Localisation et cartographie en simultané fiable, précise et robuste d'un robot sous-marin

Jérémy Nicola

THESIS / UNIVERSITY OF WESTERN BRITTANY
under the seal of the Brittany-Loire University
to obtain the title of
DOCTOR OF THE UNIVERSITY OF WESTERN BRITTANY
Branch: Robotics
Doctoral School Maths-STIC

presented by
Jeremy Nicola
Prepared at iXblue and ENSTA Bretagne
(Lab-STICC, FR)

-

Robust, Reliable and Precise Simultaneous Localization and Mapping for an underwater robot

Thesis defended on 18th September 2017
before the jury composed of :

Michel Kieffer, president of the jury
Professor, Centrale-Supélec, Paris, Fr

Nacim Ramdani, rapporteur
Professor, University of Orléans, Orléans, Fr

Louise Travé, rapporteur
Research Director, LAAS-CNRS, Toulouse, Fr

Luc Jaulin, thesis supervisor
Professor, Lab-STICC, Brest, Fr

Vincent Creuze, reviewer,
Associate Professor, LIRMM, Montpellier, Fr

Michel Legris, reviewer,
Professor, Lab-STICC, Brest, Fr

Benoit Zerr, reviewer,
Professor, Lab-STICC, Brest, Fr

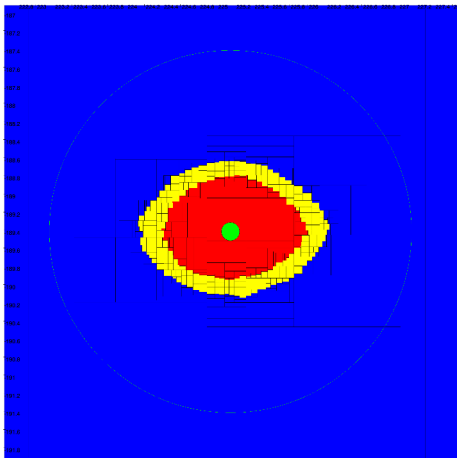
Sébastien Pennec, thesis co-supervisor
PhD, Signal Processing Research Engineer, iXblue, Brest, Fr

JEREMY NICOLA

ROBUST, PRECISE AND RELIABLE SIMULTANEOUS LOCALIZATION AND MAPPING FOR AN UNDERWATER ROBOT

ROBUST, PRECISE AND RELIABLE SIMULTANEOUS LOCALIZATION AND MAPPING FOR AN UNDERWATER ROBOT

JEREMY NICOLA



Comparison and combination of probabilistic and set-membership methods for the SLAM problem

September 2016 –

To my family.

ABSTRACT

In this thesis, we work on the problem of simultaneously localizing an underwater robot while mapping a set of acoustic beacons lying on the seafloor, using an acoustic range-meter and an inertial navigation system. We focus on the two main approaches classically used to solve this type of problem: Kalman filtering and set-membership filtering using interval analysis. The Kalman filter is optimal when the state equations of the robot are linear, and the noises are additive, white and Gaussian. The interval-based filter do not model uncertainties in a probabilistic framework, and makes only one assumption about their nature: they are bounded. Moreover, the interval-based approach allows to rigorously propagate the uncertainties, even when the equations are non-linear. This results in a high reliability in the set estimate, at the cost of a reduced precision.

We show that in a subsea context, when the robot is equipped with a high precision inertial navigation system, a part of the SLAM equations can reasonably be seen as linear with additive Gaussian noise, making it the ideal playground of a Kalman filter. On the other hand, the equations related to the acoustic range-meter are much more problematic: the system is not observable, the equations are non-linear, and the outliers are frequent. These conditions are ideal for a set-based approach using interval analysis.

By taking advantage of the properties of Gaussian noises, this thesis reconciles the probabilistic and set-membership processing of uncertainties for both linear and non-linear systems with additive Gaussian noises. By reasoning geometrically, we are able to express the part of the Kalman filter equations linked to the dynamics of the vehicle in a set-membership context. In the same way, a more rigorous and precise treatment of uncertainties is described for the part of the Kalman filter linked to the range-measurements. These two tools can then be combined to obtain a SLAM algorithm that is reliable, precise and robust. Some of the methods developed during this thesis are demonstrated on real data...

RÉSUMÉ

Dans cette thèse, nous nous intéressons au problème de la localisation et de la cartographie en simultané d'un robot sous-marin et d'un ensemble de balises disposées sur le fond marin à l'aide d'un distance-mètre et d'une centrale inertielle. Nous nous concentrons sur deux approches utilisées classiquement pour résoudre ce type de problème : le filtrage de Kalman et le filtrage ensembliste par intervalles. Le filtrage de Kalman est optimal lorsque les équations d'état du robot sont linéaires, et que les bruits sont additifs, blancs et Gaussiens. Le filtrage ensembliste par intervalles ne raisonne pas en termes de bruits probabilistes et ne fait qu'une seule hypothèse sur leur nature : ils sont bornés. De plus, l'approche par intervalles permet de propager les incertitudes de manière garantie, même lorsque les équations sont non-linéaires. Il en résulte une grande fiabilité de l'estimée ensembliste, au prix d'une précision réduite.

On montre que dans un contexte sous-marin, lorsque le robot est équipé d'une centrale inertielle de haute précision, une partie des équations d'état du SLAM peuvent être raisonnablement vues comme linéaires à bruit additif Gaussien, en faisant le terrain de jeu idéal pour un filtre de Kalman. En revanche, les équations liées au distance-mètre acoustiques sont elles beaucoup plus problématiques : le système est non-observable, les équations non-linéaires, et les mesures aberrantes fréquentes. Ces conditions sont idéales pour une approche ensembliste par intervalles.

En exploitant les propriétés des bruits Gaussiens, cette thèse réconcilie le traitement probabiliste et ensembliste des incertitudes pour les systèmes linéaires et non-linéaires à bruit additif Gaussien. Par un raisonnement géométrique, on parvient à exprimer la partie des équations du filtre de Kalman liée à la dynamique du véhicule dans un cadre ensembliste. De la même manière, un traitement plus rigoureux et précis des incertitudes est proposé pour la partie du filtre de Kalman liée aux mesures de distance. Ces deux outils peuvent par la suite être combinés pour obtenir un algorithme de SLAM fiable, précis et robuste. Certaines des méthodes développées au cours de cette thèse sont démontrées sur des données réelles...

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

[1] Vincent Drevelle and Jeremy Nicola. "VIBes: A Visualizer for Intervals and Boxes." In: *Mathematics in Computer Science* 8.3 (2014), pp. 563–572. DOI: 10.1007/s11786-014-0202-0. URL: http://dx.doi.org/10.1007/s11786-014-0202-0.

[2] L Jaulin, S Rohou, J Nicola, M Saad, F Le Bars, and B Zerr. "Distributed localization and control of a group of underwater robots using contractor programming." In: *SWIM 2015* (2015), p. 40.

[3] Luc Jaulin Jeremy Nicola. "Guaranteed localization of an underwater robot using bathymetry data and interval analysis: a preliminary study." MOQUESM 2014th edition. 2014.

[4] Luc Jaulin Jeremy Nicola. "Towards the hybridization of set-membership and probabilistic methods for the localization of an underwater vehicle." Summer Workshop on Interval Methods 2014th edition. 2014. URL: http://www.math.uu.se/swim2014/.

[5] Luc Jaulin Jeremy Nicola. "Gaussian, non-linear set-inversion." Summer Workshop on Interval Methods 2015th edition. 2015. URL: http://iti.mff.cuni.cz/series/2015/620.pdf#page=82.

[6] Luc Jaulin Jeremy Nicola. "OMNE is a Maximum Likelihood Estimator - with application to underwater mobile mapping." Summer Workshop on Interval Methods 2016th edition. 2016. URL: https://swim2016.sciencesconf.org/data/pages/Nicola_Jaulin.pdf.

[7] Fabrice Le Bars, Jeremy Nicola, and Luc Jaulin. "EASIBEX-MATLAB: a simple tool to begin with interval contractors." In: *SWIM 2015* (2015), p. 56.

[8] J. Nicola and L. Jaulin. "Guaranteed nonlinear parameter estimation with additive Gaussian noise." In: *Reliable Computing (accepted)* (2017).

[9] J. Nicola and L. Jaulin. "Marine Robotics and Applications." In: ed. by Luc Jaulin, Andrea Caiti, Marc Carreras, Vincent Creuze, Frederic Plumet, Benoit Zerr, and Annick Billon-Coat. Springer International Publishing, 2018. Chap. Comparison of Kalman and Interval Approaches for the Simultaneous Localization and Mapping of an Underwater Vehicle.

[10] Jan Sliwka, Jeremy Nicola, Remi Coquelin, Francois Becket de Megille, Benoit Clement, and Luc Jaulin. “Sailing without wind sensor and other hardware and software innovations.” In: *Robotic Sailing*. Springer Berlin Heidelberg, 2011, pp. 27–38.

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. Luc Jaulin for the continuous support of my research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to thank the iXblue Acoustic Products Business Unit in Brest and all the staff for giving me the opportunity to conduct research on the topic of underwater SLAM for the last three years, and for their so far unparalleled hospitality and welcomeness.

I would also like to thank my thesis committee: Prof. Nacim Ramdani, Prof. Louise Travé, Prof. Vincent Creuze, Prof. Michel Kieffer, Prof. Michel Legris, Prof. Benoit Zerr and and Sébastien Pennec, PhD, for their insightful comments and encouragements, but also for the challenging questions which incented me to widen my research from various perspectives.

Finally, I would like to thank the teaching staff at ENSTA Bretagne : when I was a student there, some of you have been sources of motivation. When I became a PhD student, some of you turned into mentors. Now we are colleagues. You rock.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GENERAL INTRODUCTION

The GPS does not work underwater. As every Global Navigation Satellite System (GNSS), the GPS relies on electromagnetic signals sent by artificial satellites, whose orbital parameters are known, to estimate the global 3D position of the receiver. However, high-frequency electromagnetic signals such as the ones used by the GPS (in the order of ~1GHz) do not propagate well through water. This does not only make GNSS unusable for subsea operations, but also rules out the possibility of wireless high communication rate with a submerged vehicle. Light, also made of electromagnetic waves, has a limited penetration in oceanic water, to the point where the sunlight will not penetrate at all beyond 1,000 meters in depth. If we add to that the pressure, which increases by about approximately one atmosphere for every 10 meters of water depth (and at 11,000 meters, the greatest ocean depth, that amounts to the weight of an elephant tiptoe on a stamp), it makes the ocean, together with the deep space, one of the most hostile and challenging environment for a robot to evolve in, localize itself and collect data on its surroundings.

All the same, the need for undersea positioning capability keeps growing decade after decade. This growth is caused by the continued development of fields such as oceanology, biology, wreck inspection, security, military or deep-water resource exploitation such as oil extraction or the emerging market of subsea mining. Given the aforementioned hostility of the oceanic environment, manned operations are not possible beyond a few tens of meters, and must then be delegated to robots. The most frequently type of robots used for such operations are ROV (Remotely Operated Vehicles), which are controlled by a human through the mean of a tether connecting him to the vehicle.

Localization is a primary capacity for any autonomous vehicle. Knowing where the robot is is fundamental to plan its trajectory and avoid known obstacles in order to achieve its goal. To localize itself, a robot must possess a representation of its environment, known as a map, for which there exist many types. Then, it interacts with its environment through sensors, and gathers *exteroceptive* measurements (i.e.: observations of the environment), as well as *proprioceptive* measurements (observations of values internal to the robot such as its specific acceleration, the battery level, the thrusters input voltage...), that are fed to a *localization algorithm* whose role is to estimate the state of the robot. When a map is not available, but the robot position is exactly known, it is the role of a *mapping* algorithm to create this map with sensors, also biased by noises and giving a partial information on the

map. Finally, when neither the map nor the localization of the robot are known, or some of these elements are only partially known, it is the role of a *simultaneous localization and mapping* (SLAM for short) algorithm to jointly estimate the robot's pose and the map of the environment from the sensor data.

As every measuring devices, sensors are affected by uncertainties: they never return a perfect information on the physical quantity they measure. The uncertainty on the measurement can come from environmental factors (the temperature, humidity, atmosphere's pressure...), from the sensor itself (limited resolution, noise in the electronic circuitry, misalignment of optical lenses in the case of a camera etc...). In many scenarios, a sensor only returns a partial information of the system of interest. A localization, mapping or SLAM algorithm must be therefore be able to handle those uncertainties, for which there exist several modelization, and the partial information gathered during the mission to estimate the desired parameters. In this work, we will consider the framework of *probabilistic* uncertainties and *unknown but bounded* uncertainties.

LOCALIZATION Equipped with a map, a robot will interact with its environment, to estimate its position with sensors. As all measurement system, a sensor can't return a perfect information on the state of the measured system. The uncertainty on the measurement can come from environmental factor (temperature, humidity, the atmosphere...), from the sensor itself (limited resolution, noise in the electronic circuitry, misalignment of optical lenses in the case of a camera...).

This measure will therefore be corrupted with an error or even be totally wrong. In most case, a single observation will bring a *partial* knowledge of the system. Measuring the position of an object on the image plane of a camera doesn't bring any information on the depth of this object. Measuring the distance from a sensor to an object will only tell us the object lies in a spherical shell around the sensor. It is the role of a localization algorithm to cope with uncertainties and combine several partial views of the system to provide an estimation of its pose.

MAPPING There are several types of map for representing an environment, depending on the application context. Such a map could be a road map [27], to enable the navigation of intelligent vehicles, an elevation map, used for guidance and control of missiles, submarines or terrestrial robots Desrochers, Lacroix, and Jaulin [23] and Drevelle [27], or the coordinates of some remarkable objects, called landmarks, which the robot will interact with to localize itself Davison et al. [22], Kantor and Singh [52], and Newman and Leonard [68]. Examples of landmarks are the GPS coordinates of lighthouses,

allowing a boat viewing several of these lighthouses to estimate its position by triangulation. It could also be the orbital parameters of navigation satellites Drevelle and Bonnifait [24], or some object that are easily identified in the environment such as a traffic sign, an LED or an acoustic speaker emitting a known signal.

