mathbf{\Theta}_k.$$

Assume that $\Sigma_{k|k-1}$ has the structure in (6.22). In this case, if the measurement covariance $\mathbf{\Theta}_k$ is block diagonal, i.e., there are no cross-covariances between landmark observations, it is a matter of tedious computation to arrive at

$$\Sigma\nu_{kij} = \begin{cases} \Sigma_{M_i}^{k|k-1} - \mathbf{b}_{i_k}\Sigma_{R_iM_i}^{k|k-1} - \left(\Sigma_{M_iR_i}^{k|k-1} - \mathbf{b}_{i_k}\Sigma_{R_i}^{k|k-1}\right)\mathbf{b}_{i_k}^T + \mathbf{\Theta}_{i_k}, & i = j \\ \mathbf{0}, & i \neq j \end{cases}$$

which means that the innovation covariance is also block-diagonal and its inverse can be written as $\Sigma_{\nu_k}^{-1} = \operatorname{diag}(\Sigma_{\nu_{k1}}^{-1}, \ldots, \Sigma_{\nu_{kN_O}}^{-1})$. Using this structure it is a matter of computation to arrive at a covariance matrix with the structure in (6.22), where the individual elements are

$$\begin{cases} \Sigma_{M_i}^{k|k} = \Sigma_{M_i}^{k|k-1} - \left(\Sigma_{M_i}^{k|k-1} - \Sigma_{M_iR_i}^{k|k-1}\mathbf{b}_i^T\right)\Sigma_{\nu_i}^{-1}\left(\Sigma_{M_i}^{k|k-1} - \Sigma_{M_iR_i}^{k|k-1}\mathbf{b}_i^T\right)^T \\ \Sigma_{R_i}^{k|k} = \Sigma_{R_i}^{k|k-1} - \left(\Sigma_{R_iM_i}^{k|k-1} - \Sigma_{R_i}^{k|k-1}\mathbf{b}_i^T\right)\Sigma_{\nu_i}^{-1}\left(\Sigma_{R_iM_i}^{k|k-1} - \Sigma_{R_i}^{k|k-1}\mathbf{b}_i^T\right)^T \\ \Sigma_{M_iR_i}^{k|k} = \Sigma_{M_iR_i}^{k|k-1} - \left(\Sigma_{M_i}^{k|k-1} - \Sigma_{M_iR_i}^{k|k-1}\mathbf{b}_i^T\right)\Sigma_{\nu_i}^{-1}\left(\Sigma_{R_iM_i}^{k|k-1} - \Sigma_{R_i}^{k|k-1}\mathbf{b}_i^T\right)^T \end{cases}.$$

This shows that the structure is maintained by the update process of the Kalman filter when applied to the system at hand. Recall the discrete dynamics matrix $\mathbf{F}_{F_k}$ defined in (6.20). The predicted covariance for the instant $k+1$ is then given by

$$\Sigma_{k+1|k} = \boldsymbol{\Xi}_k + \begin{bmatrix} \mathbf{R}_{k+1}^T \mathbf{R}_k \Sigma_{M_1}^{k|k} \mathbf{R}_k^T \mathbf{R}_{k+1} & \cdots & \mathbf{0} & \mathbf{R}_{k+1}^T \mathbf{R}_k \Sigma_{M_1 R_1}^{k|k} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{k+1}^T \mathbf{R}_k \Sigma_{M_N}^{k|k} \mathbf{R}_k^T \mathbf{R}_{k+1} & \mathbf{0} & \cdots & \mathbf{R}_{k+1}^T \mathbf{R}_k \Sigma_{M_N R_N}^{k|k} \\ \Sigma_{R_1 M_1}^{k|k} \mathbf{R}_k^T \mathbf{R}_{k+1} & \cdots & \mathbf{0} & \Sigma_{R_1}^{k|k} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Sigma_{R_N M_N}^{k|k} \mathbf{R}_k^T \mathbf{R}_{k+1} & \mathbf{0} & \cdots & \Sigma_{R_N}^{k|k} \end{bmatrix}.$$

Thus, using the final condition of the lemma, that requires the process covariance $\boldsymbol{\Xi}_k$ to have the same structure as (6.22), one complete step of the Kalman filter for the BO-SLAM version of system (2.9) maintains throughout the update and predict stages the structure of the covariance matrix if it was (6.22). Therefore, if the initial covariance of the filter is set to have the structure in (6.22) and the conditions of the lemma hold, then the filter covariance will maintain that structure for all time. □

B.2 Proof of Lemma 7.1, Properties of the anti-commutation matrix

Due to the recursive nature of $\bar{\mathbf{S}}(\mathbf{a})$ and $\mathbf{S}(\boldsymbol{\varpi})$, it is possible to express all the sides of these identities for dimension $n+1$ in terms of the identity for n and some extra components that are equal in both sides. Given that the identities are clearly true for the bidimensional case, these properties can be proven using mathematical induction.

Property 7.1.1, Linearity

Let $\mathbf{a}_{1:n} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T \in \mathbb{R}^n$ and $\mathbf{b}_{1:n} = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}^T \in \mathbb{R}^n$ for all integer $n \geq 2$. The proof follows by first establishing that $\bar{\mathbf{S}}(\mathbf{a}_{1:n} + \mathbf{b}_{1:n}) = \bar{\mathbf{S}}(\mathbf{a}_{1:n}) + \bar{\mathbf{S}}(\mathbf{b}_{1:n})$. Consider the definition of the anti-commutation matrix in (7.2). It is clear that, if $\bar{\mathbf{S}}(\mathbf{a}_{1:n} + \mathbf{b}_{1:n}) = \bar{\mathbf{S}}(\mathbf{a}_{1:n}) + \bar{\mathbf{S}}(\mathbf{b}_{1:n})$, then the identity also applies for vectors of dimension $n+1$. In that case, and given that for $n = 2$ the identity also applies, it is shown by induction that the identity applies for all integer $n \geq 2$.

The elements of $\bar{\mathbf{S}}(\mathbf{a}_{1:n})$ are linearly related to the elements of $\mathbf{a}_{1:n}$, and, therefore, the identity $\bar{\mathbf{S}}(\alpha \mathbf{a}_{1:2}) = \alpha \bar{\mathbf{S}}(\mathbf{a}_{1:2})$ is apparent. Supposing that this is also true for n dimensions, then it is clear that it will also apply to vectors of dimension $n+1$. Therefore, by induction, as the identity is true for $n = 2$, it is shown that $\bar{\mathbf{S}}(\alpha \mathbf{a}_{1:n}) = \alpha \bar{\mathbf{S}}(\mathbf{a}_{1:n})$ for all integer $n \geq 2$.

Combining these two results, it is possible to write

$$\begin{aligned} \bar{\mathbf{S}}(\alpha \mathbf{a}_{1:n} + \beta \mathbf{b}_{1:n}) &= \bar{\mathbf{S}}(\alpha \mathbf{a}_{1:n}) + \bar{\mathbf{S}}(\beta \mathbf{b}_{1:n}) \\ &= \alpha \bar{\mathbf{S}}(\mathbf{a}_{1:n}) + \beta \bar{\mathbf{S}}(\mathbf{b}_{1:n}), \end{aligned}$$

thus concluding the proof.

Property 7.1.2, Anti-commutativity

For $n = 2$,

$$\bar{\mathbf{S}}(\mathbf{a}_{1:2})\,\mathbf{b}_{1:2} = \begin{bmatrix} -a_2 & a_1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = -a_2 b_1 + a_1 b_2. \tag{B.1}$$

Expanding $-\bar{\mathbf{S}}(\mathbf{b}_{1:2})\,\mathbf{a}_{1:2}$, it is a simple matter of computation to see that it yields the same result, and hence the property is verified for $n = 2$. Suppose now that the property holds for arbitrary n, and compute, for $n+1$, the left hand side of Property 7.1.2

$$\begin{aligned} \bar{\mathbf{S}}(\mathbf{a}_{1:n+1})\,\mathbf{b}_{1:n+1} &= \left[\begin{array}{cccc|c} 0 & \cdots & 0 & -a_{n+1} & (-1)^0 a_n \\ \vdots & & (-1)^2 a_{n+1} & 0 & (-1)^1 a_{n-1} \\ 0 & \iddots & & \vdots & \vdots \\ (-1)^n a_{n+1} & 0 & \cdots & 0 & (-1)^{n-1} a_1 \\ \hline & & \bar{\mathbf{S}}(\mathbf{a}_{1:n}) & & \mathbf{0}_{\frac{n(n-1)}{2}\times 1} \end{array}\right] \begin{bmatrix} b_1 \\ \vdots \\ b_{n-1} \\ b_n \\ b_{n+1} \end{bmatrix} \\ &= \begin{bmatrix} -a_{n+1} b_n + a_n b_{n+1} \\ a_{n+1} b_{n-1} - a_{n-1} b_{n+1} \\ \vdots \\ (-1)^n a_{n+1} b_1 + (-1)^{n-1} a_1 b_{n+1} \\ \bar{\mathbf{S}}(\mathbf{a}_{1:n})\,\mathbf{b}_{1:n} \end{bmatrix}. \end{aligned} \tag{B.2}$$

For the right hand side similar computations lead to

$$\bar{\mathbf{S}}(\mathbf{b}_{1:n+1})\,\mathbf{a}_{1:n+1} = - \begin{bmatrix} b_{n+1} a_n - b_n a_{n+1} \\ -b_{n+1} a_{n-1} + b_{n-1} a_{n+1} \\ \vdots \\ (-1)^{n-1} b_{n+1} a_1 + (-1)^n b_1 a_{n+1} \\ -\bar{\mathbf{S}}(\mathbf{b}_{1:n})\,\mathbf{a}_{1:n} \end{bmatrix},$$

whose first n elements are clearly the same as the ones in (B.2). Also, from the induction hypothesis, it is known that $\bar{\mathbf{S}}(\mathbf{a}_{1:n})\,\mathbf{b}_{1:n} = -\bar{\mathbf{S}}(\mathbf{b}_{1:n})\,\mathbf{a}_{1:n}$. Then, if the identity is valid for n, it can be seen that it must be true for $n+1$. Given that it holds for $n = 2$, this implies that the identity holds for all $n \geq 2$ by induction, thus concluding the proof of this property.