SIMULTANEOUS LOCALIZATION AND MAPPING Simultaneous Localization and Mapping (SLAM) is one of the most fundamental problem in mobile robotics, and has been an actively researched topic since the 80s Smith, Self, and Cheeseman [87].

SLAM is the process of jointly estimating the map of the environment and the position of the vehicle. It is a chicken and egg problem: if a precise map of the environment is available, it should be easy to localize the vehicle (Localization), and if the vehicle is precisely localized it should be easy to build a map of its environment (Mapping). SLAM problem occurs when neither does the robot know its position, nor does he know the map of its environment. It is harder to solve than localization, as the map is not known, and mapping, as its pose is also unknown.

There are two approaches to SLAM: online and offline SLAM. Online SLAM is the problem of estimating the map and the robot pose at time t. Offline SLAM is the problem of estimating the map and the robot pose over the entire path of the robot, with all the measurements made during the mission available at a time.

In this thesis, we will focus on the online SLAM problem, as this is more suited for a real-time implementation that can be embedded in a sensor.

The vehicle is equipped with sensors and is therefore limited in its ability to measure physical quantities. Limitation arise from several factors. For example, a camera is limited in the resolution of its photosensor array and the quality and the alignment of its optics. Even if the optics were perfect or the array's resolution was infinite, they would not allow to see through walls, sense the depth of a specific object or measure the electric current running in a wire that would be in its field of view. Each sensor has a specific domain of application, and several of them should be combined to gather enough information to be able to extract the specific quantities we are interested in. Sensors are also affected by noise, which makes the measurements uncertain and limits the information that can be extracted.

A SLAM algorithm should be able to deal with those uncertainties and combine several sources of information all affected by different types of uncertainties. In this thesis, we will focus on two methods which allow for fusing several sources of information affected by uncertainties, that take two very different approaches in their representation of uncertainties: the Kalman Filter, and the Robust State Observer.

APPLICATION CONTEXT

Industrial application making use of subsea measurement devices are typically high-cost industries, such as the military industries or the oil and gas industries. The quantities estimated from the measurements gathered by the vehicle are used to position and design critical, expansive structures, and should satisfy two properties:

1. they should be precise: some application require that the error on the parameters to estimate should be less than centimetric.
2. they should be reliable, in the sense that the estimation method that provided the results should be trusted up to some risk-threshold: if the hypothesis and the models are right, so should the estimations be. If this is not the case, the estimation method should be able to detect a fault and its output should not be trusted.

Navigation

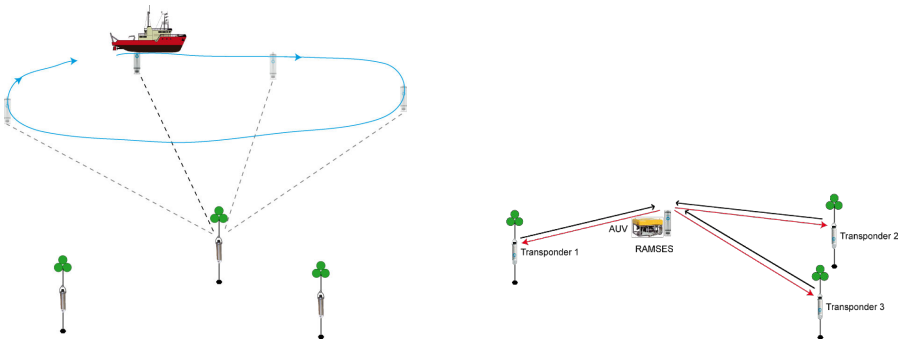
Subsea navigation is the field where an underwater vehicle evolves underwater. It includes the planning and the tracking of its trajectory, as well as its positioning.

In order to localize itself in the environment, the vehicle generally navigates in an array of acoustic beacons whose position is known from a previous calibration from the surface, as illustrated in Figure 0.1. However, a precise calibration of the sensor array takes time, and is costly, because of the surface vessel's time required to do it.

A cheaper approach is to perform a fast, gross pre-calibration of the sensors array, or use any other mean to get a reasonable initial guess on the sensors positions, and then use the underwater vehicle to refine the sensors positions at the same time as it uses them to position itself while performing its mission. That task of simultaneously navigating in a sensors array and calibrating that array is a typical SLAM task.

Metrology

The aim of subsea metrology is to estimate the relative positioning between two structures on the seabed. Generally, the objective of subsea metrology is to connect pipelines. The connecting piece has to be build on the surface, and the stresses on the connecting point is important, making for tight construction tolerances. Therefore, subsea metrology has to be both precise and reliable, so that the connecting pipes can be correctly designed, as if the connecting pipes do not meet precisely the tolerances, their life span can be significantly reduced and their replacement will induce delays in the production



(a) (b)

Figure 0.1: (a): Calibration of the sensors array from the surface, (b): Navigation in the sensors array with a subsea vehicle

as well as costs associated to the construction vessel's time to repair them.

The aim of this thesis is to propose precise, reliable and robust estimation tools that can be both used for navigation and metrology.

HANDLING UNCERTAINTIES

Probabilistic approach

Probabilistic robotics is a research topic that grew quickly in the last decade Thrun, Bugard, and Fox [89], and is the de-facto approach for many industrial application.

Set-membership approach

Also represents a whole space of guesses, but no probability is affected to each subset. Instead, each subset is classified as either compatible, not compatible, or undetermined with respect to the constraints of the problem.

There exist many representation for sets. Interval good nonlinear. Linear methods less subject to wrapping effect.

GOALS AND CONTRIBUTIONS OF THIS THESIS

As motivated in the Application Context Section page 4, in this thesis we are interested in a SLAM method that is reliable, precise and robust.

Reliability

A SLAM algorithm is said to be reliable if, given a model and a set of hypothesis for the problem, the estimates it provides are trustworthy,

i.e. if the true estimated variable is within a confidence region that contains all likely parameters up to a given risk threshold when the model and the hypothesis are correct. Additionally, under model or some hypothesis violation, a reliable algorithm should be able to detect when the estimates it provides do not contain the correct parameter anymore: such algorithm should have a fault-detection capacity. Reliability is therefore not the ability for an algorithm to provide correct estimates when the model is not correct or when some hypothesis are broken. For an algorithm to be reliable under model uncertainty or violable hypothesis, the possibility that the model is incorrect and that some hypothesis might be violated must be in the definition of the problem itself. To be able to survive hypothesis violations, an algorithm should be robust.

Robustness

In the context of this thesis, we will consider robustness of a SLAM algorithm in the sense of fault-tolerance. For example, imagine a sensor that provides the position of the robot with a precision of one meter. Due to environment constraints, there could be outliers in the measurements: the true position might be tens of meters away from the measured position. An hypothesis on the probability q that a measure is an outlier can be formulated. With this additional hypothesis in the definition of the problem and a computational mean to take this hypothesis into account, the algorithm will be robust up to q% of outliers. A desirable but not necessary property of a robust algorithm is the ability to identify a faulty measurements.

Precision

A SLAM algorithm is precise if it satisfies two conditions:

1. The error (the distance between the estimated parameters and the actual parameters) is small
2. The confidence region associated with the estimation should be tight

In that sense, a better term to qualify what we are trying to achieve would be the accuracy. There is generally no way to estimate Property 1 without knowing the actual parameter to estimate. This is why Property 2 is important, as it will reflect the quality of the estimation. A measurement that reports the position of a robot with an error of 10 centimeters with a confidence region 1 meter wide might be preferable to a measurement that reports the position of the robot with an error of 1 centimeter and a confidence region which is 10 meters wide, unless there is a way to know the error which generally implies the actual position is known.

Reliability, robustness and precision in SLAM methods

Now, probabilistic-based approaches to SLAM tend to be robust and precise, while approaches based on set-membership methods tend to be reliable and robust. Indeed, a probabilistic approach to uncertainties tend to cope well with the true nature of most uncertainties, and a correct use of them produces precise estimates. However, when the statistical distributions are not known ones, or when the models involved are nonlinear, probabilistic approaches are not reliable anymore. Indeed, in these scenarios, the distributions are approximated either with a second order approximation as in the Extended Kalman Filter, discretization of the distribution function as in the Unscented Kalman Filter and the Particle Filter, or a discretization of the search-space (grid-based approaches). None of these approaches are able to provide a guaranteed estimate in such conditions. Set-membership approaches based on Interval Analysis, on the other hand, are particularly well-suited for this type of problem. They are often seen as probability-agnostic, since they assume the perturbations are unknown but bounded, and Interval Analysis allow us to consider subsets of the search space all at once without losing a single solution. Since the image of an Interval by a nonlinear function is also an Interval, they provide a coherent framework for dealing with uncertainties. Interval Analysis provides a scheme of dealing with rare events such as outliers by computing the set of parameters compatible with all the observations except a given number q at most, by using relaxed intersection techniques. This allows for a Robust estimation. Nonetheless, when the true nature of the noises is stochastic, they generally perform worse than probabilistic approaches. Indeed, since the statistical properties of the noises are ignored, information are not modeled in the problem and they are sub-optimal.

Contribution of this thesis

Several approaches that combine probabilistic and set-membership approaches have been proposed for state estimation Abdallah, Gning, and Bonnifait [1], Chen, Wang, and Shieh [18], Neuland et al. [67], and Tran et al. [90]. The main contribution of this work is to provide a scheme combining probabilistic and set-membership approaches while strictly remaining in a set-membership framework. Therefore, the estimates provided will be reliable, since no approximations are made by linearizing the equations, guaranteed with a lower-bound on the probability that the true parameter vector is in the computed confidence region, robust thanks to q-relaxation techniques, and precise since the statistical properties of the perturbations will be modeled in the problem instead of being ignored.

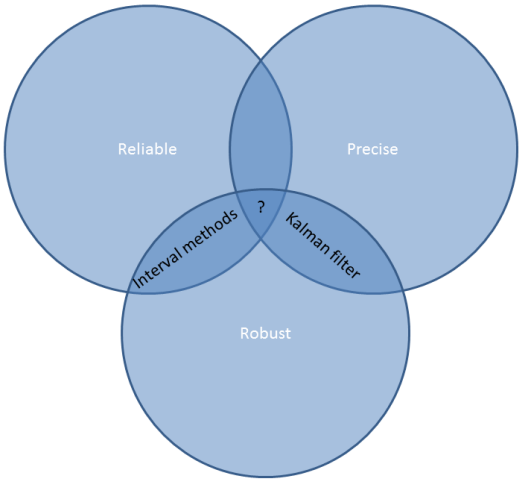


Figure 0.2: Venn diagram of reliable, precise and robust methods. We are looking for an approach that is both robust, reliable and precise.

DETAILED PLAN OF THE MANUSCRIPT

The first Part of this manuscript will present the general problem of simultaneously localizing and mapping objects underwater.

The first Chapter will be devoted to the modelling of the underwater SLAM problem. The classical SLAM formulation of the problem will be reviewed, and the different sensors used in underwater robotics will be presented, as well as the different types of positioning systems with a focus on long-baseline positioning systems (LBL). The sources of uncertainties for each of these sensors will be explored and discussed, and a convenient formulation of the problem will be proposed.

We will then present the state-of-the art algorithms used to solve our SLAM problem: the Kalman filter, a Bayesian state-estimator that belongs to the family of probabilistic state-estimators, and the Robust State Estimator (RSO) which relies on set-membership methods. The theoretical strengths and weaknesses of each approach will be discussed, and they will be compared on a real data-set. From both the theoretical and the practical analysis, we will outline the reasons why the set-membership approach fails at producing a precise estimate.

The second Part of this dissertation will be dedicated to the contributions introduced in this work.

In Chapter 3, we will focus on the equations of motion of the vehicle. The simplest case of motion equation under additive Gaussian noise will be reviewed, allowing us to sense the origin of the problem. The statistical properties of the motion noise will be discussed, and a geometrical constraint on those properties will be formulated. From this constraint, we will introduce the Kalman contractor, a contrac-

tor that reduces the search-space of the Cartesian product between two state-vectors at different times to a region that contains the true state-vectors with a minimum fixed probability. The gains obtained through this contractor will be discussed on the real data-set.

Chapter 4 will concentrate on the second set of equations for the SLAM problem: the observation equations. The observation equations model how the robot perceives its environment and therefore how it is able to dress a map of it. We will first look at these equations as a classical parameter estimation problem, and again examine the statistical properties of the observation noises. It will be shown that estimating n parameters with a probability threshold from a set of m observations subject to additive Gaussian noise amounts to computing the pre-image of an m-sphere through the observation equations. This problem is a typical set-inversion problem that can efficiently solved with Interval Analysis, even for non-linear models. The problem of inverting this sphere will be treated as-is, and an other scheme that relies on relaxed intersection techniques will be analyzed. These tools will be applied to real data in a static localization context, i.e. where we ignore the motion equation of the vehicle.

Now, these methods enable us to compute a set that contains the parameter vector of interest with a given probability, but we are also interested in a punctual estimation that is optimal according to some criterion. In Appendix C, the problem will be slightly modified to allow us to compute a Maximum Likelihood Estimation in a guaranteed manner.

Chapter 5 will conclude this thesis, and Appendix A will give a brief introduction to Interval analysis and contractor methods for the unfamiliar reader.

Part I

STATE OF THE ART IN RANGE-ONLY SLAM

In Part i, we review the state of the art in SLAM algorithms using range-only measurements. In Chapter 1, we will first model the SLAM problem for our specific use case, and in Chapter 2 we will review, compare and comment the classical algorithms used to solve the problem. From these observations, we will propose some improvements in Part ii.

1 MODELING THE SIMULTANEOUS LOCALIZATION AND MAPPING PROBLEM FOR AN UNDERWATER VEHICLE

1.1 INTRODUCTION

In this chapter we will first present the most commonly used sensors for solving the SLAM problem in a subsea context. Our SLAM problem will then be modelized and its specificity will be discussed. For a more complete state-of-the art in underwater robot modelization and perception, the reader is referred to [84].

1.2 NAVIGATION AND ACOUSTIC POSITIONING SYSTEMS

1.2.1 *Sensors used for the navigation of an underwater vehicle*

1.2.1.1 *Gyroscopes*

A gyroscope is a sensor that provides a measurement of the rotation speed. High-end gyroscopes, which are extremely precise (about $0.001°/s$), rely on the Sagnac effect: for a circular optical path revolving around its axis, a light beam traveling in the same direction as the rotation needs to travel more than one circumference around the ring before it attains its emission point. On the other hand, light traveling in the opposite direction needs to travel a shorter distance. By splitting a coherent light-beam in both direction and summing them up, a shift in the interference pattern is thus observed. The shift in the interference fringes is proportional to the optical path's angular velocity, which can therefore be measured. Using optical fibers, the optical path can be made arbitrarily long, enhancing the precision of the sensor. By combining three such gyroscopes, the full rotational motion of the robot can be estimated. Other types of gyroscopes exist, such as mechanical gyroscopes, mems-based gyroscopes or ring-laser gyroscopes.

1.2.1.2 *Accelerometers*

An accelerometer measures the instantaneous acceleration of its carrier. It can bee seen as a mass-spring system, as illustrated in Figure 1.2. Usually, three accelerometers are combined to measure the full specific acceleration (including the gravity) of the carrier.

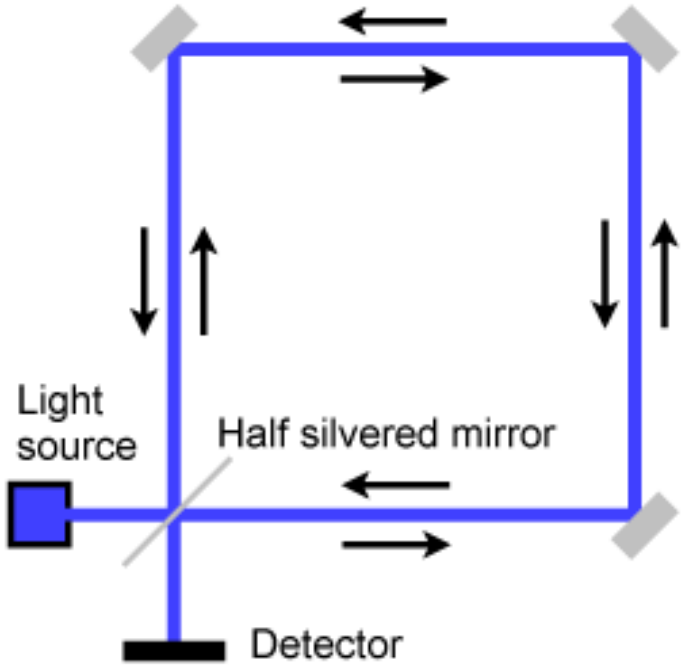


Figure 1.1: Illustration of Sagnac 's effect principle

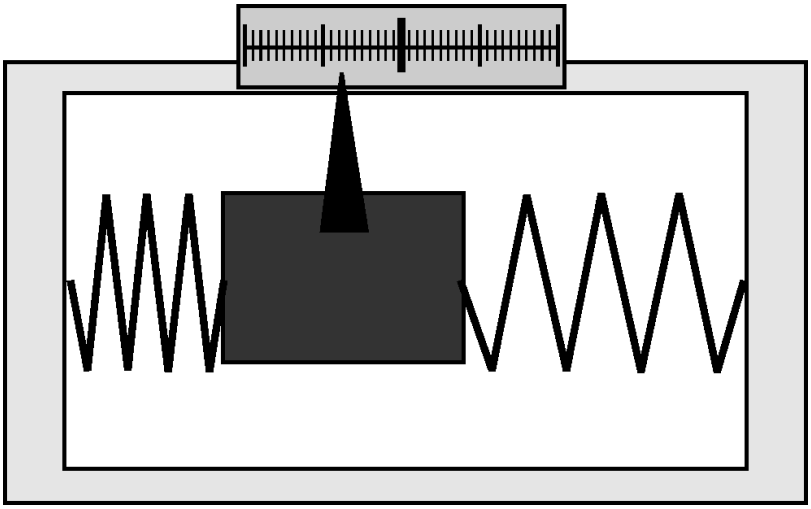


Figure 1.2: Accelerometer

Figure 1.3: Pressure sensor

1.2.1.3 *Pressure sensors*

A pressure sensor provides a measurement of the pressure applied by the water column plus the atmospheric pressure above the robot. With this information, knowing the density of the seawater, it is possible to retrieve the depth of the robot with a precision of about 1cm. For long-term navigation, the tide effects have to be taken into account and must be compensated.

1.2.1.4 *Doppler velocity log*

A Doppler Velocity Log (DVL for short) sensor projects a high-frequency acoustic signal on the seafloor, and measures the frequency shift induced on the reflected signal by the relative motion between the robot and the ground. From this frequency shift, it is possible to measure the velocity of the robot in a specific direction, and by combining several such projectors, the full velocity vector of the robot can be obtained with a precision of about 0.1 cm/s.

1.2.2 *Acoustic positioning system*

Acoustic positioning systems are sensors that use acoustic signals to retrieve information such as the bearing and/or the distance to another acoustic device (beacon).

1.2.2.1 *Ultra-Short Baseline*

A USBL is a sensor that combines several acoustic transducers whose relative positions are precisely known. One of these transducers emits a signal that is answered by a transponder. By measuring the time taken between the emission of the signal and the reception of the answer it is possible to estimate the distance between the system and the beacon. By also measuring the phase difference in the received signals in the transducers array, it is possible to estimate its angles of arrival, providing the full 3D position of the beacon in the sensor frame.

Figure 1.4: DVL. The four disks (three of which are visible here) project a high frequency acoustic signal on the seafloor and measure the received Doppler shift, allowing to measure the speed of its carrier in several directions.

Figure 1.5: The iXblue GAPS USBL positioning system

Examples of such systems are the POSIDONIA and the GAPS (illustrated on Figure 1.5), by iXblue. The later offers a pre-calibrated, tight coupling with an inertial measurement unit, to provide an accurate position of the transponder in the global frame.

1.2.2.2 *Short Baseline*

A short baseline positioning system is similar to an USBL system except for the greater distances in the transducer's array. SBL computes the distance between the transponder and each array's transducer, and unlike USBL, it computes the transponders position by trilateration. Therefore, the greater the distances in the transducer's array, the greater the precision in the transponder's positioning.

1.2.2.3 *Long Baseline*

A Long Baseline positioning System (LBL for short) is constituted of at least three transponders placed on the seabed, around the operation zone (up to $\sim$ 10km), whose absolute (or at least relative) positions are precisely known. The submarine emits an acoustic signal, and measures the time of flight of the signals answered by the transponders, and the submarine trilaterates its position. An advantage of LBL systems is that they are often deployed above the seabed, and thus avoid reflection of the signals on the water surface. Their main inconvenient is that it sometimes takes a long time to calibrate (determine the relative of absolute positions of the beacons) precisely the baseline. Due to the fact that it requires only one acoustic transducer, an LBL positioning devices such as the RAMSES developed by iXblue is cheaper than an USBL and also smaller, making it easier to integrate on a subsea vehicle. LBL positioning devices,and the RAMSES specifically, coupled with inertial navigation systems described in the next paragraph will be the center of focus of this PhD thesis.

1.2.3 *Inertial navigation system*

An Inertial Navigation System is a sensor that combines accelerometers and gyroscopes to estimate the motion of a rigid body without external reference. Sometime, they also embed a three-axis magnetometer and an estimator, often a Kalman filter, to provide an estimation of the attitude of the vehicle. In that case, The INS is referred as an AHRS (Attitude and Heading Reference System). It can also merge other information (as the speed measured from a DVL) to provide a more accurate estimation.

Figure 1.6: The iXblue RAMSES LBL positioning system

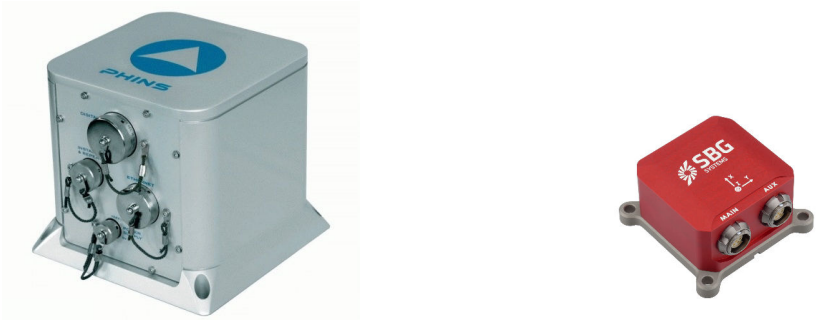


Figure 1.7: On the left an iXblue PHINS INS relying on FOG gyroscopes, on the left an SBG Systems Ellipse-E INS relying on MEMS based gyroscopes.

1.3 PROBLEM MODELIZATION

In this thesis, we will consider the case of an underwater vehicle equipped with an LBL positioning device, a DVL and a high-grade INS. Its goal will be to localize itself while localizing a set of beacons lying on the seabed. This is really interesting in an operational context, where ship time is extremely expensive. Reducing the time taken by a surface vessel to calibrate the LBL array, as illustrated in Figure 0.1, and instead report that task during the navigation phase represents important costs saving. The process of localizing a set of objects while localizing the vehicle that collects the measurements is known as "Simultaneous Localization And Mapping", or SLAM, a problem widely studied in the robotics scientific community as recalled in the General Introduction of this manual.

SLAM is a typical state estimation problem described by the following state equations:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{f}_k\left(\mathbf{x}_k, \mathbf{u}_k\right) \\ \mathbf{y}_k &= \mathbf{g}_k\left(\mathbf{x}_k\right) \end{cases} \tag{1.1}$$

where k is the time, $\mathbf{x} = (\mathbf{x}_m, \mathbf{x}_{b_1}, \mathbf{x}_{b_2}, \ldots, \mathbf{x}_{b_n})$ is the state vector, containing the position $\mathbf{x}_m$ of the mobile and the beacons $\mathbf{x}_{b_i}$ and $\mathbf{y}$ the observations. The first equation is the evolution function, modeling the dynamic part of the problem, the second one is the observation function that models the exteroceptive measurements the vehicle makes on its environment.

1.3.1 *Motion equation*

When the robot is equipped with a high-precision INS and a DVL, as will be the case in the framework of this thesis, the input vector $\mathbf{u}$ can be considered to be its Euler angles (ϕ, θ, ψ) and its velocity $\mathbf{v} = (v_x, v_y, v_z)$. A simple model of the robot motion can then be used:

$$\dot{\mathbf{x}}_m = \mathbf{R}_{\phi,\theta,\psi} \cdot \mathbf{v} \tag{1.2}$$

where $\mathbf{R}_{\phi,\theta,\psi}$ is the rotation matrix that maps the velocity vector $\mathbf{v}$ measured in the robot frame to the global frame.

This model is a kinematic model, as no second order derivatives of the state variables are involved. For a dynamic modelization of the robot the reader can refer to [32]. An Euler discretization yields the following evolution equation:

$$\mathbf{x}_{m_{k+1}} = \mathbf{x}_{mk} + \delta t \cdot \mathbf{R}_{\phi,\theta,\psi} \cdot \mathbf{v} + \boldsymbol{\omega}_k \tag{1.3}$$

where $\boldsymbol{\omega}_k$ is some noise.

As the landmarks are tied on the seafloor, we have the following evolution equation for the ith landmark:

$$\mathbf{x}_{b_{i_{k+1}}} = \mathbf{x}_{b_{i_k}}. \tag{1.4}$$

1.3.1.1 *Uncertainties on the robot motion*

The Euler angles are provided by the INS with a high precision, in the order of magnitude of 0.001deg. In this work we will assume the main source of uncertainties is the DVL, and we will consider the noise on the estimated Euler angles to be negligible in front of the DVL noise. This assumption will allow us to later use a Kalman filter in nominal conditions, as no multiplicative noise will be involved.

The uncertainties are therefore modeled as a centered Gaussian perturbation $\boldsymbol{\alpha}_v$ with covariance matrix $\boldsymbol{\Gamma}_{\alpha_v}$that comes from the DVL. The final robot evolution equation will be:

$$\mathbf{x}_{m_{k+1}} = \mathbf{x}_{mk} + \delta t \cdot \mathbf{R}_{\phi,\theta,\psi} \cdot \mathbf{v} + \delta t \cdot \mathbf{R}_{\phi,\theta,\psi} \cdot \boldsymbol{\alpha}_v. \tag{1.5}$$

Since the knowledge of $\mathbf{R}_{\phi,\theta,\psi}$ is supposed to be perfect, we can simplify our model by considering a centered, Gaussian perturbation $\boldsymbol{\alpha}$ with covariance matrix $\boldsymbol{\Gamma}_\alpha = \delta t^2 \cdot \mathbf{R}_{\phi,\theta,\psi} \cdot \boldsymbol{\Gamma}_{\alpha_v} \cdot \mathbf{R}^T_{\phi,\theta,\psi}$ corresponding to the DVL speed measurement noise projected in the global frame, and writing:

$$\mathbf{x}_{m_{k+1}} = \mathbf{x}_{mk} + \delta t \cdot \mathbf{R}_{\phi,\theta,\psi} \cdot \mathbf{v} + \boldsymbol{\alpha}. \tag{1.6}$$

In subsequent part of the thesis, we will also sometime consider that the input of the system is the mobile's velocity directly expressed in the global frame, noted $\boldsymbol{V}$ for example:

$$\mathbf{x}_{m_{k+1}} = \mathbf{x}_{mk} + \delta t \cdot \boldsymbol{V} + \boldsymbol{\alpha}. \tag{1.7}$$

1.3.2 *Observation equation*

The RAMSES, every time of flight acoustic measuring device, emits an interrogation at time t_e and receives the answer of the ith transponder at time t_{r_i}. By measuring the time elapsed between the instant t_e when the interrogation is sent by the RAMSES, and the instant t_{r_i} when the answer of the beacon is received, and assuming a constant sound velocity c, we have the following observation equation that links the the position of the robot, the position of the i^{th} beacon and the instants t_e, t_{r_i} :

$$y_i = \sqrt{\left(x_m - x_{b_i}\right)^2 + \left(y_m - y_{b_i}\right)^2 + \left(z_m - z_{b_i}\right)^2} = \frac{c}{2} \cdot \left(t_{r_i} - t_e - \delta_{TAT}\right) \tag{1.8}$$

with δ_{TAT} the time elapsed between the instant when the transponder receivers its interrogation and the instant where it emits its answers.