Property 7.1.3, Anti-commutation with a skew-symmetric matrix

Consider the vector $\boldsymbol{\varpi}_n = \begin{bmatrix} \varpi_{n_1} & \cdots & \varpi_{n_{n-1}} & \boldsymbol{\varpi}_{n-1}^T \end{bmatrix}^T \in \mathbb{R}^{\frac{n(n-1)}{2}}$, where $\boldsymbol{\varpi}_{n-1} \in \mathbb{R}^{\frac{(n-1)(n-2)}{2}}$, ϖ_{n_i} represents the i-th element of vector $\boldsymbol{\varpi}_n$, and $\boldsymbol{\varpi}_2 = \varpi_{2_1}$. In order to be able to use $\boldsymbol{\varpi}_n$ to parametrize a skew-symmetric matrix, note that in the definition presented in (7.1), the elements of the vector $\boldsymbol{\varpi}$ are used in reverse order, from the last element, ϖ_{n_p}, down to the first, ϖ_1. Hence, it is possible to write $\boldsymbol{\varpi}_2 = \varpi_{n_p}$, $\boldsymbol{\varpi}_3 = \begin{bmatrix} \varpi_{n_p-2} & \varpi_{n_p-1} & \varpi_{n_p} \end{bmatrix}^T$, $\boldsymbol{\varpi}_4 = \begin{bmatrix} \varpi_{n_p-5} & \cdots & \varpi_{n_p} \end{bmatrix}^T$, and so on and so forth up to $\boldsymbol{\varpi}_n = \begin{bmatrix} \varpi_1 & \cdots & \varpi_{n_p} \end{bmatrix}^T$. Noting this, it is possible to redefine the skew-symmetric matrix in a recursive way, as it is convenient for the use of mathematical induction in this proof. The recursive definition is

$$\mathbf{S}\left(\varpi_{(n+1)}\right) = \left[\begin{array}{c|c} \mathbf{S}(\varpi_n) & \begin{matrix}(-1)^n\varpi_{(n+1)_n}\\ \vdots \\ (-1)\varpi_{(n+1)_1}\end{matrix} \\ \hline * & 0\end{array}\right],$$

where

$$\mathbf{S}(\varpi_2) = \begin{bmatrix} 0 & -\varpi_{2_1} \\ \varpi_{2_1} & 0 \end{bmatrix}.$$

For the bidimensional case, the left hand side of the property in analysis is

$$\mathbf{S}(\varpi_2)\,\mathbf{a}_{1:2} = \begin{bmatrix} 0 & -\varpi_{2_1} \\ \varpi_{2_1} & 0 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -\varpi_{2_1}a_2 \\ \varpi_{2_1}a_1 \end{bmatrix}. \tag{B.3}$$

The right hand side of the property is equal to $\begin{bmatrix} -a_2 & a_1 \end{bmatrix}^T \varpi_{2_1}$, which is the same as (B.3). The proof follows naturally by assuming the induction hypothesis that the identity is valid for an arbitrary n, and showing that, in that case, it is also valid for $n+1$. Consider the left hand side of the property for the $(n+1)$-dimensional case:

$$\begin{aligned}\mathbf{S}(\varpi_{n+1})\,\mathbf{a}_{1:n+1} &= \left[\begin{array}{c|c} \mathbf{S}(\varpi_n) & \begin{matrix}(-1)^n\varpi_{(n+1)_n}\\ \vdots \\ (-1)\varpi_{(n+1)_1}\end{matrix} \\ \hline * & 0\end{array}\right]\begin{bmatrix} \mathbf{a}_{1:n} \\ a_{n+1}\end{bmatrix} \\ &= \begin{bmatrix} \mathbf{S}(\varpi_n)\,\mathbf{a}_{1:n} + \begin{bmatrix}(-1)^n\varpi_{(n+1)_n}a_{n+1} \\ \vdots \\ (-1)\varpi_{(n+1)_1}a_{n+1}\end{bmatrix} \\ \\ -(-1)^n\varpi_{(n+1)_n}a_1 - \cdots - (-1)\varpi_{(n+1)_1}a_n \end{bmatrix}\end{aligned}$$

The right hand side is

$$\bar{\mathbf{S}}^T(\mathbf{a}_{1:n+1})\,\varpi_{n+1} = \left[\begin{array}{cccc|c} 0 & \cdots & 0 & (-1)^n a_{n+1} & \\ \vdots & & \iddots & 0 & \bar{\mathbf{S}}^T(\mathbf{a}_{1:n}) \\ 0 & (-1)^2 a_{n+1} & & \vdots & \\ (-1)a_{n+1} & 0 & \cdots & 0 & \\ \hline (-1)^0 a_n & (-1)^1 a_{n-1} & \cdots & (-1)^{n-1}a_1 & \mathbf{0}_{1\times\frac{n(n-1)}{2}} \end{array}\right]\begin{bmatrix} \varpi_{(n+1)_1} \\ \vdots \\ \varpi_{(n+1)_n} \\ \varpi_n \end{bmatrix},$$

which, after multiplication yields

$$\bar{\mathbf{S}}^T(\mathbf{a}_{1:n+1})\,\varpi_{n+1} = \begin{bmatrix} \begin{bmatrix}(-1)^n a_{n+1}\varpi_{(n+1)_n} \\ \vdots \\ (-1)a_{n+1}\varpi_{(n+1)_1}\end{bmatrix} + \bar{\mathbf{S}}^T(\mathbf{a}_{1:n})\,\varpi_n \\ \\ (-1)^0 a_n\varpi_{(n+1)_1} + \cdots + (-1)^{n-1}a_1\varpi_{(n+1)_n} \end{bmatrix}.$$

Notice that, under the induction hypothesis, these two expressions are exactly the same. Hence, if Property 7.1.3 holds for any n, then it will hold for $n+1$. Given that it holds for $n = 2$, by mathematical induction it follows that it holds for all $n \geq 2$.

Property 7.1.4, Relation to unskew operator

Consider the argument of the unskew operator in the bidimensional case, given by

$$\mathbf{b}_{1:2}\mathbf{a}_{1:2}^T - \mathbf{a}_{1:2}\mathbf{b}_{1:2}^T = \begin{bmatrix} 0 & a_2b_1 - a_1b_2 \\ a_1b_2 - a_2b_1 & 0 \end{bmatrix}.$$

Applying the unskew operator to this identity yields

$$\mathbf{S}^{-1}\Big(\mathbf{b}_{1:2}\mathbf{a}_{1:2}^T - \mathbf{a}_{1:2}\mathbf{b}_{1:2}^T\Big) = -a_2b_1 + a_1b_2,$$

which is clearly the same as $\bar{\mathbf{S}}(\mathbf{a}_{1:2})\mathbf{b}_{1:2}$ (see (B.1)).

For an arbitrary i, the argument of the unskew operator is

$$\mathbf{b}_{1:i}\mathbf{a}_{1:i}^T - \mathbf{a}_{1:i}\mathbf{b}_{1:i}^T = \begin{bmatrix} 0 & a_2b_1 - a_1b_2 & a_3b_1 - a_1b_3 & \cdots & a_ib_1 - a_1b_i \\ * & 0 & a_3b_2 - a_2b_3 & \cdots & a_ib_2 - a_2b_i \\ * & * & \ddots & & \vdots \\ * & * & * & 0 & a_ib_{i-1} - a_{i-1}b_i \\ * & * & * & * & 0 \end{bmatrix},$$

and the result of unskewing is

$$\mathbf{S}^{-1}\Big(\mathbf{b}_{1:i}\mathbf{a}_{1:i}^T - \mathbf{a}_{1:i}\mathbf{b}_{1:i}^T\Big) = \left[\begin{array}{c} -(a_ib_{i-1} - a_{i-1}b_i) \\ \vdots \\ (-1)^{i-1}(a_ib_1 - a_1b_i) \\ \hline -(a_{i-1}b_{i-2} - a_{i-2}b_{i-1}) \\ \vdots \\ (-1)^{i-2}(a_{i-1}b_1 - a_1b_{i-1}) \\ \vdots \\ -(a_2b_1 - a_1b_2) \end{array}\right].$$

It can be seen that the lower components are the same as $\mathbf{S}^{-1}\Big(\mathbf{b}_{1:i-1}\mathbf{a}_{1:i-1}^T - \mathbf{a}_{1:i-1}\mathbf{b}_{1:i-1}^T\Big)$, which is important for the sequel.

Assume now that Property 7.1.4 holds for an arbitrary n. Then, the left hand side of the identity in Property 7.1.4 for $(n+1)$-dimensional space is

$$\mathbf{S}^{-1}\Big(\mathbf{b}_{1:n+1}\mathbf{a}_{1:n+1}^T - \mathbf{a}_{1:n+1}\mathbf{b}_{1:n+1}^T\Big) = \begin{bmatrix} -a_{n+1}b_n + a_nb_{n+1} \\ \vdots \\ (-1)^n a_{n+1}b_1 + (-1)^{n-1}a_1b_{n+1} \\ \mathbf{S}^{-1}\Big(\mathbf{b}_{1:n}\mathbf{a}_{1:n}^T - \mathbf{a}_{1:n}\mathbf{b}_{1:n}^T\Big) \end{bmatrix}.$$

Comparison of this expression with (B.2) confirms the identity on the assumption that it holds for n. It was shown that the property is true for $n = 2$, and that if it holds for n, then it will also hold for $n+1$. Thus, Property 7.1.4 must be true for all n, by induction. □

B.3 The error/perturbation model for orthogonal matrices

This section addresses the proof of Lemma 7.2, which requires the introduction of the following property.

Lemma B.1 (Perturbed determinant). *Let* $\mathbf{S}(\varpi) \in \mathfrak{so}(n)$ *and* $\varpi \in \mathbb{R}^{\frac{n(n-1)}{2}}$. *The determinant of the matrix* $\mathbf{I}_n + \epsilon\mathbf{S}(\varpi)$ *is* $1 + \mathcal{O}(\epsilon^2)$, *for all integer* $n \geq 2$.

Proof. Recall the definition of the determinant of an arbitrary square matrix $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^n$, given by the Leibniz formula [Lüt96], i.e.,

$$|\mathbf{A}| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma_i} \tag{B.4}$$

where σ_i is the i-th element of the permutation σ belonging to the set S_n of permutations of $\{1,\ldots,n\}$, and $\operatorname{sgn}(\sigma)$ is the signature of the permutation, corresponding to 1 for an even number of permuted pairs and -1 for an odd number. The first parcel of (B.4) corresponds to the trivial permutation $\sigma^* = \{1,\ldots,n\}$, which means that

$$|\mathbf{A}| = \prod_{i=1}^{n} a_{ii} + \sum_{\sigma \in S_n/\sigma^*} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma_i}.$$

In the particular case of $\mathbf{A} = \mathbf{I} + \epsilon\mathbf{S}(\varpi)$, where all the diagonal elements are equal to 1, the first parcel becomes $\prod_{i=1}^{n} a_{ii} = 1$. Moreover, every non-trivial permutation will lead to at least two off-diagonal elements appearing in the product of the $a_{i\sigma_i}$ for all $i = 1,\ldots,n$. Thus, due to the nature of this matrix, all the parcels other than the trivial permutation will be of the order of ϵ^2 or higher. This means that

$$|\mathbf{I} + \epsilon\mathbf{S}(\varpi)| = 1 + \mathcal{O}(\epsilon^2),$$

and thus the proof is concluded. ☐

Is is now possible to proceed to the proof of Lemma 7.2.

Proof of Lemma 7.2

The proof is done by exploring the constraints imposed by the matrix space O(n) and using the error model (7.9).