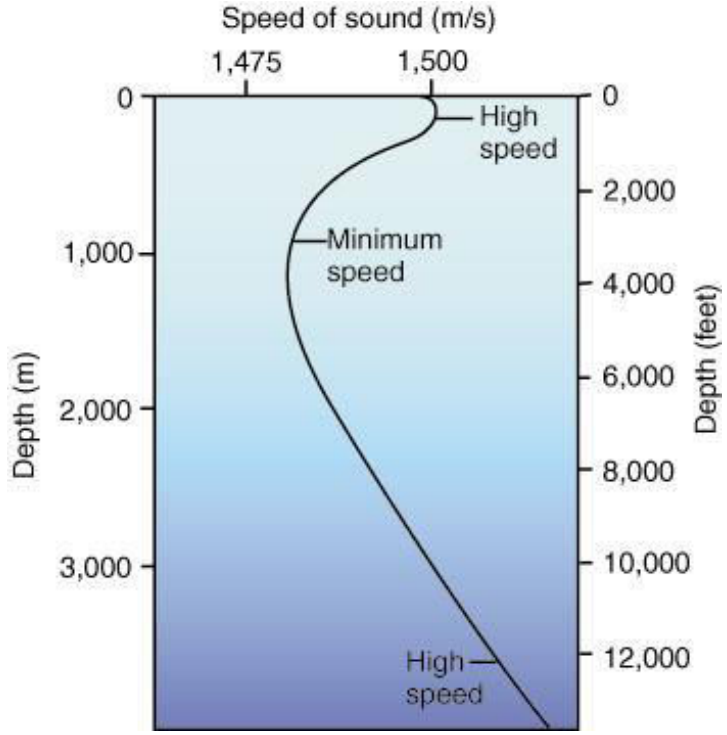


Figure 1.8: An example of sound velocity profile: the velocity varies with the water depth.

1.3.2.1 *Uncertainties on the time-of-flight measurements*

The proposed model is naive. The underwater acoustic channel is highly non-linear, dispersive, and suffers from multi-path propagation, frequency fading and many other effects. The electronic circuitry may also induce uncertainties in the time of flight measurements.

PROPAGATION DELAYS There are uncertainties in the time it takes for the signal to go through the hardware and the transducer before reaching the propagation channel. However, these uncertainties can be precisely determined in the laboratory.

VELOCITY PROFILE The assumption that the celerity is fixed is gross [17]. Ocean water is a dispersive medium, which means it impacts the speed of the acoustic signal. A typical approach is to use a layered model: the celerity is assumed to depend only on the immersion, and is constant between two layers z_i and z_{i+1}. When the sound speed velocity is dependent on the depth it has two different effects:

1. The angle of reception of the signal is affected
2. The trajectory of the acoustic signal is modified (it is not a straight line anymore and there might be multiple path connecting the emitter and the receiver)

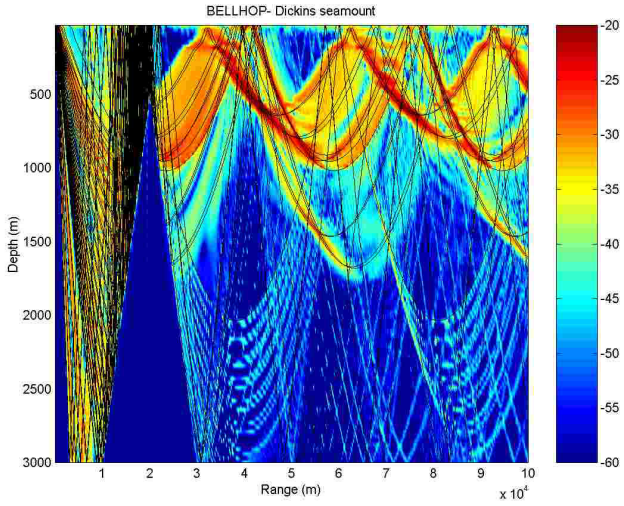


Figure 1.9: Example of paths that could be taken by an acoustic signal. Some are reflected on the seafloor or the water surface, and the dispersivity of the channel is obvious. The figure has been generated by BELLHOP [85], an acoustic ray-tracing program.

For the scope of this thesis, we will consider the case when the submarine is approximately in the same velocity layer as the beacons, and the distance between the submarine and the beacons is less than a kilometer. Therefore, the model described in Equation 1.8 is reasonable as a first order approximation, and we will not use more complex methods for computing the path traveled by the acoustic signals.

MULTI-PATH The path traveled by the acoustic signal may not be the shortest path between the source and the transponder: it may be reflected on the seafloor, the water surface, or simply result from the fact that the acoustic channel is dispersive and there may exist several acoustic paths between the transponders.

MEASUREMENT NOISES The received signal must be amplified, and therefore the noise will also be. We will consider that those noises are filtered out by the electronics at a low-level, early in the processing chain of the signals.

DOPPLER SHIFT The relative motion between the robot and the transponder may induce a Doppler shift in the signal, and "interfere" with the detection process. A Doppler shift might for example distort the data association of the acoustic signal, in which case the time of flight will be considered to come from the wrong beacon.

In this thesis we will use the constant celerity hypothesis, and model all aforementioned sources of uncertainties as an additive perturbation β, resulting in the observation model of Equation :

$$y_i = \frac{2}{c} \cdot \sqrt{\left(x_m - x_{b_i}\right)^2 + \left(y_m - y_{b_i}\right)^2 + \left(z_m - z_{b_i}\right)^2} + \delta_{TAT} + \beta. \quad (1.9)$$

β will be modeled as a centered Gaussian distribution whose standard deviation grows linearly with the measured time-of-flight (0.03 for example).

1.4 CONCLUSION

In this chapter, we presented the class of sensors that have been used in the scope of this work, and described their main sources of uncertainties. It must be noted that there exist many other sensors in a subsea context, such as echo-sounders or multi-beam echo-sounders. For a more detailed review of these sensors, the reader is referred to [84]. Our robot will be equipped with a high-precision INS, a DVL, a pressure-sensor, and a LBL acoustic positioning device. A simple but effective model of motion for our robot was proposed, as well as a simple observation model for measuring the time-of-flight between the robot and the beacons. In the next section, we will present the most commonly used method to solve the problem of simultaneously localizing a robot and a set of objects using time-of-flight sensors.

2

CLASSICAL APPROACHES TO SOLVE THE SLAM PROBLEM

2.1 INTRODUCTION

SLAM is a typical estimation problem which, as many estimation problem, faces the issue of dealing with uncertainties. From imperfect, partial, and sometimes erroneous measurements, a SLAM algorithm has to provide an estimation of the parameter vector of interest. The first key to solving this problem is to chose how to model uncertainties. The second one is to chose how to deal with uncertainties. In this thesis, we will focus on the two main approaches to dealing with uncertainties: probabilities and sets. This chapter is organized as follow. First, principles of probabilities to treat uncertainties in the SLAM problem will be recalled. From this, we will recall the Kalman filter algorithm [50] and its adaptations to deal with nonlinear problems [74, 94]. Secondly, we will present the set-membership approach for dealing with uncertainties. We will show that when the uncertainties are not stochastic anymore but unknown and bounded, the SLAM problem can be formulated as a set-inversion problem, which is a typical use case for set-membership methods [43][79]. Furthermore, an interval representation of these uncertainties allows for a propagation of uncertainties. The q-relaxed intersection, a set-theoretic operator, will be introduced, which allows to compute a solution set even when there are outliers in the measurements, whose maximum number is predefined. Finally, a comparison of solving the SLAM problem with a Kalman filter and a set-membership state estimator based on interval analysis will be made on a real data-set, and the results will be discussed [70].

2.2 PROBABILISTIC APPROACHES

Probabilistic approaches model uncertainties explicitly using the calculus of probability theory. Instead of relying on a punctual estimation, probabilistic estimation algorithm represent information by probability distributions. Such a representation allows for representing ambiguities and computing a likelihood associated to a solution. When the models are linear and the distributions are well known (Gaussian for example), probabilistic methods are easy to manipulate. However, when the probability density functions are not known, or the model are nonlinear (as is often the case on real world systems), the resulting distributions are not easily represented by parametrized

functions. When this kind of situation arises, strategies must be put in place to approximate the true probability distributions. A strategy can be to discretize the parameter space and associate a likelihood to each cell of the resulting array. As the measurements arrive, the array is updated and the probability distribution is refined. This type of approach is known as "grid-based", and works well but is subject to the curse of dimensionality: when the size of the parameter vector to estimate grows, the memory used to represent the distribution and the computer time to update it grows exponentially with that size. An other non-parametric approach known as "particle filter" also consists of search space discretization, except that instead of discretizing the whole search space, the distribution is approximated by a set of states, called particles, with a likelihood associated to each. Such a representation also allows for representing a broad set of distributions, and is more flexible in the sense that particles can be concentrated on an area of the search-space that is more likely to contain the true solution instead of being spread over the entire search space. Also, it is easy to propagate this sample based representation through nonlinear transformation. Finally, the approach often taken is to approximate the distributions by Gaussian distributions, which are easy to represent and manipulate, since they are only parametrized by their mean and covariance matrices. Even though the image of a Gaussian distribution through a nonlinear function might not be a Gaussian distribution anymore, a Gaussian approximation can be computed either by linearization of the function, or by using more complex schemes such as the Unscented Transform. This class of SLAM algorithm is known as Kalman filters, and will be the our reference probabilistic method in this thesis.

2.2.1 *Kalman filtering*

The Kalman filter [51] is the most widely used and studied technique for state estimation. It is a two step algorithm that applies to linear systems subject to additive, white Gaussian noise of the form:

$$\begin{cases} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k \cdot \boldsymbol{x}_k + \boldsymbol{B}_k \cdot \boldsymbol{u}_k + \boldsymbol{\alpha}_k \\ \boldsymbol{y}_k &= \boldsymbol{C} \cdot \boldsymbol{x}_k + \beta_k \end{cases} \tag{2.1}$$

where $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are respectively the evolution, command and observation matrices, and where $\boldsymbol{\alpha}_k, \boldsymbol{\beta}_k$ are some centered, independent Gaussian perturbation happening at instant k. With $\hat{\boldsymbol{x}}_0, \boldsymbol{\Gamma}_0$ an initial guess of the esperance and covariance matrix of the Gaussian distribution followed by $\boldsymbol{x}$ initially, and $\boldsymbol{\Gamma}_\alpha, \boldsymbol{\Gamma}_\beta$ the covariance matrices of the noises $\boldsymbol{\alpha}, \boldsymbol{\beta}$, the Kalman filter is able to estimate the esperance and covariance $\hat{\boldsymbol{x}}_k, \boldsymbol{\Gamma}_k$ of the distribution followed by $\boldsymbol{x}$ at any instant k. To do so, it acts in a two step process: the prediction, where the proprioceptive

measurements are integrated to project the current probability distribution at instant k onto the next instant $k+1$, and an update step, where an exteroceptive measurement $\boldsymbol{y}_k$, that brings some knowledge on the actual state $\boldsymbol{x}_k$ is processed to reduce the uncertainty on its estimate.

PREDICTION EQUATION The Kalman filter prediction equation can be written as:

$$\begin{cases} \hat{\boldsymbol{x}}_{k+1} &= \boldsymbol{A}_k \cdot \hat{\boldsymbol{x}}_k + \boldsymbol{B}_k \cdot \boldsymbol{u}_k \\ \boldsymbol{\Gamma}_{k+1} &= \boldsymbol{A}_k \cdot \boldsymbol{\Gamma}_k \cdot \boldsymbol{A}_k^T + \boldsymbol{\Gamma}_\alpha \end{cases}. \tag{2.2}$$

It can be seen that the Kalman filter prediction equation simply computes the image of a Gaussian distribution through an affine function.

UPDATE EQUATION When a measurement $\boldsymbol{y}_k$ on the robot is made at time k, that new knowledge is integrated in the estimate by the use of the update equation:

$$\begin{cases} \hat{\boldsymbol{x}}_k &= \hat{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k \cdot (\boldsymbol{y}_k - \boldsymbol{C}_k \cdot \hat{\boldsymbol{x}}_{k-1}) \\ \boldsymbol{\Gamma}_k &= (\boldsymbol{I} - \boldsymbol{K}_k \cdot \boldsymbol{C}_k) \cdot \boldsymbol{\Gamma}_{k-1} \\ \boldsymbol{K}_k &= \boldsymbol{\Gamma}_k \cdot \boldsymbol{C}_k^T \left(\boldsymbol{C}_k \cdot \boldsymbol{\Gamma}_{k-1} \cdot \boldsymbol{C}_k + \boldsymbol{\Gamma}_\beta\right)^{-1} \end{cases} \tag{2.3}$$

DRAWBACK OF THE KALMAN FILTER The Kalman Filter is optimal for linear systems subject to white, additive Gaussian noises. No optimality can be proved for systems which are subject to different kind of noises. Also, most systems encountered in real life are nonlinear. The system can then be linearized in its evolution equation, its observation equation or both in order to perform a linear first order approximation of the system, in a way that enables us to apply the Kalman filter equations. Such method is known as the Extended Kalman filter (EKF for short). An other method consists, instead of approximating the function with its tangent, to approximate directly the first and second order moments of the distributions by propagating a set of carefully chosen hypotheses through the nonlinear models. The hypothesis are then merged to recover the mean and covariance matrices. This method is referred as the Unscented Kalman Filter [49, 94] (UKF for short), and can be thought of a special case of particle filter, except that instead of randomly choosing the hypothesis, there are rules to select the points to ensure the second order approximation of the distribution is precise, by making use of the Unscented Transform.