For $n \times n$ matrices in the orthogonal group, one has, by definition, that

$$\mathbf{M}^T\mathbf{M} = \mathbf{I} \tag{B.5}$$

and

$$|\mathbf{M}| = \pm 1. \tag{B.6}$$

The last relation is derived from the first, and as such is redundant, but it is convenient to state it for the sequel. Substituting the error model (7.9) in (B.5) yields, after expansion,

$$\mathbf{M}^{(0)^T}\mathbf{M}^{(0)} + \epsilon\left(\mathbf{M}^{(0)^T}\mathbf{M}^{(1)} + \mathbf{M}^{(1)^T}\mathbf{M}^{(0)}\right) = \mathbf{I},$$

where terms in ϵ^2 or higher order were not included. From this expression, it can be seen that

$$\mathbf{M}^{(0)^T}\mathbf{M}^{(0)} = \mathbf{I}, \tag{B.7}$$

which means that $\mathbf{M}^{(0)} \in \mathrm{O}(n)$, and that

$$\mathbf{0} = \mathbf{M}^{(0)^T}\mathbf{M}^{(1)} + \mathbf{M}^{(1)^T}\mathbf{M}^{(0)}$$

or, equivalently,

$$\mathbf{0} = \mathbf{M}^{(1)}\mathbf{M}^{(0)^T} + \mathbf{M}^{(0)}\mathbf{M}^{(1)^T}.$$

From these last identities, it comes that both $\mathbf{M}^{(1)}\mathbf{M}^{(0)^T}$ and $\mathbf{M}^{(0)^T}\mathbf{M}^{(1)}$ are skew-symmetric matrices. Let $\mathbf{S}(\boldsymbol{\varpi})$ and $\mathbf{S}(\bar{\boldsymbol{\varpi}})$ be two $n \times n$ skew-symmetric matrices parametrized by the vectors $\boldsymbol{\varpi} \in \mathbb{R}^{\frac{n(n-1)}{2}}$ and $\bar{\boldsymbol{\varpi}} \in \mathbb{R}^{\frac{n(n-1)}{2}}$. Then, it is possible to write

$$\mathbf{M}^{(0)^T}\mathbf{M}^{(1)} = \mathbf{S}(\bar{\boldsymbol{\varpi}}), \tag{B.8}$$

and

$$\mathbf{M}^{(1)}\mathbf{M}^{(0)^T} = \mathbf{S}(\boldsymbol{\varpi}) \tag{B.9}$$

which, left and right multiplying (B.8) and (B.9) by $\mathbf{M}^{(0)}$, respectively, and using (B.7) on both, can be rearranged to give

$$\mathbf{M}^{(1)} = \mathbf{M}^{(0)}\mathbf{S}(\bar{\boldsymbol{\varpi}})$$

and

$$\mathbf{M}^{(1)} = \mathbf{S}(\boldsymbol{\varpi})\mathbf{M}^{(0)}. \tag{B.10}$$

It remains to analyse the determinant of $\mathbf{M}^{(0)}$. For that purpose, consider (B.6), and its expansion according to the error model (7.9), yielding

$$\left|\mathbf{M}^{(0)} + \epsilon\mathbf{M}^{(1)}\right| = |\mathbf{M}|,$$

where the terms of order 2 and higher were discarded. Recalling (B.10), and rearranging the argument of the determinant on the left hand side yields

$$|\mathbf{I} + \epsilon\mathbf{S}(\boldsymbol{\varpi})|\left|\mathbf{M}^{(0)}\right| = |\mathbf{M}|.$$

At this time, Lemma B.1 is of use, leading to

$$\left|\mathbf{M}^{(0)}\right| = |\mathbf{M}|. \tag{B.11}$$

This concludes the proof that $\mathbf{M}^{(0)}$ does indeed belong to $\mathrm{O}(n)$, and that it retains the determinant of the perturbed version.

The final part of the proof follows naturally from this. If $\mathbf{M} \in \mathrm{SO}(n) \subsetneq \mathrm{O}(n)$, then its determinant is 1. Using (B.11), it comes that the determinant of $\mathbf{M}^{(0)}$ is also 1, thus belonging to $\mathrm{SO}(n)$ too. □

C

Simulated Environment

This appendix explains the simulated environment used throughout the thesis (Sections 4.5, 5.5, 6.5, and 8.4), detailing the vehicle trajectories, the motion model, and the landmark environment.

C.1 The map

The generic map used for the simulations in this thesis is inspired on the fifth floor of the North Tower at IST, as shown in Figure C.1. The landmarks are placed in 36 corners and doors with random heights.

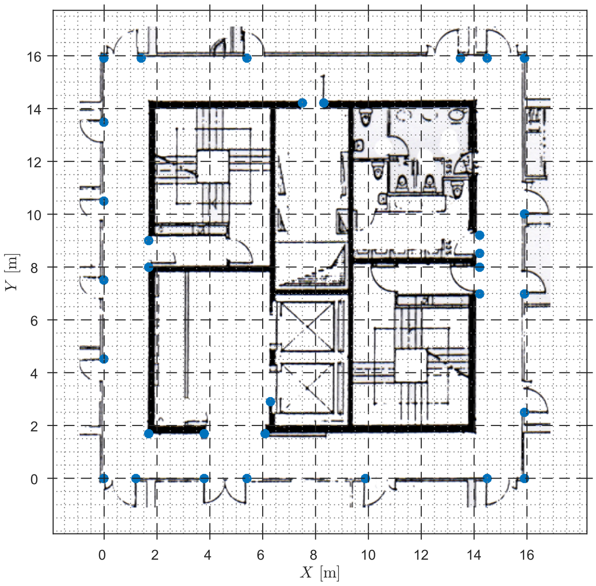


Figure C.1: The top-view of the Earth-fixed map.

C.2 Motion model

The motion model used in the simulations relies on a discrete kinematic model, described by

$$\begin{cases} \mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(T_s \mathbf{S}(\boldsymbol{\omega}_k)\right) \\ {}^E\mathbf{p}_{k+1} = {}^E\mathbf{p}_k + T_s {}^E\mathbf{v}_k \end{cases}.$$

This is completely generic in terms of the type of vehicle used and fits the kinematic model used for the landmarks in system design (see Section 2.1.2). The linear velocity is provided by the differentiation of the trajectory designed in the next section. The structure of the rotation matrix is enforced to keep the vehicle always moving forward in its body-fixed reference frame, as well as to maintain the horizontal axis perpendicular to the force that would have to be applied to the vehicle for that particular trajectory to hold, the force per mass vector is simply

$${}^E\mathbf{F}_k = \frac{\mathrm{d}^2 {}^E\mathbf{p}(t)}{\mathrm{d}t^2}\Big|_{t=t_k} + \mathbf{g}$$

where $\mathbf{g} \approx \begin{bmatrix} 0 & 0 & 9.81 \end{bmatrix}^T$ m/s^2 is the acceleration of gravity. The resulting rotation matrix is then

$$\mathbf{R}_k = \begin{bmatrix} \frac{{}^E\mathbf{v}_k}{\|{}^E\mathbf{v}_k\|} & \frac{{}^E\mathbf{v}_k \times {}^E\mathbf{F}_k}{\|{}^E\mathbf{v}_k \times {}^E\mathbf{F}_k\|} & \frac{{}^E\mathbf{v}_k \times {}^E\mathbf{v}_k \times {}^E\mathbf{F}_k}{\|{}^E\mathbf{v}_k \times {}^E\mathbf{v}_k \times {}^E\mathbf{F}_k\|} \end{bmatrix} \tag{C.1}$$

C.3 Trajectory design

The path was designed in order to allow the vehicle to cycle counter-clockwise through the four corridors that compose the map. Furthermore, since most of the algorithms require a trajectory rich in all directions for their observability and convergence, a possible "wandering" motion in the horizontal and vertical planes was added. The vehicle starts at rest for some time and then takes-off, until it reaches constant speed at the desired altitude. The forward speed is maintained while in a corridor with the possibility of vertical and horizontal "wandering"with spatial sinusoidal motion, and changing direction when the vehicle enters a corner. Figure C.2 depicts the chosen trajectory in bidimensional views.

The trajectory of the vehicle is divided in various phases, namely:

- take-off;
- flight through a corridor; and
- making a turn at a corner.

To simplify design the following equations of motion are defined with respect to the initial time at which the vehicle enters each of the corresponding phases. Consider $\{c_j, l_j, z_j\}$ as the frame aligned with the corridor that starts at corner j, c being the direction of forward movement in a corridor, l pointing towards the left of each corridor, and z defined in order to obtain a right-handed reference frame. An example for frame A is shown in Figure C.2.

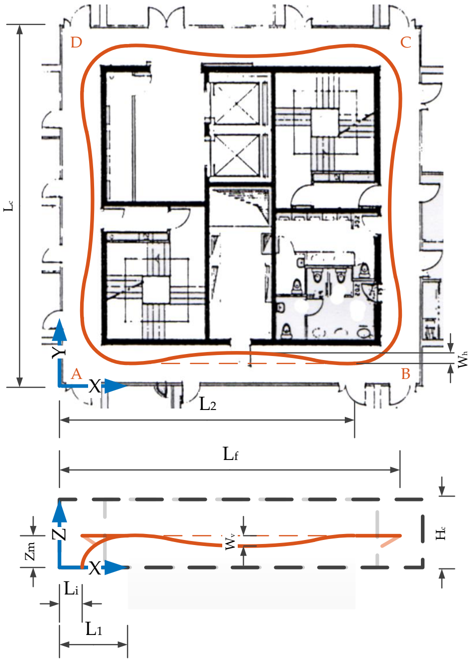


Figure C.2: Bidimensional path followed by the vehicle during simulation.

The parameters are the dimensions of the corridors, the time allotted for each phase which determines the forward corridor velocity $V_C > 0$, the steady flight altitude $Z_m > 0$, the number of cycles in the corridor wandering $k_w \in \mathbb{N}$, and the amplitudes of the horizontal and vertical sinusoidal wandering motions, respectively $W_h \in \mathbb{R}$ and $W_v \in \mathbb{R}$.

Corridors In the corridors, the idea is to have the vehicle have a constant forward speed while also moving along a sinusoidal path, both horizontally and vertically. Given the number of oscillations wanted, it is possible to compute the spatial angular velocity of that sinusoidal path, $\Omega = \frac{2k_w\pi}{L_2-L_1}$, and to have φ as a phase parameter to align the trajectory as intended.

Since the forward speed in c_j is set to V_C, it only remains to compute the other two velocities, given by

$$\begin{cases} \dfrac{\mathrm{d}l_j}{\mathrm{d}t}(t) = -\Omega V_C W_h \cos\Big(\Omega(c_j(t) - L_1) + \varphi\Big) \\ \dfrac{\mathrm{d}z_j}{\mathrm{d}t}(t) = -\Omega V_C W_v \cos\Big(\Omega(c_j(t) - L_1) + \varphi\Big) \end{cases}$$

where all quantities are expressed in the local corridor j reference frame. Given that the rotation matrix (C.1) depends both on the linear velocity and on the total acceleration (linear acceleration with gravity), it is necessary to compute it as well. The component in c_j is naturally zero, and the other two are given by

$$\begin{cases} \dfrac{\mathrm{d}^2 l_j}{\mathrm{d}t^2}(t) = \Omega^2 V_C^2 W_h \sin\Big(\Omega(c_j(t) - L_1) + \varphi\Big) \\ \dfrac{\mathrm{d}^2 z_j}{\mathrm{d}t^2}(t) = \Omega^2 V_C^2 W_v \sin\Big(\Omega(c_j(t) - L_1) + \varphi\Big) \end{cases}$$

For the corners and take-off phases, the three coordinates are described by the following 7 degree polynomials.

$$\begin{cases} c_j(t) = \alpha_c t^7 + \beta_c t^6 + \gamma_c t^5 + \delta_c t^4 + \varepsilon_c t^3 + \zeta_c t^2 + \eta_c t + \vartheta_c \\ l_j(t) = \alpha_l t^7 + \beta_l t^6 + \gamma_l t^5 + \delta_l t^4 + \varepsilon_l t^3 + \zeta_l t^2 + \eta_l t + \vartheta_l \\ z_j(t) = \alpha_z t^7 + \beta_z t^6 + \gamma_z t^5 + \delta_z t^4 + \varepsilon_z t^3 + \zeta_z t^2 + \eta_z t + \vartheta_z \end{cases} \tag{C.2}$$

The coefficients depend on the flight phase and will be computed next, bearing in mind that, since the filters proposed in this dissertation rely on kinematics, the main objective is to have continuity in the linear and angular velocities throughout the whole trajectory. Since the angular velocity is computed through the rotation matrix, which depends on the linear velocity and acceleration, the polynomial model was chosen to ensure continuity in these two variables.

The corners Flying through the corners is made at constant altitude, and varying c and l. There are 8 boundary conditions for every coordinate, as shown in Table C.1 where t_i and t_f are the time instants when the vehicle enters and leaves the corner, respectively.

Table C.1: Boundary conditions for motion in corners.

	$(.)$	$\frac{\mathrm{d}}{\mathrm{d}t}(.)$	$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(.)$	$\frac{\mathrm{d}^3}{\mathrm{d}t^3}(.)$
$c_j(t_i)$	L_2	V_C	0	0
$c_j(t_f)$	L_f	0	$\Omega^2 V_C^2 W_h \sin(\Omega(L_1 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_h \cos(\Omega(L_1 - L_1) + \varphi)$
$l_j(t_i)$	L_i	0	$\Omega^2 V_C^2 W_h \sin(\Omega(L_2 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_h \cos(\Omega(L_2 - L_1) + \varphi)$
$l_j(t_f)$	L_1	V_C	0	
$z_j(t_i)$	Z_m	0	$\Omega^2 V_C^2 W_z \sin(\Omega(L_2 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_z \sin(\Omega(L_2 - L_1) + \varphi)$
$z_j(t_f)$	Z_m	0	$\Omega^2 V_C^2 W_z \sin(\Omega(L_1 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_z \sin(\Omega(L_1 - L_1) + \varphi)$

Take-off Take-off is very similar to the corners, the difference being that l is constant, and z varies from 0 to the flight altitude Z_{max}. The boundary conditions are given in Table C.2.

Table C.2: Boundary conditions for motion in the take-off phase.

	$(.)$	$\frac{\mathrm{d}}{\mathrm{d}t}(.)$	$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(.)$	$\frac{\mathrm{d}^3}{\mathrm{d}t^3}(.)$
$c_A(t_i)$	L_i	0	0	0
$c_A(t_f)$	L_1	V_C	0	0
$l_A(t_i)$	L_i	0	0	0
$l_A(t_f)$	L_i	0	$\Omega^2 V_C^2 W_h \sin(\Omega(L_1 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_h \cos(\Omega(L_1 - L_1) + \varphi)$
$z_A(t_i)$	0	0	0	0
$z_A(t_f)$	Z_m	0	$\Omega^2 V_C^2 W_z \sin(\Omega(L_1 - L_1) + \varphi)$	$\Omega^3 V_C^3 W_z \cos(\Omega(L_1 - L_1) + \varphi)$

It is a matter of tedious computation to arrive at expressions for the coefficients of (C.2), using the boundary conditions in Tables C.1 and C.2. The result is not relevant enough to warrant reproduction here, but it can be obtained by solving a simple linear matrix equation.

Given the individual trajectories in each of the phases, it is possible to express them all in a common Earth-fixed frame according to Figure C.2. The remaining quantities can then be computed with the kinematic model, which is initialized with the following initial conditions

$$\mathbf{R}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad {}^{E}\mathbf{p}_0 = \begin{bmatrix} L_i & L_i & 0 \end{bmatrix}^T, \quad \omega_0 = \mathbf{0}, \quad \text{and} \quad {}^{E}\mathbf{v}_0 = \mathbf{0},$$

where L_i is defined in Figure C.2.

D

Experimental Setup: SCORE Lab

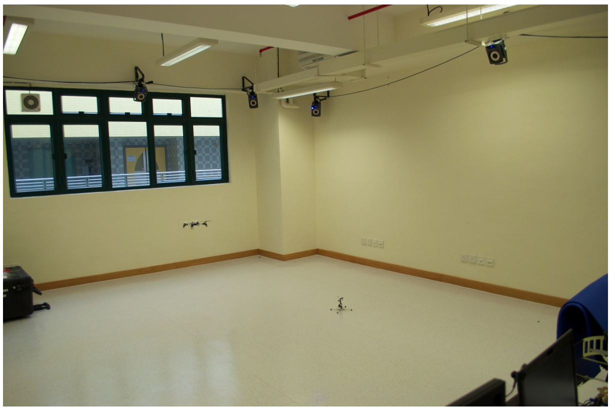


Figure D.1: The SCORE Lab.

At the time of the experiments of Sections 3.5, 4.6, and 8.5 (RB-SLAM and RO-SLAM), the facilities of the Sensor-based Cooperative Robotics Research (SCORE) Lab[1] at the Department of Electrical and Computer Engineering of the Faculty of Science and Technology, University of Macau, consisted of a 6m×6m×3m room, approximate dimensions, with a usable area of around 16 m^2, see Fig. D.1. The room was equipped with a *VICON* motion capture system, which provides accurate estimates of the position, attitude, and linear and angular velocities of any vehicle placed inside the working area with the correct markers. In terms of aerial robots, the lab has available several quadrotors of different dimensions, such as a *AscTec Pelican* quadrotor. For the purpose of these experiments, the *AscTec Pelican* was instrumented with a *Microsoft Kinect* RGB-D camera, for the extraction of visual features of the room, a *Microstrain 3DM-GX3-25* inertial measurement unit, and a *Crossbow Cricket* receiver. The objective of the experiment was to assess the SLAM algorithms with real data, resorting to the ground truth data provided by the *VICON* system. For that purpose, the quadrotor was equipped with several *VICON* markers. The following sections detail the equipment utilized.

---

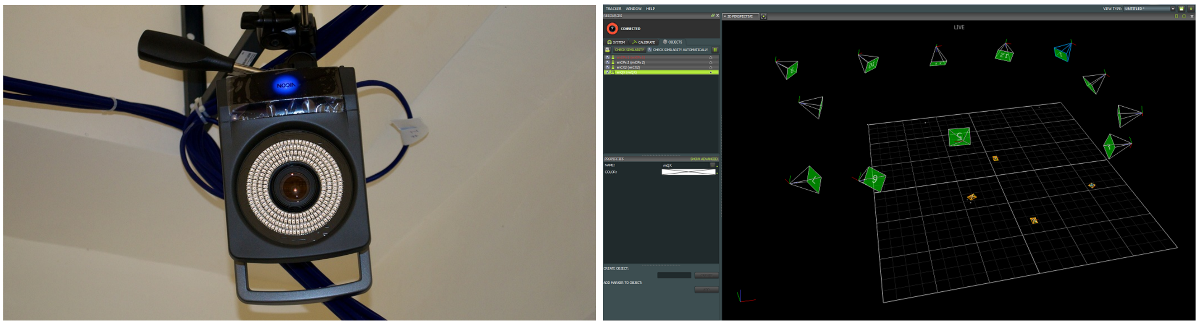
[1] `https://score.fst.umac.mo/`

D.1 *VICON* Motion Capture System

The *VICON* Motion Capture System [VIC], as its name suggests, is a state-of-the-art optical system that records the movement of objects or people with millimetric resolution in 3-D, by means of infrared marker-tracking. Its primary uses involve military, entertainment[2], sports, medical, computer vision, and robotic applications. For the last two, it provides a very good basis for validation. The high-accuracy of the tracking achieved by *VICON*, see [WGM08] for details, enables its use as ground truth for experimental validation of positioning and attitude estimation algorithms.

Optical motion capture uses a series of high-resolution, high-rate, infrared cameras setup around a working area, viewing the scene from a variety of angles[3]. Reflective markers are placed on the object to be tracked, and are recognized by the software, which then combines the information of the full set of cameras to generate a tridimensional picture of the markers[4]. Identifying a set of matrices as a rigid body allows to track the object, and the placement of the markers has to be chosen carefully, in order to maximize the performance of the capture.

The *VICON* instance present in the SCORE Lab is capable of acquiring more than 100 frames per second, providing the position, attitude, and linear and angular velocities of the objects therein in that fast rate. The latter quantities are obtained by finite differences of the former, and are thus less smoother than the fundamental quantities that determine the pose of the object. An example of the *VICON* cameras used and a screenshot of the proprietary software can be found in Fig. D.2.



(a) A *VICON* camera.

(b) The *VICON* software.

Figure D.2: The *VICON* system of the SCORE Lab.

D.2 The instrumented *AscTec Pelican*

The experiments detailed in the chapters mentioned above were conducted, in a hand-driven fashion, using as platform the quadrotor of Fig. D.3, equipped with an Intel Atom 1.6 GHz

---

[2] `http://www.awn.com/news/technology/ilm-employs-vicon-motion-capture-avengers`

[3] `http://physbam.stanford.edu/cs448x/old/Optical_Motion_Capture_Guide.html`

[4] `http://www.cs.utah.edu/~halzahaw/MotionCapture_main.html`

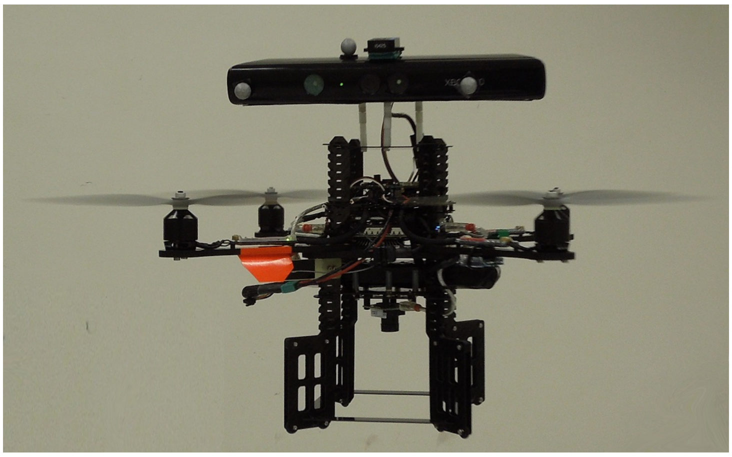


Figure D.3: The *AscTec Pelican* equipped with a *Microsoft Kinect*, a *Microstrain 3DM-GX3-25* and *VICON* markers.

processor with 1 GB DDR2 soldered RAM, and into which was added a *Microstrain 3DM-GX3-25* inertial measurement unit working at 200 Hz, a *Microsoft Kinect* camera, at 30 Hz, and a *Crossbow Cricket* receiver, at 1 Hz. The *AscTec Pelican* is a very versatile platform with the following characteristics[5]:

- Minimum take-off weight of 630 g;
- Maximum payload of 650 g;
- Maximum flight time with full payload of 15 minutes.