The Kalman filter suffers from the fact that it has to maintain a correlated map in memory, in the form of the covariance matrix $\boldsymbol{\Gamma}$. As the number N of landmarks grows, so does the state vector (linearly with N) and the covariance matrix (quadratically with N). While this is a

severe limitation in scenarios where there is a huge number of landmarks, such as large-scale visual SLAM, in an underwater context where the number of beacons is reasonable that quadratic growth is sustainable. However, in a subsea environment $\mathbf{\Gamma}$ has high chances to be sparse: the more distant the beacons are from each others, the less they are correlated. In that case, sparse methods such as the Sparse Information Filter [89], which can be seen as the complementary of the Kalman filter might be a better approach.

Additionally, in its extended form, the Kalman filter makes use of the current state estimate $\hat{\boldsymbol{x}}_k$ to compute the Kalman gain $\boldsymbol{K}_k$, as it necessary to compute the Jacobian of the observation model. This dependency in the current state is not present in the linear case. Therefore, a wrong choice for the initial guess of the system state can lead to a divergence in the estimations produced by the Kalman filter which is problematic when no initial knowledge on the system is available.

Finally, in its basic formulation, the Kalman filter is not built with a mechanism to accommodate for faulty measurements, aka outliers. However, criterion on the innovation $\boldsymbol{y}_k - \boldsymbol{C}_k \cdot \hat{\boldsymbol{x}}_{k-1}$, that is, the difference between the expected observation and the actual observation, can be chosen in order to accept or reject a measurement. It consists of choosing a threshold and computing a Mahalanobis distance.

2.3 SET-MEMBERSHIP SLAM

The study of the Kalman filter outlined the reliability problem: when the state equations are non linear or when the noise is not white, Gaussian, the estimates may be inconsistent, meaning that the true solution might not be in a given confidence region, even a high-confidence one. We are interested in set-membership methods based on interval analysis relying on the assumption that the noises are unknown but bounded. Indeed, interval analysis does not require the system to be linearized, and no approximations have to be made: the image of an interval through a nonlinear function is an interval, whereas for example the image of a Gaussian distribution through the same function is not a Gaussian distribution anymore. Given that the models are correct, and the assumptions about the uncertainties are satisfied, the estimates resulting from the set-membership method can be trusted, i.e. this method is reliable . We will see that the assumption that there are outliers in the measurements can be incorporated in the definition of the problem, and the method can be made tolerant to such outliers, making it both robust and reliable. For a short introduction to interval analysis, the reader can refer to A.

Set-membership methods consist in characterizing parameters that are compatible with the observation and their associated bounds. Therefore, contrarily to probabilistic methods, we do not want to find the parameter that minimizes an error criterion, but instead the set

of parameters that are consistent with the model, the observations and their bounds. Also, no likelihood is associated to a given subset of parameters: a subset of parameters is either compatible, incompatible, or no decision can be made about its compatibility with the observations.

It is assumed that the original position $\mathbf{x}(0)$ of the mobile belongs to a known set $\mathbb{X}_0$ and that $\boldsymbol{\alpha}_k \in [\boldsymbol{\alpha}_k], \boldsymbol{\beta}_k \in [\boldsymbol{\beta}_k]$, i.e. the noises are bounded.

2.3.1 *Localization with a Robust State Observer*

RECURSIVE BOUNDED STATE ESTIMATION The feasible set $\mathbb{X}(k)$at any instant k corresponding to the set of all pose and beacon vectors $\mathbf{x}(k)$ that are consistent with the past can be computed recursively [9, 54, 55] from the relation :

$$\mathbb{X}(k+1) = \mathbf{f}_k\left(\mathbb{X}(k) \cap \mathbf{g}_k^{-1}(\mathbb{Y}(k))\right) \tag{2.4}$$

.

Therefore, solving the SLAM problem in a recursive manner amounts to solving the following CSP at each instant k:

$$\mathscr{H}_{RBESE}\begin{cases}\mathbf{x}_k & = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) \\ \mathbf{y}_k & = \mathbf{g}_k(\mathbf{x}_k) \\ \mathbf{x}_k \in & \mathbb{X}(k), \mathbf{y}(k) \in \mathbb{Y}(k)\end{cases}. \tag{2.5}$$

$\mathbf{y}_k$ is guaranteed to be an outlier if we have:

$$\mathbb{X}_k \cap \mathbf{g}^{-1}(\mathbb{Y}_k) = \emptyset. \tag{2.6}$$

Using Equation 2.6, if an outlier is detected it is simply discarded, and the estimator continues to iterate. However, the first problem is that it is possible that an outlier is present without being detected. In such situation the estimator becomes inconsistent and it may take several iterations before the inconsistency is detected. Another problem is that the state vector for the SLAM problem is high dimensional (i.e.: dimension $3 \cdot (N+1)$ where N is the number of landmarks). It then quickly becomes prohibitive to perform bisections as the dimension of the state vector grows. A more robust scheme using a sliding data horizon can be used instead.

CONTRACTOR ON A SLIDING WINDOW In [92], the authors compare the aforementioned recursive bounded estimation with a method based on contractors [16] propagating constraints over a sliding horizon called CPSLAM.

Define:

$$\mathbf{f}_k^i = \mathbf{f}_k \circ \mathbf{f}_{k-1} \circ \cdots \circ \mathbf{f}_{k-i}. \tag{2.7}$$

With an horizon of the l last exteroceptive measurements we can write:

$$\mathbb{X}(k+1) = \mathbf{f}_k(\mathbb{X}_k) \cap \bigcap_{i\in\{0,\dots,l\}} \mathbf{f}_k^i \circ \mathbf{g}_k^{-1}(\mathbb{Y}(k-i)). \tag{2.8}$$

The goal of the CPSLAM approach is then to compute the interval hull $[\mathbb{X}(k+1)]$ of $\mathbb{X}(k+1)$. By memorizing and re-applying several times the same past observation, it is possible to reach a good precision without performing any bisection if desired. In this formulation, outliers are still not accounted for and will eventually become inconsistent as outliers are added.

THE ROBUST STATE OBSERVER In [38], the author describes a more robust alternative to the recursive set estimator using the q-relaxed intersection operator, denoted $\bigcap^{\{q\}}$, that computes the intersection of m sets except q at most. The definition is recalled in Definition A.6.3 in Appendix A. Using relation 2.4, it is straightforward to deduce the following equation:

$$\mathbb{X}(k+1) = \mathbf{f}_k(\mathbb{X}_k) \cap \bigcap^{\{q\}}_{i\in\{0,\dots,l\}} \mathbf{f}_k^i \circ \mathbf{g}_k^{-1}(\mathbb{Y}(k-i)). \tag{2.9}$$

With this formulation, our estimator is guaranteed to survive to q outliers among the last l measurements, independently of the nature or the amplitude of these outliers. Should a greater number of outliers be added to the last l measurements, the estimator will lead to the empty $\emptyset$ set and the inconsistency will be detected. The robust observer is the interval-based observer that we will use, alongside with the Kalman filter in the framework of this thesis.

2.4 COMPARISON

In this section we compare the two approaches on a data-set acquired in the La Ciotat bay, France in February 2014. A ship equipped with a PHINS [76], a RAMSES [81], a DVL, a GPS and a pressure sensor performs a survey of a zone where 4 acoustic beacons lie on the seafloor. The position of these 4 landmarks are precisely known. The trajectory of the ship and the position of the landmarks are depicted in Figure 2.1.

2.4.1 *Methodology*

The perfect knowledge on the vehicle's and beacons positions gives us the ground truth. We will first run a test-case where the robots position is assumed to be perfectly known at the start of the mission, and the beacons positions are known up to a bias in the order of

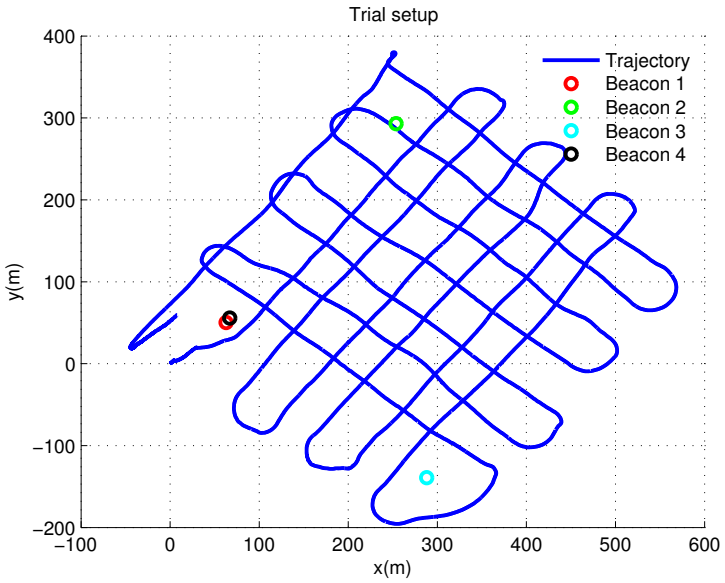


Figure 2.1: Setup of the trial

magnitude of 30 meters in the X/Y plane, and in the order of 1 meter in the Z direction. We will then run a test-case where the robots position is still assumed to be perfectly known, but the landmarks X/Y coordinates are supposedly totally unknown. In this case, the Kalman filter for the landmark is initialized on the initial position of the vehicle with a very large covariance matrix. In both scenarios, the altitude of the beacons is known with a precision of about one meter. We will compare the two approaches for the estimation of the vehicle's and the landmark's positions.

CONSISTENCY An estimation is consistent if the true value of the estimated quantity is contained in the confidence domain provided by the filter. For the Kalman filter, we will say that an estimated position is consistent if the true position is inside its 99% confidence ellipse. For the RSO, the estimation is consistent if it is contained in its sub-paving. We will compare the rate of consistency for both filters during all the mission.

ERROR We define the error of an estimation as the Euclidean distance that separates it from the true value. For the Kalman filter, it will be the distance between its mean and the true position, for the RSO it will be the distance between the center of mass of the sub-paving and the true position.

Vehicle's precision	UKF	RSO
Max. Error - m	1.15	71.48
Final Error - m	0.469	9.57
Consistency - %	100	100

Table 2.1: Precision for the vehicle's localization when a small bias is added to the beacon's positions

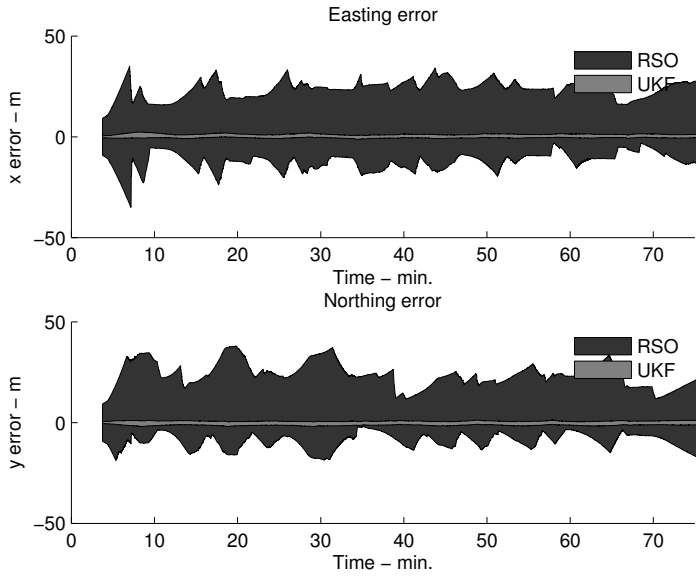


Figure 2.2: Confidence plot for the Easting and the Northing of the mobile when the base is biased

2.4.2 *SLAM with some initial knowledge*

Figure 2.2 displays the confidence domains for the error on the x (Easting) and y (Northing) position of the vehicle in the local frame for the Kalman filter and the RSO. The Kalman filter gives an estimate with a decimetric error, while the RSO gives a decametric error. Both approaches contain the real position during all the mission.

Regarding the estimated positions of the beacons shown in Figure 2.3, the Kalman filter converges in less than 10 minutes to a decimetric precision, while the RSO's precision quickly reaches a decametric precision and doesn't improve after that. Both approaches give estimations that contain the true positions of the beacons during all the mission.

Beacons precision		UKF	RSO
Beacon 1	Final Error - m	0.246	10.99
	Initial bias - m	28.7	
	Consistency - %	100	100
Beacon 2	Final Error - m	0.180	7.45
	Initial bias - m	22.8	
	Consistency - %	100	100
Beacon 3	Final Error - m	0.221	7.56
	Initial bias - m	35.1	
	Consistency - %	100	100
Beacon 4	Final Error - m	0.657	9.05
	Initial bias - m	30.1	
	Consistency - %	40.63	100

Table 2.2: Precision for the beacons localization when a small bias is added to their initial positions

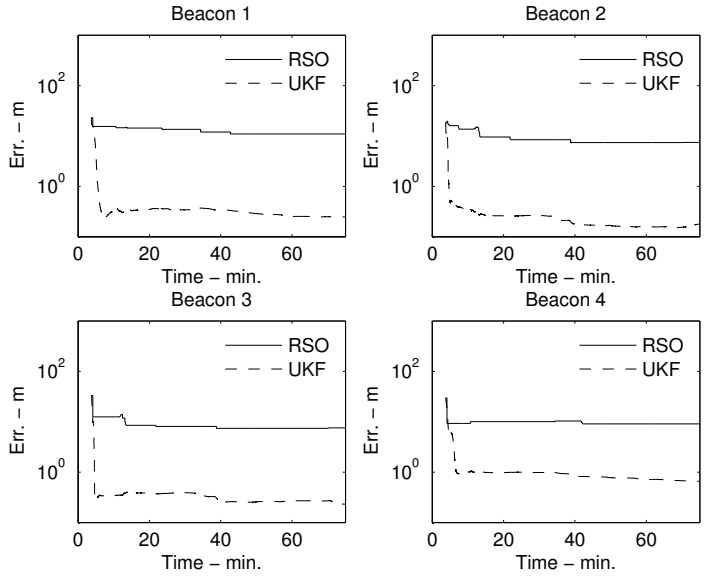


Figure 2.3: Error plot for the estimation of the beacons positions when their initial positions are biased

Vehicle's precision	UKF	RSO
Max. Error - m	4.76	71.48
Final Error - m	3.73	19.0
Consistency - %	83.9	100

Table 2.3: Precision for the vehicle's localization when the beacons positions are unknown

2.4.3 *SLAM without initial knowledge*

Figure 2.4 shows that the confidence on the error of the vehicle's position is higher than for the previous scenario. However, both filters are consistent during the whole mission: their 99% confidence domain for the vehicle's position contain the true position. Regarding the estimation of the beacons positions, the Kalman filter is able to locate the second beacon with a final error of about 4 meters, while it converges to wrong positions for the other beacons, whose 99% confidence domain do not contain the true beacons positions. The RSO on the other hand, is able to locate the beacons with a final error of about 10 meters for all the beacons, and the true beacon's positions are contained in its estimate.

	Beacons precision	UKF	RSO
Beacon 1	Final Error - m	458.2	6.40
	Initial bias - m	319.8	
	Consistency - %	11.6	100
Beacon 2	Final Error - m	4.29	16.85
	Initial bias - m	80.4	
	Consistency - %	68.9	100
Beacon 3	Final Error - m	332.7	17.83
	Initial bias - m	87.1	
	Consistency - %	24.6	100
Beacon 4	Final Error - m	48.7	13.52
	Initial bias - m	387.6	
	Consistency - %	17.5	100

Table 2.4: Precision for the beacons localization when their initial location is unknown

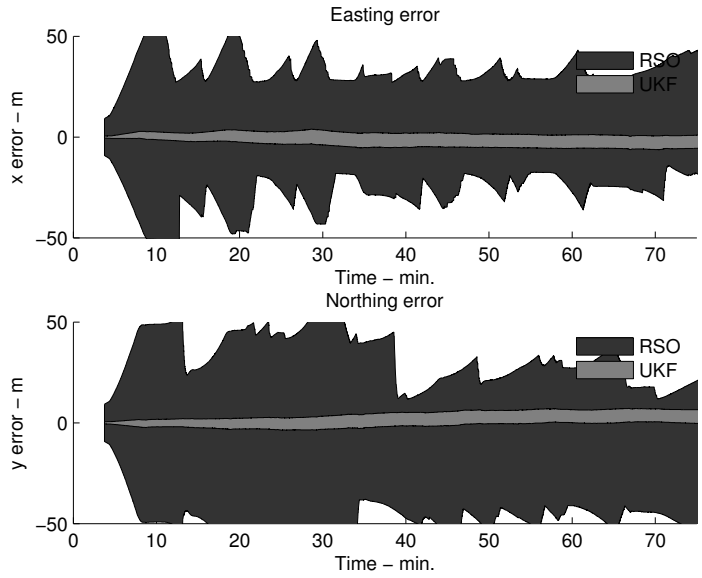


Figure 2.4: Confidence plot for the Easting and the Northing of the mobile when the base is unknown

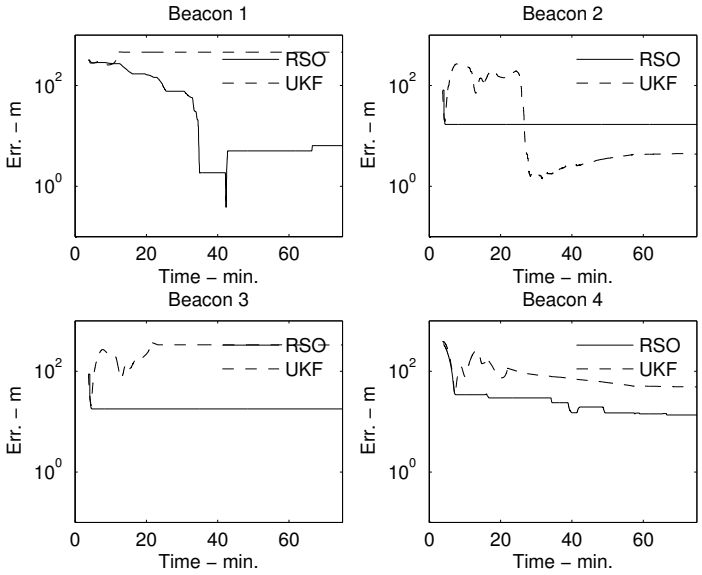


Figure 2.5: Confidence plot for the Easting and the Northing of the mobile when the base is unknown

2.5 DISCUSSION ON THE PESSIMISM OF THE RSO

Section 2.4 shown that the RSO's precision, despite having a 100% consistency, is quite unsatisfying, and one might argue that its high consistency comes from its high level of pessimism. In this section, we describe two factors that have been identified as a unnecessary source of pessimism.

2.5.1 *Wrapping effect*

Definition 1. A transformation T is said to be box-conservative if the image of an axis-aligned box through T is also an axis-aligned box.

Remark 2. The evolution model for the vehicle involves a rotation that maps the speed measured in the local vehicle's frame to the global frame. A rotation is generally not box-conservative, and this rotation adds pessimism known as wrapping-effect [65] [80].

Since the evolution model is linear, interval are not the best representation for sets. For this kind of models, other methods such as zonotopes [61] or ellipsoids are a more attractive representation. To reduce the wrapping induced by the rotation, a combination of linear and interval methods could be used.

2.5.2 *Non-compensation of noises*

Assume the robot moves in a fixed direction. We have $\forall k, \mathbf{R_k} = \mathbf{R}, \mathbf{\Gamma}_{\alpha_k} = \mathbf{\Gamma}_\alpha$. At the beginning the robot knows its exact position, so $\mathbf{\Gamma}_0 = 0$. If the robot does not make any observation, we have the following properties.

Proposition 3. *Without exteroceptive measurements, the precision given by a Kalman filter for the position of the system described previously grows as a square-root of the time.*

Proof. We have

$$\begin{cases} \mathbf{\Gamma}_1 &= \mathbf{\Gamma}_\alpha \\ \mathbf{\Gamma}_2 &= \mathbf{\Gamma}_1 + \mathbf{\Gamma}_\alpha = 2 \cdot \mathbf{\Gamma}_\alpha \\ \vdots & \\ \mathbf{\Gamma}_k &= k \cdot \mathbf{\Gamma}_\alpha \end{cases}. \tag{2.10}$$

Given a fixed time-step dt between instant k and instant $k+1$, the elapsed time is $t = k \cdot dt$, and we have

$$\mathbf{\Gamma}_k = \frac{t}{dt} \cdot \mathbf{\Gamma}_\alpha. \tag{2.11}$$

A confidence domain with confidence η for a random variable $\boldsymbol{x}$ following a normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Gamma}_k$ is an ellipse described by the following inequality

$$\sqrt{(\boldsymbol{x}_k - \boldsymbol{\mu}_k)^T \cdot \boldsymbol{\Gamma}_k^{-1} \cdot (\boldsymbol{x}_k - \boldsymbol{\mu}_k)} \leq \alpha(\eta) \tag{2.12}$$

with $\alpha(\eta)$ the radius of the ellipse, a function that gives the confidence threshold α for a confidence η.

Consequently, introducing Equation 2.11, we have:

$$\sqrt{(\boldsymbol{x}_k - \boldsymbol{\mu}_k)^T \cdot \boldsymbol{\Gamma}_\alpha^{-1} \cdot (\boldsymbol{x}_k - \boldsymbol{\mu}_k)} \leq \alpha(\eta) \cdot \sqrt{\frac{t}{dt}} \tag{2.13}$$

which means that the radius of the ellipse grows as a square root of time. □

Proposition 4. *Without exteroceptive measurements, the precision given by the RSO for the position of the system described above grows linearly with time.*

Proof. The perturbation on the integrated speed is supposed to belong to some interval $[\mathbf{w_{ff}}]$.

From that we have

$$\begin{cases} [\boldsymbol{x}_{m_1}] & = \boldsymbol{x}_{m_0} + \mathbf{R} \cdot \mathbf{u} + [\mathbf{w}_\alpha] \\ [\boldsymbol{x}_{m_2}] & = [\boldsymbol{x}_{m_1}] + \mathbf{R} \cdot \mathbf{u} + [\mathbf{w}_\alpha] = \boldsymbol{x}_{m_0} + 2 \cdot \mathbf{R} \cdot \mathbf{u} + 2 \cdot [\mathbf{w}_\alpha] \\ \vdots & \\ [\boldsymbol{x}_{m_2}] & = \boldsymbol{x}_{m_0} + k \cdot \mathbf{R} \cdot \mathbf{u} + k \cdot [\mathbf{w}_\alpha] \end{cases} \tag{2.14}$$

Given a fixed time-step dt between instant k and instant $k+1$, the elapsed time is $t = k \cdot dt$, and with $w([\boldsymbol{x}])$ the width of the largest interval component of $[\boldsymbol{x}]$, we have

$$w\left(\left[\boldsymbol{x}_{m_{k_i}}\right]\right) = \frac{t}{dt} \cdot w([\boldsymbol{w}_{\alpha_i}])$$

which concludes the proof. □

Propositions 3 and 4 shows that the Kalman filter is much more precise when integrating the proprioceptive measurements than its interval counterpart. By taking into account that the proprioceptive noises are Gaussian, it is able to provide a precision that grows as a square root of the time, whereas the RSO provides an estimate whose precision grows linearly with time.

2.6 CONCLUSION

In this chapter, we compared the use of a Kalman filter against an interval filter for the SLAM problem of an underwater vehicle on a real data-set. As expected, the Kalman filter gives estimates that are of high precision when we have a reasonable prior knowledge on the beacons positions, and might converge towards wrong solutions when no such knowledge is available. The RSO, on the other hand, is consistent during all the mission with or without good prior knowledge, but its precision is quite poor. To improve the precision of the Kalman filter, we could use an approach similar to [11], where the initial state is first estimated with an interval method, and then a Kalman filter is initialized from this estimate. An other approach would be to overcome the sources of pessimism described in Section 2.5. Since the evolution model is linear, the wrapping effect could be reduced by combining an interval approach with a linear approach such as ellipsoidal methods [9, 19, 77] or zonotopes [61]. An other way of drastically reducing the pessimism of the RSO would be to make it able to integrate the uncertainties as a square-root of time, which would make the propagation-retro-propagation described in Equation 2.7 much more precise. This approach will be described in the next chapter.

Part II

CONTRIBUTIONS

In Part i we reviewed, compared and analyzed the performances of two SLAM algorithms on a real data set. We observed that set-based methods tend to be less precise than their probabilistic counterparts, with the benefit of a greater robustness. Part ii will be dedicated to the study of the sources of pessimism that makes interval approaches less precise, and tools to overcome this problem will be presented. Thus, Chapter 3 will focus on diminishing the pessimism when propagating information between states $\mathbf{x}_k, \mathbf{x}_l$ at two instants k, l in the trajectory. Chapter 4 will focus on the static aspect of SLAM, more precisely on the fusion of exteroceptive measurements, and how to make it more precise. Finally, the results of this thesis will be discussed and Chapter 5 will conclude this manuscript.

3

THE KALMAN CONTRACTOR

3.1 INTRODUCTION

In Chapter 2 we have seen how, for a simple open-loop motion model, a Kalman filter's confidence domain grows as a square root of time whereas an interval filter's grows linearly with time. We will put aside the observation and update equations of the Kalman filter to focus only on the prediction equation. This focus is motivated by the model of our underwater robot. In Chapter 1, Equation 1.6, we have seen that the motion model of the robot is linear, and that the uncertainties applied on the motion can reasonably be modeled as Gaussian. This is the perfect use case for a Kalman filter, as has been discussed in Section 2.2.1. On the other hand, the observation model described in Equation 1.9 is non-linear and non-invertible, which means that the full robot's state cannot be recovered from a single observation. Additionally, the sources of noises on the observations are highly uncontrolled, and we have seen that modeling it as an additive, white Gaussian noise is a coarse simplification. Therefore, the observations are not well suited for being processed by a Kalman filter. The goal of this chapter is therefore to study how a Kalman filter achieves a good processing of the motion uncertainties, and see how we can take advantage of its properties in a set-membership framework [70]. Note that combining the Kalman filter with set membership uncertainties has already been considered by Jauberthie et al. [34] [96] [90].

By studying a simple example, we will highlight how we can give the ability to integrate the uncertainties as a square root of time to an interval filter, to finally result in an elegant and compact tool to achieve that goal.

3.2 CASE STUDY: THE RANDOM WALK

Let's consider the simple motion model of the random walk:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{w}_k \tag{3.1}$$

.

Equation 3.1 could for example describe the motion of a submarine that stays fixed above the seafloor, where its imperfect DVL reports a velocity which is centered and normally distributed with a covariance matrix $\boldsymbol{\Gamma}_w$.

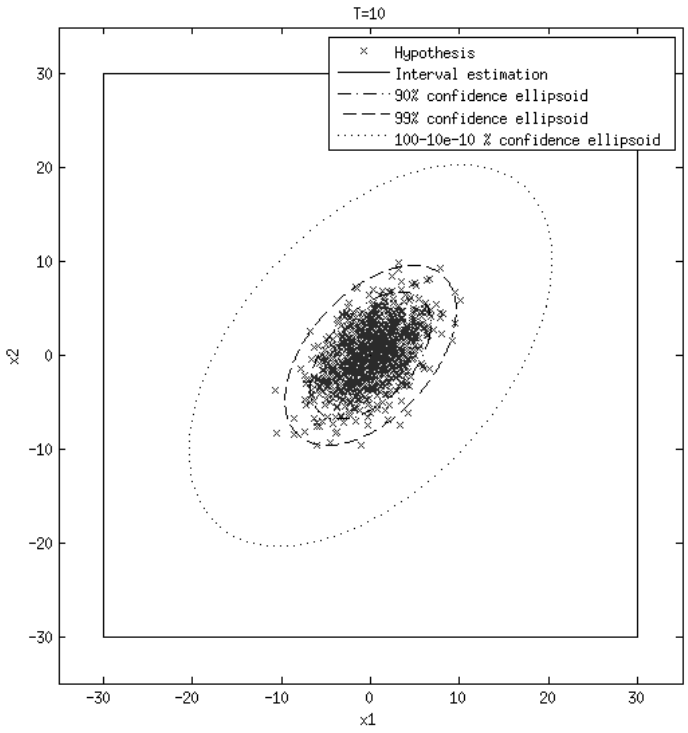


Figure 3.1: Monte-Carlo propagation for Equ. 3.1 after 10 seconds: the cross are the hypothesis, the ellipse are the confidence domain for various confidence thresholds, the inner square is the confidence region for an interval propagation.