The *Microsoft Kinect* is a motion sensing device that consists in fact in a depth sensor with 11-bit VGA resolution and a RGB camera with 8-bit VGA resolution. The depth sensor is composed by an infrared pattern projector combined with a monochrome CMOS sensor, i.e., an active pixel sensor. Figure D.4 shows a small schematic of the components of this device.

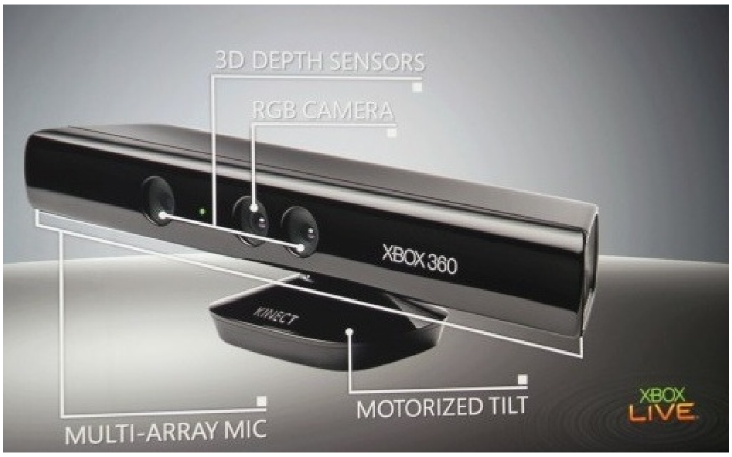


Figure D.4: A diagram of the technologies in the *Microsoft Kinect*. Reproduced from a slide in Microsoft's E3 Conference.

---

[5] `http://www.asctec.de/uav-applications/research/products/asctec-pelican/`

Cricket transceivers use a combination of radio and ultrasound technologies to compute distances between communicating nodes. With each RF advertisement, a Cricket in (active) beacon mode transmits an ultrasonic pulse. A Cricket in (passive) receiver mode will then be able to compute the time the ultrasonic pulse took to arrive, assuming instantaneous transmission of the RF information. The receiver then computes the actual distance using a temperature-aware method.

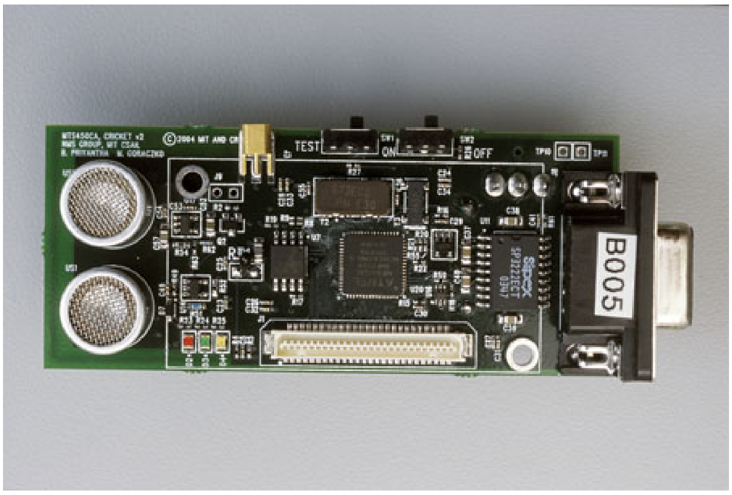


Figure D.5: A *Crossbow Cricket* receiver. Reproduced from `http://cricket.csail.mit.edu/`.

In the internal board of the quadrotor a version of the Linux distribution Ubuntu was installed, and on top of that, the Robot Operating System (ROS)[6], a framework for robot software development that provides hardware abstraction, device drivers, libraries, visualizers, message-passing, package management within other features, see [QCG+09] for further details. A driver for the inertial measurement unit developed by the Dynamical Systems and Ocean Robotics (DSOR) Lab team of the Institute for Systems and Robotics was employed, as well as community-based drivers[7] and software[8] for interfacing with the *Microsoft Kinect*, and a ROS package custom made by the author, for interfacing with the Cricket receiver and beacons (allowing the periodic cycling of beacon transmissions).

---

[6] `http://www.ros.org`

[7] `http://www.ros.org/wiki/openni_camera`

[8] `http://www.ros.org/wiki/openni_launch`

Bibliography

[ACS04] J. Andrade-Cetto and A. Sanfeliu. The effects of partial observability in SLAM. In *Proc. of the IEEE International Conference on Robotics and Automation*, volume 1, pages 397–402, Barcelona, Spain, April 2004. 1.3.1

[Adm16] Federal Aviation Administration. *Pilot's Handbook of Aeronautical Knowledge*. US Dept. of Transportation., 2016. 1.1

[AHB87] K.S. Arun, T. S. Huang, and S. D. Blostein. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987. 7.1

[AHWD11] A. Ahmad, Shoudong Huang, J. J. Wang, and G. Dissanayake. A new state vector for range-only SLAM. In *Proc. of the 2011 Chinese Control and Decision Conference (CCDC)*, pages 3404–3409, May 2011. 1.2.1

[AM79] Brian D.O. Anderson and John B. Moore. *Optimal Filtering*. Information and System Sciences Series. Prentice Hall, 1979. 8.2

[AM08] Karl Johan Astrom and Richard M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, Princeton, NJ, USA, 2008. 9.2

[AM10] Joshua N. Ash and Randolph L. Moses. Self-localization of sensor networks. In S. Haykin and K.J.R. Liu, editors, *Handbook on Array Processing and Sensor Networks*, Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. Wiley, 2010. 1.2.1, 5.1

[And71] B. D. O. Anderson. Stability properties of Kalman-Bucy filters. *Journal of the Franklin Institute*, 291(2):137–144, 1971. 2.3.1, 2.3.1

[And84] T.W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section Series. Wiley, 1984. 7.5

[Bai02] Tim Bailey. *Mobile robot localisation and mapping in extensive outdoor environments*. PhD thesis, Australian Centre for Field Robotics, The University of Sidney, New South Wales, Australia, 2002. 1.2.2

[Bai03] Tim Bailey. Constrained initialisation for bearing-only SLAM. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, volume 2, pages 1966–1971. IEEE, 2003. 1.2.1

[BB15] Axel Barrau and Silvère Bonnabel. Invariant filtering for Pose EKF-SLAM aided by an IMU. In *Proc. of the 2015 IEEE 54th Annual Conference on Decision and Control*, pages 2133–2138, Osaka, Japan, December 2015. 1.3.1

[BBF$^+$06] Andrea Bonarini, Wolfram Burgard, Giulio Fontana, Matteo Matteucci, Domenico Giorgio Sorrenti, and Juan Domingo Tardos. Rawseeds: Robotics advancement through web-publishing of sensorial and elaborated extensive data sets. In *Proceedings of IROS'06 Workshop on Benchmarks in Robotics Research*, 2006. 1.8, 6.1, 6.4.1, 6.6, 8.6(d), 8.5.2, 9.1

[BCK06] K.E. Bekris, M. Click, and E.E. Kavraki. Evaluation of algorithms for bearing-only SLAM. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 1937–1943, May 2006. 1.2.1

[BDW06] T. Bailey and H. Durrant-Whyte. Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics & Automation Magazine*, 13(3):108–117, 2006. 1.2

[BETG08] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, 2008. 1.2.1, 3.4.1, 4.6.1, 6.4.1, 8.5.2

[BFMG08] J.-L. Blanco, J.-A. Fernandez-Madrigal, and J. Gonzalez. Efficient probabilistic Range-Only SLAM. In *Proc. 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'08*, pages 1017–1022, September 2008. 1.2.1

[BG07] I. Borg and P.J.F. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. Springer, 2007. 7.1

[BH97] R.G. Brown and P.Y.C. Hwang. *Introduction to random signals and applied Kalman filtering*. Wiley, 1997. 8.2

[BJ09] A.N. Bishop and P. Jensfelt. A Stochastically Stable Solution to the Problem of Robocentric Mapping. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 1615–1622, Kobe, Japan, may 2009. 1.3.2

[BNG$^+$06] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot. Consistency of the EKF-SLAM Algorithm. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3562–3568, oct. 2006. 1.2.2, 1.3.1

[Bou15] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab. `http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/parameters.html`, October 2015. Online: version of October 14th, 2015. 6.4.1

[Bow02] Nathaniel Bowditch. *The American Practical Navigator*. Defense Mapping Agency Hydrographic Topographic Center, 2002. 1

[BPL12] Adrien Bartoli, Daniel Pizarro, and Marco Loog. Stratified Generalized Procrustes Analysis. *International Journal of Computer Vision*, 101(2):227–253, 2012. 7.1, 7.3

[Bro70] R.W. Brockett. *Finite Dimensional Linear Systems*. Series in decision and control. John Wiley & Sons, 1970. A.1

[BS92] C.A. Beattie and S.W. Smith. Optimal matrix approximants in structural identification. *Journal of Optimization Theory and Applications*, 74(1):23–56, 1992. 7.1

[BSC13] B. Bacca, J. Salvi, and X. Cufí. Long-term mapping and localization using feature stability histograms. *Robotics and Autonomous Systems*, 61(12):1539–1558, 2013. 1.2.2

[BSL$^+$11] Simon Baker, Daniel Scharstein, J.P. Lewis, Stefan Roth, MichaelJ. Black, and Richard Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, 2011. 4.6.1

[BSO10] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Optimal position and velocity navigation filters for autonomous vehicles. *Automatica*, 46(4):767–774, 2010. 3.3

[BSO11a] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. On the observability of linear motion quantities in navigation systems. *Systems & Control Letters*, 60(2):101–110, 2011. A.1

[BSO11b] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Single range aided navigation and source localization: Observability and filter design. *Systems & Control Letters*, 60(8):665–673, 2011. 1.6.1, i, 2.1, 2.2.1, 3.3, 4, 4.2.2, 4.3, 5.1, 5.2.2, 9.1

[BSO12] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. A GES Attitude Observer with Single Vector Observations. *Automatica*, 48(2):388–395, February 2012. 1.6.1

[BSO13a] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. GES source localization based on discrete-time position and single range measurements. In *Proceedings of the 21st Mediterranean Conference on Control Automation (MED)*, pages 1248–1253, June 2013. 2.3.2.1

[BSO13b] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Globally exponentially stable filters for source localization and navigation aided by direction measurements. *Systems & Control Letters*, 62(11):1065–1072, November 2013. 1.6.1, i, 6.1, 6.3, 9.1

[BSO14] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Sensor-based Long Baseline Navigation: observability analysis and filter design. *Asian Journal of Control*, 16(4):974–994, July 2014. i, 5.1, 5.3.2

[BWPJ11] Stephen Barkby, Stefan B. Williams, Oscar Pizarro, and Michael V. Jakuba. A featureless approach to efficient bathymetric slam using distributed particle mapping. *Journal of Field Robotics*, 28(1):19–39, 2011. 1.2

[CB02] Fabio Crosilla and Alberto Beinat. Use of generalised procrustes analysis for the photogrammetric block adjustment by independent models. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 56(3):195–209, 2002. 7.1

[CCC+16] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6):1309–1332, Dec 2016. 1.2

[CDKN+14] Luca Carlone, Jingjing Du, Miguel Kaouk Ng, Basilio Bona, and Marina Indri. Active SLAM and Exploration with Particle Filters Using Kullback-Leibler Divergence. *Journal of Intelligent & Robotic Systems*, 75(2):291–311, 2014. 9.2

[CDM08] Javier Civera, Andrew J. Davison, and J. M. Martínez Montiel. Inverse depth parametrization for monocular slam. *IEEE Transactions on Robotics*, 24(5):932–945, Oct 2008. 1.2.1, 6.3.1, 6.6.1

[CEBW14] J. Christmas, R.M. Everson, J. Bell, and C.P. Winlove. Inexact bayesian point pattern matching for linear transformations. *Pattern Recognition*, 47(10):3265–3275, 2014. 7.1

[CFG+09] Simone Ceriani, Giulio Fontana, Alessandro Giusti, Daniele Marzorati, Matteo Matteucci, Davide Migliore, Davide Rizzi, Domenico G. Sorrenti, and Pierluigi Taddei. Rawseeds ground truth collection systems for indoor self-localization and mapping. *Autonomous Robots*, 27(4):353–371, 2009. 1.8, 6.1, 6.4.1, 6.6, 8.6(d), 8.5.2

[CFJS00] Alessandro Chiuso, Paolo Favaro, Hailin Jin, and Stefano Soatto. 3-D motion and structure from 2-D motion causally integrated over time: Implementation. In *Proceedings of the European Conference on Computer Vision*, pages 734–750. Springer Berlin Heidelberg, 2000. 1.2.1

[CGDM10] Javier Civera, Oscar G. Grasa, Andrew J. Davison, and J. M. M. Montiel. 1-Point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry. *Journal of Field Robotics*, 27(5):609–631, 2010. 6.6.1

[CGM09] Javier Civera, Oscar García, and José M. Martínez Montiel. Benchmark Solution "Inverse Depth EKF based Visual Odometry using 1Point RANSAC". `http://www.rawseeds.org/rs/solutions/view/48`, November 2009. (document), 6.9

[CMC07] John L. Crassidis, F. Landis Markley, and Yang Cheng. Survey of Nonlinear Attitude Estimation Methods. *Journal of Guidance, Control, and Dynamics*, 30(1):12–28, January 2007. 8.3.1

[CMCTN07] J.A. Castellanos, R. Martinez-Cantin, J.D. Tardós, and J. Neira. Robocentric map joining: Improving the consistency of EKF-SLAM. *Robotics and Autonomous Systems*, 55(1):21–29, 2007. 1.2, 1.3.1, 1.4, 1.6.1, 3.1

[ČN04] K. Čapek and C. Novack. *R.U.R. (Rossum's Universal Robots)*. Penguin Classics. Penguin Books, 2004. 1

[Coo05] Aron J. Cooper. A Comparison of Data Association Techniques for Simultaneous Localization and Mapping. Master's thesis, Massachussets Institute Of Technology, Massachussetts, U.S.A., June 2005. 1.2.2

[CPKL12] Kiwan Choi, Jiyoung Park, Yeon-Ho Kim, and Hyoung-Ki Lee. Monocular SLAM with undelayed initialization for an indoor robot. *Robotics and Autonomous Systems*, 60(6):841–851, 2012. 1.2.1

[Cra05] J.J. Craig. *Introduction to Robotics: Mechanics and Control*. Addison-Wesley series in electrical and computer engineering: control engineering. Pearson/Prentice Hall, 2005. 2.1

[CSK+17] Gordon Christie, Adam Shoemaker, Kevin Kochersberger, Pratap Tokekar, Lance McLean, and Alexander Leonessa. Radiation search operations using scene understanding with autonomous UAV and UGV. *Journal of Field Robotics*, 2017. 1

[CUDW96] Michael Csorba, Jeffrey K. Uhlmann, and Hugh F. Durrant-Whyte. New approach to simultaneous localization and dynamic map building. In *Proc. of the SPIE Aerosense*, volume 2738, pages 26–36, 1996. 1.2

[DB15] G. Dubbelman and B. Browning. COP-SLAM: Closed-Form Online Pose-Chain Optimization for Visual SLAM. *IEEE Transactions on Robotics*, 31(5):1194–1213, Oct 2015. 1.2.1

[DNDW+01] G. Dissanayake, P. Newman, H.F. Durrant-Whyte, S. Clark, and M. Csobra. A solution to the simultaneous localisation and mapping (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3):229–241, 2001. 1.3.2

[Dor05] Leo Dorst. First order error propagation of the procrustes method for 3D attitude estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):221–229, February 2005. 1.6.2, 7.1, 7.4, 7.4.1

[DRMS07] A.J. Davison, I.D. Reid, N.D. Molton, and O. Stasse. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067, 2007. 1.2.1

[DS12] Joseph Djugash and Sanjiv Singh. Motion-aided network SLAM with range. *The International Journal of Robotics Research*, 31(5):604–625, 2012. 1.2.1, 5.1

[DW88] H.F. Durrant-Whyte. Uncertain geometry in robotics. *IEEE Journal of Robotics and Automation*, 4(1):23–31, feb 1988. 1.2

[DWB06] Hugh Durrant-Whyte and Tim Bailey. Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. *IEEE Robotics & Automation Magazine*, 13(2):99–110, 2006. 1.2

[EHE+12] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard. An Evaluation of the RGB-D SLAM System. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 1691–1696, St. Paul, MA, USA, May 2012. 1.2, 1.2.1

[FLS99] Hans Jacob S. Feder, John J. Leonard, and Christopher M. Smith. Adaptive mobile robot navigation and mapping. *The International Journal of Robotics Research*, 18(7):650–668, 1999. 9.2

[FPRARM15] Jorge Fuentes-Pacheco, José Ruiz-Ascencio, and Juan Manuel Rendón-Mancha. Visual simultaneous localization and mapping: a survey. *Artificial Intelligence Review*, 43(1):55–81, 2015. 1.2

[FW01] J.M. Fitzpatrick and J.B. West. The Distribution of Target Registration Error in Rigid-body Point-based Registration. *IEEE Transactions on Medical Imaging*, 20(9):917–927, September 2001. 1.8, 7.1, 7.4

[Gan60] Felix R. Gantmacher. *The theory of matrices*, volume 1. Chelsea Pub. Co., 1960. 2.1

[GBSO12] Bruno J. Guerreiro, Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Sensor-based Simultaneous Localization and Mapping - Part II: Online Inertial Map and Trajectory Estimation. In *Proceedings of the 2012 American Control Conference*, pages 6334–6339, Montréal, Canada, June 2012. 1.6.2

[GBSO13] Bruno J. Guerreiro, Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Globally Asymptotically Stable Sensor-based Simultaneous Localization and Mapping. *IEEE Transactions on Robotics*, 29(6):1380–1395, December 2013. 1.4, 1.6.1, i, 2.2, 6.3.2, 7.5, 9.1

[GD04] John C. Gower and Garmt B. Dijksterhuis. *Procrustes Problems*. Oxford Statistical Science Series. Oxford University Press, 2004. 2.5, 7.1, 7.3

[Gel74] A. Gelb. *Applied Optimal Estimation*. MIT Press, 1974. 2.3.2.2, 4.4, 5.4, B.1

[GH95] D. Goryn and S. Hein. On the estimation of rigid body rotation from noisy data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1219–1220, December 1995. 7.1

[GLGP13] Yunchao Gong, S. Lazebnik, A Gordo, and F. Perronnin. Iterative Quantization: A Procrustean Approach to Learning Binary Codes for Large-Scale Image Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2916–2929, December 2013. 7.1

[GMG+08] Michael A. Goodrich, Bryan S. Morse, Damon Gerhardt, Joseph L. Cooper, Morgan Quigley, Julie A. Adams, and Curtis Humphrey. Supporting wilderness search and rescue using a camera-equipped mini UAV. *Journal of Field Robotics*, 25(1-2):89–110, 2008. 1

[Goo91] Colin Goodall. Procrustes Methods in the Statistical Analysis of Shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991. 7.1, 7.3, 7.4, 8.2

[GSO14] Bruno J. Guerreiro, Carlos Silvestre, and Paulo Oliveira. Automatic 2-D LiDAR geometric calibration of installation bias. *Robotics and Autonomous Systems*, 62(8):1116–1129, 2014. 9.2

[Gue13] Bruno J. Guerreiro. *Sensor-based Control and Localization of Autonomous Systems in Unknown Environments*. PhD thesis, Instituto Superior Técnico, University of Lisbon, December 2013. 1.6.2

[Gui08] E. Guizzo. Three Engineers, Hundreds of Robots, One Warehouse. *IEEE Spectrum*, 45(7):26–34, July 2008. 1.6.3

[HD07] Shoudong Huang and Gamini Dissanayake. Convergence and Consistency Analysis for Extended Kalman Filter Based SLAM. *IEEE Transactions on Robotics*, 23(5):1036–1049, October 2007. 1.3.1, 1.3.2

[HD16] Shoudong Huang and Gamini Dissanayake. A critique of current developments in simultaneous localization and mapping. *International Journal of Advanced Robotic Systems*, 13(5):1–13, September 2016. 1.2

[HHN88] Berthold K. P. Horn, H.M. Hilden, and Shariar Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society America*, 5(7):1127–1135, 1988. 7.1

[HK77] R. Hermann and Arthur J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, October 1977. 1.3.2

[HMK07] Henry Huang, Frederic D. Maire, and Narongdech Keeratipranon. Bearing-only simultaneous localization and mapping for vision-based mobile robots. In Goro Obinata and Ashish Dutta, editors, *Vision Systems-Applications*, pages 335–360. I-Tech Education and Publishing, Vienna, Austria, 2007. 1.2.1

[HMR08] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis. Analysis and improvement of the consistency of extended Kalman filter based SLAM. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 473–479, Pasadena, CA, USA, May 2008. 1.3.1

[HMR10] Guoquan Huang, Anastasios I. Mourikis, and Stergios I. Roumeliotis. Observability-based rules for designing consistent ekf slam estimators. *The International Journal of Robotics Research*, 29(5):502–528, 2010. 1.3.1, 1.6.1

[HS97] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997*, pages 1106–1112, Jun 1997. 6.4.1

[HSS14] Kejun Huang, N.D. Sidiropoulos, and A Swami. Non-Negative Matrix Factorization Revisited: Uniqueness and Algorithm for Symmetric Decomposition. *IEEE Transactions on Signal Processing*, 62(1):211–224, January 2014. 7.1

[IPSE+14] Laura Igual, Xavier Perez-Sala, Sergio Escalera, Cecilio Angulo, and Fernando De la Torre. Continuous generalized procrustes analysis. *Pattern Recognition*, 47(2):659–671, 2014. 7.1, 7.3

[JB16] Tor A. Johansen and Edmund Brekke. Globally Exponentially Stable Kalman Filtering for SLAM with AHRS. In *Proceedings of the 19th International Conference on Information Fusion (FUSION 2016)*, pages 909–916, Heidelberg. Germany, July 2016. 1.6.1, 8.3.1

[JIS+11] E. Jauregi, I. Irigoien, B. Sierra, E. Lazkano, and C. Arenas. Loop-closing: A typicality approach. *Robotics and Autonomous Systems*, 59(3-4):218–227, 2011. 1.2.2