Figure 3.1 shows an integration of Equation 3.1 for 10 time steps with a Monte-Carlo method and an interval integration. It is clear that the interval method is extremely conservative, and provides a confidence region which contains $\boldsymbol{x}_{10}$ with a confidence greater than $1 - 10^{-10}$.

Figure 3.2 shows how the confidence region grows with time both with a Kalman filter and an interval filter integration of Equation 3.1, which confirms what we observed in the previous chapter.

3.3 THE KALMAN CONTRACTOR

In this section, we will take a closer look on the prediction equations of statistical relationship that links the states $\boldsymbol{x}_k, \boldsymbol{x}_l$ at two instants k, l of a system with linear motion model subject to additive, white Gaussian noise. An IMU generally provides inertial data at a high rate. As has been explained in Section 2.3.1, solving the SLAM problem with a robust state observer requires to propagate information between the current instant k and previous instants where exteroceptive measurements were gathered. This requires us to compute the

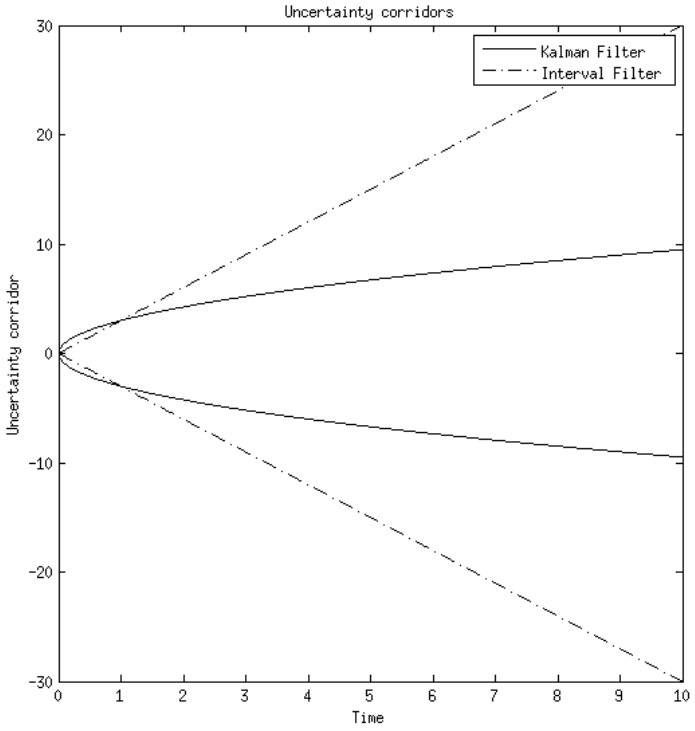


Figure 3.2: 1D projection of the confidence region as a function of time

composed function $f_k^i = f_k \circ f_{k-1} \circ \cdots \circ f_{k-i}$ such that $x_k = f_k^i(x_i)$. For the general case, there is no known method to find f_k^i by other means than actually computing the composition. As there is at least as many composition to perform as there were proprioceptive measurements between instant k and instant l, and as an IMU generally provide attitude data at a high rate, computing f_k^i will prove to be an expensive operation, forcing us to use very short sliding windows to perform real-time estimations. For example, for an IMU that produces orientations at a rate of 100Hz, computing $x_k = f_k^i(x_i)$ with x_k and x_l distant by 10 seconds will require us to perform about 1000 operations. Now, interval constraint propagation methods are based on fix-point methods, which means that information must be propagated and retro-propagated many times between x_k and x_l. Since there is also no way to easily compute f_i^k, the inverse of f_k^i for the general case, we can see how quickly fix-point methods will become computationally too expensive for real-time use. For our specific use-case, where the attitude of the system can be seen as inputs and is not part of the state vector, and where the evolution equation is linear, we will show that it is possible to "pre-integrate" the motion equations of the vehicle. This means that it is possible to efficiently compute the transformation f_k^i and its inverse, allowing us to compute x_k from x_l

and reverse in a single operation. To do so, we will unroll the Kalman filter prediction equations and rewrite them in an "integral" form.

3.3.1 *Principle*

Let's iterate the Kalman filter prediction equation for the following general linear system:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A} \cdot \boldsymbol{x}_k + \boldsymbol{u}_k + \boldsymbol{w}_k \tag{3.2}$$

.

For $k \in [0, 3]$, we have the predicted mean $\hat{\boldsymbol{x}}_k$:

$$\begin{cases} \hat{\boldsymbol{x}}_1 &= \boldsymbol{A}_0 \cdot \hat{\boldsymbol{x}}_0 + \boldsymbol{u}_0 \\ \hat{\boldsymbol{x}}_2 &= \boldsymbol{A}_1 \cdot \boldsymbol{A}_0 \cdot \hat{\boldsymbol{x}}_0 + \boldsymbol{A}_1 \cdot \boldsymbol{u}_0 + \boldsymbol{u}_1 \\ \hat{\boldsymbol{x}}_3 &= \boldsymbol{A}_2 \cdot \boldsymbol{A}_1 \cdot \boldsymbol{A}_0 \cdot \hat{\boldsymbol{x}}_0 + \boldsymbol{A}_2 \cdot \boldsymbol{A}_1 \cdot \boldsymbol{u}_0 + \boldsymbol{A}_1 \cdot \boldsymbol{u}_1 + \boldsymbol{u}_2 \end{cases} \tag{3.3}$$

and the predicted covariance matrix $\boldsymbol{\Gamma}_k$:

$$\begin{cases} \boldsymbol{\Gamma}_1 &= \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_0 \cdot \boldsymbol{A}_0^T + \boldsymbol{\Gamma}_{\boldsymbol{w}} \\ \boldsymbol{\Gamma}_2 &= \boldsymbol{A}_1 \cdot \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_0 \left(\boldsymbol{A}_0 \cdot \boldsymbol{A}_1\right)^T + \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}} \cdot \boldsymbol{A}_0^T + \boldsymbol{\Gamma}_{\boldsymbol{w}} \\ \boldsymbol{\Gamma}_3 &= \boldsymbol{A_2} \cdot \boldsymbol{A}_1 \cdot \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_0 \left(\boldsymbol{A}_0 \cdot \boldsymbol{A}_1 \cdot \boldsymbol{A}_2\right)^T + \boldsymbol{A}_1 \cdot \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}} \left(\boldsymbol{A}_0 \cdot \boldsymbol{A}_1\right)^T \\ &\quad + \boldsymbol{A}_0 \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}} \left(\boldsymbol{A}_0\right)^T + \boldsymbol{\Gamma}_{\boldsymbol{w}} \end{cases} \tag{3.4}$$

.

Definition 5. Now define the transition matrix $\mathbf{P}_k^i$ as:

$$\begin{cases} \mathbf{P}_k^i = & \mathbf{A}_{k-1} \cdot \mathbf{A}_{k-2} \ldots \mathbf{A}_i, \text{ if } k > i \\ \mathbf{P}_k = & \mathbf{P}_k^0 = \mathbf{I} \\ \mathbf{P}_k^i = & \mathbf{P}_i^k, \text{ if } k < i \end{cases} \tag{3.5}$$

The transition matrix satisfies the Chasle equation:

$$\mathbf{P}_k^i = \mathbf{P}_k^l \cdot \mathbf{P}_l^i \tag{3.6}$$

and

$$\mathbf{P}_k^i = \mathbf{P}_k \cdot \mathbf{P}_i^{-1} \tag{3.7}$$

which means that matrices $\mathbf{P}_k^i$ can all be obtained from the $\mathbf{P}_k$.

Remark 6. The state at any time k can be expressed only as a function of $\mathbf{x}_0$:

$$\mathbf{x}_k = \mathbf{P}_k \mathbf{x}_0 + \sum_{i=0}^{k-1} \mathbf{P}_k^i \cdot (\mathbf{u}_i + \boldsymbol{\alpha}_i) \tag{3.8}$$

Proof. We have:

$$\begin{array}{lccccc} \mathbf{x}_1 = & \mathbf{A}_0\mathbf{x}_0 + \mathbf{u}_0 + \boldsymbol{\alpha}_0 & & & & \\ \mathbf{x}_2 = & \mathbf{A}_1\mathbf{A}_0\mathbf{x}_0 + & \mathbf{A}_1 \cdot (\mathbf{u}_0 + \boldsymbol{\alpha}_0) + & \mathbf{u}_1 + \boldsymbol{\alpha}_1 & & \\ \mathbf{x}_3 = & \mathbf{A}_2\mathbf{A}_1\mathbf{A}_0\mathbf{x}_0 + & \mathbf{A}_2\mathbf{A}_1 \cdot (\mathbf{u}_0 + \boldsymbol{\alpha}_0) + & \mathbf{A}_2\mathbf{u}_1 + \mathbf{A}_2\boldsymbol{\alpha}_1 + & \mathbf{u}_2 + \boldsymbol{\alpha}_2 & \\ \mathbf{x}_4 = & \mathbf{P}_3\mathbf{x}_0 + & \mathbf{P}_3 \cdot (\mathbf{x}_0 + \boldsymbol{\alpha}_0) + & \mathbf{P}_3^1 (\mathbf{u}_1 + \boldsymbol{\alpha}_1) + & \mathbf{P}_3^2 \cdot (\mathbf{u}_2 + \boldsymbol{\alpha}_2) + & \mathbf{u}_3 + \boldsymbol{\alpha}_3 \\ \cdots = & \cdots & & & & \\ \mathbf{x}_k = & \mathbf{P}_k\mathbf{x}_0 + & \sum_{i=0}^{k-1} \mathbf{P}_k^i \cdot (\mathbf{u}_i + \boldsymbol{\alpha}_i) & & & \end{array} \tag{3.9}$$

□

Then from Equation 3.9, $\hat{\boldsymbol{x}}_k, \boldsymbol{\Gamma}_k$ can be obtained for any instant k:

$$\begin{cases} \hat{\boldsymbol{x}}_k &= \boldsymbol{P}_k^0 \cdot \hat{\boldsymbol{x}}_0 + \sum_{i=0}^{k-1} \boldsymbol{P}_k^{i+1} \boldsymbol{u}_i \\ \boldsymbol{\Gamma}_k &= \boldsymbol{P}_k^0 \cdot \boldsymbol{\Gamma}_0 \cdot \left(\boldsymbol{P}_k^0\right)^T + \sum_{i=1}^{k} \boldsymbol{P}_k^i \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}_i} \cdot \left(\boldsymbol{P}_k^i\right)^T \end{cases} \tag{3.10}$$

.

Equation 3.10 can also be written as:

$$\begin{cases} \hat{\boldsymbol{x}}_k &= \boldsymbol{P}_k^0 \cdot \left(\hat{\boldsymbol{x}}_0 + \sum_{i=0}^{k-1} \left(\boldsymbol{P}_{i+1}^0\right)^{-1} \boldsymbol{u}_i\right) \\ \boldsymbol{\Gamma}_k &= \boldsymbol{P}_k^0 \cdot \left(\boldsymbol{\Gamma}_0 + \sum_{i=1}^{k} \left(\boldsymbol{P}_i^0\right)^{-1} \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}_i} \cdot \left(\boldsymbol{P}_i^0\right)^{-1^T}\right) \cdot \left(\boldsymbol{P}_k^0\right)^T \end{cases} \tag{3.11}$$

and by introducing the matrices:

$$\begin{cases} \boldsymbol{Q}_k &= \sum_{i=0}^{k-1} \left(\boldsymbol{P}_{i+1}^0\right)^{-1} \boldsymbol{u}_i \\ \boldsymbol{R}_k &= \sum_{i=1}^{k} \left(\boldsymbol{P}_i^0\right)^{-1} \cdot \boldsymbol{\Gamma}_{\boldsymbol{w}_i} \cdot \left(\boldsymbol{P}_i^0\right)^{-1^T} \end{cases} \tag{3.12}$$

we get a very compact writing for Equation 3.10:

$$\begin{cases} \hat{\boldsymbol{x}}_k &= \boldsymbol{P}_k^0 \cdot (\hat{\boldsymbol{x}}_0 + \boldsymbol{Q}_k) \\ \boldsymbol{\Gamma}_k &= \boldsymbol{P}_k^0 \cdot (\boldsymbol{\Gamma}_0 + \boldsymbol{R}_k) \cdot \left(\boldsymbol{P}_k^0\right)^T \end{cases} . \tag{3.13}$$

Now, if we look at the states $\boldsymbol{x}_k, \boldsymbol{x}_l$ taken at two different instants k, l:

$$\begin{cases} \hat{\boldsymbol{x}}_k &= \boldsymbol{P}_k^0 \cdot \hat{\boldsymbol{x}}_0 + \sum_{i=0}^{k-1} \boldsymbol{P}_k^{i+1} (\boldsymbol{u}_i + \boldsymbol{w}_i) = \boldsymbol{P}_k^0 \left(\hat{\boldsymbol{x}}_0 + \sum_{i=0}^{k-1} \left(\boldsymbol{P}_{i+1}^0\right)^{-1} (\boldsymbol{u}_i + \boldsymbol{w}_i)\right) \\ \hat{\boldsymbol{x}}_l &= \boldsymbol{P}_l^0 \cdot \hat{\boldsymbol{x}}_0 + \sum_{i=0}^{l-1} \boldsymbol{P}_l^{i+1} (\boldsymbol{u}_i + \boldsymbol{w}_i) = \boldsymbol{P}_l^0 \left(\hat{\boldsymbol{x}}_0 + \sum_{i=0}^{l-1} \left(\boldsymbol{P}_{i+1}^0\right)^{-1} (\boldsymbol{u}_i + \boldsymbol{w}_i)\right) \end{cases} \tag{3.14}$$

we can see that $\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{x}}_l$ are linear combinations of the same, independent, normally distributed random variables which implies that the random variable $\hat{\boldsymbol{x}}_{kl}$ resulting of the concatenation of $\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{x}}_l$ is also normally distributed:

$$\begin{pmatrix} \hat{\boldsymbol{x}}_k \\ \hat{\boldsymbol{x}}_l \end{pmatrix} \sim N\left(\hat{\boldsymbol{x}}_{kl}, \boldsymbol{\Gamma}_{kl}\right) \tag{3.15}$$

.