[JKFB06] P. Jensfelt, D. Kragic, J. Folkesson, and M. Bjorkman. A framework for vision based bearing only 3D SLAM. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 1944–1950, May 2006. 1.2.1

[JS11] Eagle Jones and Stefano Soatto. Visual-inertial navigation, mapping and localization: A scalable real-time causal approach. *The International Journal of Robotics Research*, 30(4):407–430, 2011. 1.2.1, 1.3.2, 3.1

[JU01] S.J. Julier and J.K. Uhlmann. A counter example to the theory of simultaneous localization and map building. In *Proc. of the 2001 IEEE Int. Conf. on Robotics and Automation (ICRA)*, volume 4, pages 4238–4243, Seul, South Korea, May 2001. 1.3.1

[Kab76] W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5):922–923, September 1976. 7.1

[Kal60] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960. 2.3.2

[Kan94] Kenichi Kanatani. Analysis of 3-D rotation fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):543–549, May 1994. 7.1

[KB61] Rudolph Emil Kalman and Richard S. Bucy. New results in linear filtering and prediction theory. *Transactions of the ASME–Journal of Basic Engineering*, 83(Series D):95–107, 1961. 2.3.1

[KBH08] D. Kingston, R. W. Beard, and R. S. Holt. Decentralized Perimeter Surveillance Using a Team of UAVs. *IEEE Transactions on Robotics*, 24(6):1394–1404, December 2008. 1

[KC11] O. Kermorgant and F. Chaumette. Multi-sensor data fusion in sensor-based control: Application to multi-camera visual servoing. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 4518–4523, May 2011. 1.4

[KD04] N. M. Kwok and G. Dissanayake. An efficient multiple hypothesis filter for bearing-only SLAM. In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 736–741, Sept 2004. 1.2.1

[KE13] A. Kim and R. M. Eustice. Real-Time Visual SLAM for Autonomous Underwater Hull Inspection Using Visual Saliency. *IEEE Transactions on Robotics*, 29(3):719–733, June 2013. 1.2.1

[Ken12] Farid Kendoul. Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. *Journal of Field Robotics*, 29(2):315–378, 2012. 1

[Kha02] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, 3rd edition, 2002. 2.3.1, 2.3.1, 2.3.1

[KKG09] A.T. Klesh, P.T. Kabamba, and A.R. Girard. Optimal path planning for uncertain exploration. In *Proc. of the 2009 American Control Conference*, pages 2421–2426, June 2009. 9.2

[KKM03] Shun'ichi Kaneko, Tomonori Kondo, and Atsushi Miyamoto. Robust matching of 3d contours using iterative closest point algorithm improved by m-estimation. *Pattern Recognition*, 36(9):2041–2047, 2003. 7.1

[KPL10] J. Katrasnik, F. Pernus, and B. Likar. A Survey of Mobile Robots for Distribution Power Line Inspection. *IEEE Transactions on Power Delivery*, 25(1):485–493, January 2010. 1

[KS11] Jonathan Kelly and Gaurav S Sukhatme. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. *The International Journal of Robotics Research*, 30(1):56–79, 2011. 1.2.1, 1.3.2, 3.1, 3.5.1

[LBF$^+$12] Zhengrong Li, Troy S. Bruggemann, Jason J. Ford, Luis Mejias, and Yuee Liu. Toward automated power line corridor monitoring using advanced aircraft control and multisource feature fusion. *Journal of Field Robotics*, 29(1):4–24, 2012. 1

[LDW91] J.J. Leonard and H.F. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. In *Proc. of the IEEE/RSJ International Workshop on Intelligent Robots and Systems (IROS)*, volume 3, pages 1442–1447, 1991. 1.1, 1.2

[LFR$^+$13] Stefan Leutenegger, Paul Furgale, Vincent Rabaud, Margarita Chli, Kurt Konolige, and Roland Siegwart. Keyframe-based visual-inertial slam using nonlinear optimization. In *Proceedings of Robotics: Science and Systems*, Berlin, Germany, June 2013. 1.2.1, 9.2

[LHD06] C. Leung, Shoudong Huang, and G. Dissanayake. Active SLAM using Model Predictive Control and Attractor based Exploration. In *Proc. of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5031, Oct 2006. 9.2

[LKPY11] K. Lekadir, N.G. Keenan, D.J. Pennell, and Guang-Zhong Yang. An Inter-Landmark Approach to 4-D Shape Extraction and Interpretation: Application to Myocardial Motion Assessment in MRI. *IEEE Transactions on Medical Imaging*, 30(1):52–68, January 2011. 7.1

[LLM08] Jun Liu, Xiangqian Liu, and Xiaoli Ma. First-order perturbation analysis of singular vectors in singular value decomposition. *IEEE Transactions on Signal Processing*, 56(7):3044–3049, July 2008. 7.4.1

[LLS05] T. Lemaire, S. Lacroix, and J. Solà. A practical 3D bearing-only SLAM algorithm. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2449–2454, Aug 2005. 1.2.1

[LM97] F. Lu and E. Milios. Globally Consistent Range Scan Alignment for Environment Mapping. *Autononmous Robots*, 4(4):333–349, October 1997. 1.2

[Low99] D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157, 1999. 1.2.1

[LTH$^+$15] B.W.-K. Ling, Nili Tian, C.Y.-F. Ho, Wan-Chi Siu, Kok-Lay Teo, and Qingyun Dai. Maximally decimated paraunitary linear phase fir filter bank design via iterative svd approach. *IEEE Transactions on Signal Processing*, 63(2):466–481, Jan 2015. 7.1

[Lüt96] H. Lütkepohl. *Handbook of Matrices*. John Wiley & Sons, 1996. B.3

[MA09] Mehdi Hedjazi Moghari and Purang Abolmaesumi. Distribution of target registration error for anisotropic and inhomogeneous fiducial localization error. *IEEE Transactions on Medical Imaging*, 28(6):799–813, June 2009. 7.1, 7.4

[MAMT15] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics*, 31(5):1147–1163, Oct 2015. 1.2.1

[MAMT17] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics*, 33(5):1255–1262, Oct 2017. 1.2.1

[Mas83] W. S. Massey. Cross products of vectors in higher dimensional euclidean spaces. *The American Mathematical Monthly*, 90(10):697–701, December 1983. 7.5

[MCMdD$^+$06] Luis Merino, Fernando Caballero, J.R. Martínez-de Dios, Joaquín Ferruz, and Aníbal Ollero. A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires. *Journal of Field Robotics*, 23(3-4):165–184, 2006. 1

[MDH$^+$09] Stefan May, David Droeschel, Dirk Holz, Stefan Fuchs, Ezio Malis, Andreas Nüchter, and Joachim Hertzberg. Three-dimensional mapping with time-of-flight cameras. *Journal of Field Robotics*, 26(11-12):934–965, 2009. 1.2.1

[MFA07] Guoqiang Mao, Barış Fidan, and Brian D.O. Anderson. Wireless sensor network localization techniques. *Computer Networks*, 51(10):2529 – 2553, 2007. 1.2.1

[MTKW02] Michael Montemerlo, Sebastian Thrun, Daphne Koller, and Ben Wegbreit. Fast-SLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem. In *Proc. of the AAAI National Conference on Artificial Intelligence*, pages 593–598, Edmonton, Alberta, Canada, 2002. 1.2

[Mus10] Naval Undersea Museum. Torpedo History - Part 1 - Historical Background. `http://www.history.navy.mil/museums/keyport/html/part1.htm`, June 2010. note = Accessed: 2017-02-27. 1

[MV06] Graham McNeill and Sethu Vijayakumar. Hierarchical procrustes matching for shape retrival. In *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 885–894, San Diego, CA, USA, June 2006. CVPR. 7.1

[MZZ$^+$09] E. Menegatti, A. Zanella, S. Zilli, F. Zorzi, and E. Pagello. Range-only SLAM with a mobile robot and a Wireless Sensor Networks. In *Proc. 2009 IEEE International Conference on Robotics and Automation, ICRA'09*, pages 8–14, May 2009. 1.2.1

[NT01] J. Neira and J.D. Tardós. Data Association in Stochastic Mapping Using the Joint Compatibility Test. *IEEE Transactions on Robotics and Automation*, 17(6):890–897, dec 2001. 1.2.2, 3.1, 4.6.1

[OK98] Naoya Ohta and Kenichi Kanatani. Optimal estimation of three-dimensional rotation and reliability evaluation. *IEICE Transactions on Information and Systems*, E81-D(11):1247–1252, 1998. 7.1

[OLT06] E. Olson, J.J. Leonard, and S. Teller. Robust range-only beacon localization. *IEEE Journal of Oceanic Engineering*, 31(4):949–958, 2006. 1.2.1

[PCB00] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The Cricket Location-Support System. In *Proc. of the Sixth Annual ACM International Conference on Mobile Computing and Networking (MOBICOM)*, August 2000. (document), 4.1, 4.1, 4.6.1

[PRSF00] Erwin Prassler, Arno Ritter, Christoph Schaeffer, and Paolo Fiorini. A short history of cleaning robots. *Autonomous Robots*, 9(3):211–226, December 2000. 1.6.3

[QCG$^+$09] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully B. Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*, 2009. D.2

[QSS$^+$16] Juntong Qi, Dalei Song, Hong Shang, Nianfa Wang, Chunsheng Hua, Chong Wu, Xin Qi, and Jianda Han. Search and rescue rotary-wing uav and its application to the lushan ms 7.0 earthquake. *Journal of Field Robotics*, 33(3):290–321, 2016. 1, 1.6.3

[RH10] Julian Ryde and Huosheng Hu. 3d mapping with multi-resolution occupied voxel lists. *Autonomous Robots*, 28(2):169–185, 2010. 1.2

[Ros06] M.E. Rosheim. *Leonardo's Lost Robots*. Springer, 2006. 1

[RRKB11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: An efficient alternative to SIFT or SURF. In *2011 International Conference on Computer Vision*, pages 2564–2571, Nov 2011. 1.2.1

[Rug96] Wilson J. Rugh. *Linear System Theory*. Prentice Hall, New Jersey, NJ, USA, 2nd edition, 1996. 2.2, 2.3

[S9̈3] Inge Söderkvist. Perturbation analysis of the orthogonal procrustes problem. *BIT Numerical Mathematics*, 33(4):687–694, 1993. 7.1

[SA68] L. M. Silverman and B. D. O. Anderson. Controllability, observability and stability of linear systems. *SIAM Journal on Control*, 6(1):121–130, 1968. 2.4

[SC70] Peter Schönemann and Robert Carroll. Fitting One Matrix to Another Under Choice of a Central Dilation and a Rigid Motion. *Psychometrika*, 35:245–255, 1970. 2.1, 2.4, 7.1

[SC86] Randall Smith and Peter Cheeseman. On the representation and estimation of spatial uncertainty. *International Journal of Robotics Research*, 5(4):56–68, December 1986. 1.1, 1.2, 9.2

[Sch66] Peter H. Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, 1966. 1.6.2, 7.1

[SD82] S. S. Sastry and C. A. Desoer. The robustness of controllability and observability of linear time-varying systems. *IEEE Transactions on Automatic Control*, 27(4):933–939, August 1982. 2.4

[SHCJ10] Qingjiang Shi, Chen He, Hongyang Chen, and Lingge Jiang. Distributed Wireless Sensor Network Localization Via Sequential Greedy Optimization Algorithm. *IEEE Transactions on Signal Processing*, 58(6):3328–3340, June 2010. 1.2.1

[Sib78] Robin Sibson. Studies in the robustness of multidimensional scaling: Procrustes statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(2):234–238, 1978. 7.1

[Sib79] Robin Sibson. Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):217–229, 1979. 7.1, 7.4

[SLL02] Stephen Se, David Lowe, and Jim Little. Mobile Robot Localization and Mapping with Uncertainty using Scale-invariant Visual Landmarks. *The International Journal of Robotics Research*, 21(8):735–758, 2002. 1.2, 1.2.1, 1.2.2

[SMD12] Hauke Strasdat, J.M.M. Montiel, and Andrew J. Davison. Visual slam: Why filter? *Image and Vision Computing*, 30(2):65–77, 2012. 1.2.1

[SMDL05] J. Solà, A. Monin, M. Devy, and T. Lemaire. Undelayed initialization in bearing only SLAM. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2499–2504, Aug 2005. 1.2.1

[SoA99] Wilhekn Schmidt and Heron of Alexandria. *Pneumatica et Automata*. Bibliotheca scriptorum Graecorum et Romanorum Teubneriana. B.G. Teubner, 1899. 1

[SR05] R. Sim and N. Roy. Global A-Optimal Robot Exploration in SLAM. In *Proc. of the 2005 IEEE International Conference on Robotics and Automation*, pages 661–666, April 2005. 9.2

[SSC90] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In Ingemar J. Cox and Gordon T. Wilfong, editors, *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag New York, Inc., New York, NY, USA, 1990. 1.2

[Tay71] E. G. R. Taylor. *The Haven-Finding Art: A History of Navigation from Odysseus to Captain Cook*. The Bodley Head Ltd, 1971. 1.1

[TBF05] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics*. The MIT Press, 2005. ii

[Tes98] Nikola Tesla. Method of and apparatus for controlling mechanism of moving vessels or vehicles, November 1898. US Patent 613,809. 1

[TGdDO14] A. Torres-González, J.R. Martinez de Dios, and A. Ollero. Exploiting multi-hop inter-beacon measurements in ro-slam. *Procedia Computer Science*, 32(0):1101 – 1107, 2014. The 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014). 1.2.1

[TGdDO17] Arturo Torres-González, José Ramiro Martinez de Dios, and Anibal Ollero. Range-only SLAM for robot-sensor network cooperation. *Autonomous Robots*, 2017. 4.1

[TGDO14] A. Torres-González, J.R.M.-D. Dios, and A. Ollero. Efficient robot-sensor network distributed seif range-only slam. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 1319–1326, May 2014. 1.2.1

[TI10] O. Tekdas and V. Isler. Sensor placement for triangulation-based localization. *IEEE Transactions on Automation Science and Engineering*, 7(3):681–685, July 2010. 9.2

[TLK$^+$04] S. Thrun, Y. Liu, D. Koller, A. Ng, and H. Durrant-Whyte. Simultaneous localization and mapping with sparse extended information filters. *The International Journal of Robotics Research*, 23(7-8):693–716, 2004. 1.2

[TM06] Sebastian Thrun and Michael Montemerlo. The graph slam algorithm with applications to large-scale mapping of urban structures. *The International Journal of Robotics Research*, 25(5-6):403–429, 2006. 1.2

[TMHF00] Bill Triggs, Philip F. McLauchlan, Richard I. Hartley, and Andrew W. Fitzgibbon. *Vision Algorithms: Theory and Practice*, volume 1883 of *Lecture Notes in Computer Science*, chapter Bundle Adjustment — A Modern Synthesis, pages 298–372. Springer, Berlin, Heidelberg, 2000. 1.2.1

[TNNL02] J.D. Tardós, J. Neira, P.M. Newman, and J.J. Leonard. Robust mapping and localization in indoor environments using SONAR data. *The International Journal of Robotics Research*, 21(4):311–330, 2002. 1.2

[TR06] Nikolas Trawny and Stergios I. Roumeliotis. A unified framework for nearby and distant landmarks in bearing-only SLAM. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 1923–1929. IEEE, 2006. 1.2.2

[TRB$^+$06] Christopher Taylor, Ali Rahimi, Jonathan Bachrach, Howard Shrobe, and Anthony Grue. Simultaneous localization, calibration, and tracking in an ad hoc sensor network. In *Proc. of the 5th International Conference on Information Processing in Sensor Networks*, pages 27–33. ACM, 2006. 1.2.1

[TSZ14] Jun Tang, Ling Shao, and Xiantong Zhen. Robust point pattern matching based on spectral context. *Pattern Recognition*, 47(3):1469–1484, 2014. 7.1

[Ull79] S. Ullman. The interpretation of structure from motion. *Proceedings of the Royal Society of London B: Biological Sciences*, 203(1153):405–426, 1979. 1.2.1

[Ume91] Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991. 7.1, 7.3, 7.3, 7.4.1, 7.4.1, 8.2

[VCBS$^+$07] T. Vidal-Calleja, M. Bryson, S. Sukkarieh, Alberto Sanfeliu, and J. Andrade-Cetto. On the Observability of Bearing-only SLAM. In *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, pages 4114–4119, April 2007. 1.6.1

[VIC] VICON Motion Systems. Essentials of motion capture v1.2. `http://www.udel.edu/PT/Research/MAL/essentials_of_motion_capture_v1_2.pdf`. D.1

[VRB00] A.C. Victorino, Patrick Rives, and J.-J. Borrelly. Localization and map building using a sensor-based control strategy. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 937–942, 2000. 1.4

[Wah65] Grace Wahba. Problem 65-1: A least squares estimate of satellite attitude. *SIAM Review*, 7(3):409, July 1965. 7.1

[WAL$^+$13] Stephan Weiss, Markus W. Achtelik, Simon Lynen, Michael C. Achtelik, Laurent Kneip, Margarita Chli, and Roland Siegwart. Monocular vision for long-term micro aerial vehicle state estimation: A compendium. *Journal of Field Robotics*, 30(5):803–831, 2013. 1.2.1

[WCN$^+$09] Brian Williams, Mark Cummins, José Neira, Paul Newman, Ian Reid, and Juan Tardós. A comparison of loop closing techniques in monocular SLAM. *Robotics and Autonomous Systems*, 57(12):1188–1197, 2009. 1.2.2

[WGM08] Markus Windolf, Nils Götzen, and Michael Morlock. Systematic accuracy and precision analysis of video motion capturing systems - exemplified on the vicon-460 system. *Journal of Biomechanics*, 41(12):2776–2780, 2008. D.1

[WLFP08] Andrew D. Wiles, Alexander Likholyot, Donald D. Frantz, and Terry M. Peters. A Statistical Model for Point-Based Target Registration Error With Anisotropic Fiducial Localizer Error. *IEEE Transactions on Medical Imaging*, 27(3):378–390, March 2008. 7.1, 7.1, 7.4

[WM08] Chang Wang and Sridhar Mahadevan. Manifold Alignment Using Procrustes Analysis. In *Proceedings of the 25th International Conference on Machine Learning*, pages 1120–1127, New York, NY, USA, 2008. ACM. 1.8, 7.1

[WSN87] Lee Weiss, Arthur C Sanderson, and C. P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE Journal on Robotics and Automation*, 3(5):404–417, October 1987. 1.4

[WTT$^+$07] Chieh-Chih Wang, Charles Thorpe, Sebastian Thrun, Martial Hebert, and Hugh Durrant-Whyte. Simultaneous localization, mapping and moving object tracking. *The International Journal of Robotics Research*, 26(9):889–916, 2007. 1.2.1

[XWL10] Longxu Xiao, Shihui Wei, and Hongbin Lin. Research on star-tracking correction technology for near-earth flight vehicles. *Systems Engineering and Electronics, Journal of*, 21(3):485–490, 2010. 1.1

[ZHD13] Liang Zhao, Shoudong Huang, and G. Dissanayake. Linear SLAM: A linear solution to the feature-based and pose graph SLAM based on submap joining. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 24–30, November 2013. 1.4

[ZWS$^+$17] T. Zhang, K. Wu, J. Song, S. Huang, and G. Dissanayake. Convergence and Consistency Analysis for a 3-D Invariant-EKF SLAM. *IEEE Robotics and Automation Letters*, 2(2):733–740, April 2017. 1.3.1

LIST OF PUBLICATIONS

[1] Pedro Lourenço, Pedro Batista, Paulo Oliveira, Carlos Silvestre, and C. L. Philip Chen. Sensor-based Globally Asymptotically Stable Range-Only Simultaneous Localization and Mapping. In *Proc. of the 52nd IEEE Conference on Decision and Control*, pages 5692–5697, Florence, Italy, December 2013. 1.7

[2] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. 3-D Inertial Trajectory and Map Online Estimation: Building on a GAS Sensor-based SLAM filter. In *Proc. of the 2013 European Control Conference*, pages 4214–4219, Zurich, Switzerland, July 2013. 1.7

[3] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Preliminary Results on Globally Asymptotically Stable Simultaneous Localization and Mapping in 3-D. In *Proc. of the 2013 American Control Conference*, pages 3093–3098, Washington D.C., USA, June 2013. 1.7

[4] Pedro Lourenço, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. A Globally Exponentially Stable filter for Bearing-Only Simultaneous Localization and Mapping in 3-D. In *Proc. of the 2015 European Control Conference*, pages 2817–2822, Linz, Austria, July 2015. 1.7

[5] Pedro Lourenço, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Simultaneous Localization and Mapping in Sensor Networks: a GES sensor-based filter with moving object tracking. In *Proc. of the 2015 European Control Conference*, pages 2359–2364, Linz, Austria, July 2015. 1.7

[6] Pedro Lourenço, Pedro Batista, Paulo Oliveira, Carlos Silvestre, and C. L. Philip Chen. Sensor-based Globally Exponentially Stable Range-Only Simultaneous Localization and Mapping. *Robotics and Autonomous Systems*, 68:72–85, June 2015. 1.6.1, 1.7

[7] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Simultaneous Localization and Mapping for Aerial Vehicles: a 3-D sensor-based GAS filter. *Autonomous Robots*, 40:881–902, June 2016. 1.6.1, 1.7

[8] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. New Design Techniques for Globally Convergent Simultaneous Localization and Mapping: Analysis and Implementation. In Thor I. Fossen, Kristin Y. Pettersen, and Henk Nijmeijer, editors, *Sensing and Control for Autonomous Vehicles: Applications to Land, Water and Air Vehicles*, Lecture Notes in Control and Information Sciences. Springer, 2017. 1.7

[9] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Uncertainty Characterization of the Orthogonal Procrustes Problem with Arbitrary Covariance Matrices. *Pattern Recognition*, 61:210–220, January 2017. 1.7

[10] Pedro Lourenço, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. A Globally Exponentially Stable filter for Bearing-Only Simultaneous Localization and Mapping with Monocular vision. *Robotics and Autonomous Systems*, 100:61–77, February 2018. 1.7

[11] Pedro Lourenço, Bruno J. Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Earth-fixed Trajectory and Map Online Estimation: Building on GES Sensor-based SLAM filters. under review. 1.7