We can fully characterize its distribution by computing its first and second order moments:

$$E\left(\boldsymbol{x}_{kl}\right) = \hat{\boldsymbol{x}}_{kl} = \begin{pmatrix} E\left(\boldsymbol{x}_k\right) \\ E\left(\boldsymbol{x}_l\right) \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{x}}_k \\ \hat{\boldsymbol{x}}_l \end{pmatrix} \tag{3.16}$$

$$E\left(\boldsymbol{x}_{kl} \cdot \boldsymbol{x}_{kl}^T\right) = \boldsymbol{\Gamma}_{kl} = \begin{pmatrix} E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_k^T\right) & E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_l^T\right) \\ E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_l^T\right) & E\left(\boldsymbol{x}_l \cdot \boldsymbol{x}_l^T\right) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Gamma}_k & E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_l^T\right) \\ E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_l^T\right) & \boldsymbol{\Gamma}_l \end{pmatrix} \tag{3.17}$$

.

Now, $\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{x}}_l, \boldsymbol{\Gamma}_k, \boldsymbol{\Gamma}_l$ can naturally be obtained from Equation 3.13, while a quick calculation for the of-diagonal block-matrices of $\boldsymbol{\Gamma}_{kl}$ gives:

$$E\left(\boldsymbol{x}_k \cdot \boldsymbol{x}_l^T\right) = \boldsymbol{P}_k^0 \left(\boldsymbol{\Gamma}_0 + \boldsymbol{x}_0 \cdot \boldsymbol{Q}_l^T + \boldsymbol{Q}_k \cdot \boldsymbol{x}_0^T + \boldsymbol{R}_l\right) \boldsymbol{P}_l^{0^T} \tag{3.18}$$

.

From the now fully characterized Gaussian distribution of Equation 3.15 we can compute a confidence region.

Definition 7. The inter-temporal confidence ellipsoid $\boldsymbol{\Sigma}_{kl}$ containing the states $\boldsymbol{x}_k, \boldsymbol{x}_l$ at instants k, l with a probability η is the set:

$$\boldsymbol{\Sigma}_{kl} = \left\{ \boldsymbol{x} \middle| \left(\boldsymbol{x} - \hat{\boldsymbol{x}}_{kl}\right) \cdot \boldsymbol{\Gamma}_{kl}^{-1} \cdot \left(\boldsymbol{x} - \hat{\boldsymbol{x}}_{kl}\right)^T - \alpha^2\left(\eta\right) \leq 0 \right\} \tag{3.19}$$

with α a confidence threshold.

A contractor can be defined for the set described in Definition 7:

Definition 8. The Kalman contractor $\boldsymbol{C}_{k,l} : \mathbb{IR}^{2n} \to \mathbb{IR}^{2n}$ contracts $[\boldsymbol{x}] = [\boldsymbol{x}_k] \times [\boldsymbol{x}_l]$ with respect to the constraint $\boldsymbol{x} \subseteq \boldsymbol{\Sigma}_{kl}$.

The Kalman contractor enables us to propagate information between two states at different instants k, l with a single constraint, without having to re-integrate the motion equations between k, l. Figure 3.3 illustrates the contraction of a box $[\boldsymbol{x}]$ made of the concatenation of two boxes $[\boldsymbol{x}_k], [\boldsymbol{x}_l]$ with respect to their inter-temporal confidence region $\boldsymbol{\Sigma}_{kl}$.

Selecting the right confidence threshold is important: there is a probability of $1 - \eta$ that the Kalman contractor will eliminate the correct value. Also, care should be taken when composing several Kalman contractors. Say we apply successively the Kalman contractor for different pairs: $\boldsymbol{C}_{k,l}, \boldsymbol{C}_{k,m}, \boldsymbol{C}_{k,n}$. In that case, $[\boldsymbol{x}_k]$ would contain $\boldsymbol{x}_k$ with a probability $\boldsymbol{\eta}^3$. If we selected for example $\boldsymbol{\eta} = 0.9$, that means that $[\boldsymbol{x}_k]$ now contains $\boldsymbol{x}_k$ with a probability 0.729, a 17% decrease. An obvious solution is to use the confidence threshold $\eta' = \eta^{1/n}$, with n the number of different Kalman contractors acting on $\boldsymbol{x}_k$. We could also construct a "big" inter-temporal ellipse $\boldsymbol{\Sigma}_{klm\ldots}$ in the trajectory

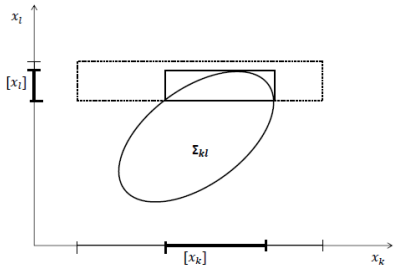


Figure 3.3: Illustration of an inter-temporal confidence ellipsoid $\mathbf{\Sigma}_{kl}$ and the contraction of a box $[\boldsymbol{x}] = [\boldsymbol{x}_k] \times [\boldsymbol{x}_l]$ with respect to $\mathbf{\Sigma}_{kl}$.

space with an appropriate radius $\alpha(\eta)$ taking into account the dimension of the problem, and then compute the inter-temporal ellipsoids as as many projections of the bigger ellipsoid onto each desired pair of instants.

The approach taken here suffers from the fact that it is not feasible to solve the global localization problem, as the mean $\hat{\mathbf{x}}_0$ and covariance matrix $\mathbf{\Gamma}_0$ must be known to derivate the mean and covariance matrix for any other state $\mathbf{x}_k$ at instant k. To model the total lack of knowledge about the initial state when there is no way to compute a reasonable initial guess, the covariance matrix is generally chosen very large so that its associated confidence ellipse covers a large area over search-space of interest. For nonlinear functions, linearizing and sigma-point methods tend to fail in such situation because they will linearize the model or propagate sigma-points around a point that is very far away from the correct one.

In the next section we will derive a relative formulation of the Kalman contractor presented here to get rid of the initial condition knowledge problem.

Moreover, when $\mathbf{\Sigma}_{kl}$ is not diagonal, according to [40] a contractor based on forward-backward propagation will not be optimal, as there will be several variable repetition in the expression of Equation 3.19. The next section will detail how the multi-occurrence of variables can be solved in this context, and Appendix B will develop the method in details.

Remark 9. When the initial state is not known and the estimate is initialized with a very large confidence matrix, one might think that the Kalman contractor from Definition 8 will not efficiently propagate information between two states $\boldsymbol{x}_k, \boldsymbol{x}_l$. Indeed, since $\mathbf{\Gamma}_0$ is large, the diagonal terms of $\mathbf{\Gamma}_{kl}$ will also be large, and the ellipsoid will be huge. However, the information gain will be carried by the correlation between $\boldsymbol{x}_k, \boldsymbol{x}_l$, which is carried by the off-diagonal elements of $\mathbf{\Gamma}_{kl}$ given by 3.18. For a random walk, two states $\boldsymbol{x}_k, \boldsymbol{x}_l$ with k, l close will be highly correlated, resulting in a tight ellipsoid $\mathbf{\Sigma}_{kl}$. The further $\boldsymbol{x}_l$ will be from $\boldsymbol{x}_k$, the less correlated the two states will be, and the more

$\mathbf{\Sigma}_{kl}$ will be shaped like a ball. The ability to propagate information precisely will thus depend on the tightness of $\mathbf{\Sigma}_{kl}$.

3.3.2 *The Kalman contractor without initial conditions*

We are now giving an alternative, relative formulation for the Kalman contractor which does not require the knowledge of the initial condition $\hat{\mathbf{x}}_0$. Such a formulation allows for example to solve the problem of global localization, where no prior assumption is available on the initial localization of our vehicle. Modeling such an absence of information in a Gaussian framework is generally solved by setting $\mathbf{\Gamma}_0 = \lambda \cdot \boldsymbol{I}$, with $\boldsymbol{I}$ the identity matrix and λ some huge value. The following formulation enables us to overcome this limitation by focusing on a relative representation of the state equations.

Proposition 10. *The signal $\mathbf{c}_k^l = \mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l$ is Markovian, and a Kalman filter can be applied.*

Proof. We have:

$$\begin{cases} \mathbf{c}_{k+1}^l & = \mathbf{x}_{k+1} - \mathbf{P}_{k+1}^l \cdot \mathbf{x}_l \\ & = \mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{u}_k + \boldsymbol{\alpha}_k - \mathbf{A}_k \cdot \mathbf{P}_k^l \cdot \mathbf{x}_0 \\ & = \mathbf{A}_k \cdot \left(\mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_0\right) + \mathbf{u}_k + \boldsymbol{\alpha}_k \\ & = \mathbf{A}_k \cdot \mathbf{c}_k^l + \mathbf{u}_k + \boldsymbol{\alpha}_k \end{cases} \tag{3.20}$$

$\mathbf{c}_k^l$ is indeed Markovian, and a Kalman filter may be applied. □

Since $\mathbf{c}_l^l = \mathbf{0}$ and $\mathbf{\Gamma}_c^l(l) = \mathbf{0}$, a Kalman filter gives the following prediction equations:

$$\begin{cases} \hat{\mathbf{c}}_k^l & = \sum_{i=l}^{k-1} \mathbf{P}_k^i \cdot \mathbf{u}_i \\ \mathbf{\Gamma}_{\mathbf{c}}^l(k) & = \sum_{i=l}^{k-1} \mathbf{P}_k^i \cdot \mathbf{\Gamma}_{\boldsymbol{\alpha}} \cdot \left(\mathbf{P}_k^i\right)^T \end{cases}. \tag{3.21}$$

Which means a confidence ellipsoid for $\mathbf{c}_k^l$ is:

$$\left(\mathbf{c}_k^l - \hat{\mathbf{c}}_k^l\right)^T \cdot \left(\mathbf{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{c}_k^l - \hat{\mathbf{c}}_k^l\right) \leq a \tag{3.22}$$

Theorem 11. *A confidence region containing the inter-temporal state vector $\mathbf{x}_k^l = (\mathbf{x}_k, \mathbf{x}_l)$ with a probability η is the ellipsoid $\mathbb{E}_k^l(\eta)$ defined as:*

$$\mathbb{E}_k^l(\eta) = \left\{ (\mathbf{x}_k, \mathbf{x}_l) \mid \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix}^T \cdot \begin{pmatrix} \left(\mathbf{\Gamma}_c^l\right)^{-1}(k) & \left(\mathbf{\Gamma}_c^l\right)^{-1}(k) \cdot \left(-\mathbf{P}_k^l\right) \\ -\mathbf{P}_k^l \cdot \left(\mathbf{\Gamma}_c^l\right)^{-1}(k) & \left(\mathbf{P}_k^l\right)^T \cdot \left(\mathbf{\Gamma}_c^l\right)^{-1}(k) \cdot \mathbf{P}_k^l \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix} \leq \alpha^2(\eta) \right\} \tag{3.23}$$

with $\alpha(\eta)$ a function that gives the radius of the confidence ellipse with probability η for a normally distributed random variable of dimension $d = dim(\mathbf{x}_k) + dim(\mathbf{x}_l)$.

Proof. From Equation 3.22, and since $\mathbf{c}_k^l = \mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l$, a confidence region containing $\mathbf{x}_k^l$ with a probability η is:

$$
\begin{aligned}
& \left(\mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l - \hat{\mathbf{c}}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l - \hat{\mathbf{c}}_k^l\right) \leq \alpha^2(\eta) \\
\Leftrightarrow \quad & \left(\mathbf{x}_k - \hat{\mathbf{c}}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l - \hat{\mathbf{c}}_k^l\right) - \left(\mathbf{P}_k^l \cdot \mathbf{x}_l\right)^T \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{x}_k - \mathbf{P}_k^l \cdot \mathbf{x}_l - \hat{\mathbf{c}}_k^l\right) \leq \alpha^2(\eta) \\
\Leftrightarrow \quad & \left(\mathbf{x}_k - \hat{\mathbf{c}}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{x}_k - \hat{\mathbf{c}}_k^l\right) - \left(\mathbf{x}_k - \hat{\mathbf{c}}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{P}_k^l \cdot \mathbf{x}_l\right) \\
& + \mathbf{x}_l^T \cdot \mathbf{P}_k^{lT} \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{P}_k^l \cdot \mathbf{x}_l\right) - \mathbf{x}_l^T \cdot \mathbf{P}_k^{lT} \cdot \left(\boldsymbol{\Gamma}_{\mathbf{c}}^l\right)^{-1}(k) \cdot \left(\mathbf{x}_k - \hat{\mathbf{c}}_k^l\right) \leq \alpha^2(\eta) \\
\Leftrightarrow \quad & \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix}^T \cdot \begin{pmatrix} \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \left(-\mathbf{P}_k^l\right) \\ -\mathbf{P}_k^l \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\mathbf{P}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \mathbf{P}_k^l \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix} \leq \alpha^2(\eta)
\end{aligned}
$$

□

Definition 12. A contractor that contracts the box $[\mathbf{x}_l] \times [\mathbf{x}_k]$ with respect to $\mathbb{E}_k^l(\eta)$ is a Kalman contractor.

Theorem 13. *The problem of contracting optimally the box $[\mathbf{x}_l] \times [\mathbf{x}_k]$ with respect to Equation 3.23 amounts to solving $2 \cdot d$ semi-definite programming problems, and an optimal contractor for this constraint can be built using the LMI contractor.*

Proof. Schur complement theorem states that for a set of matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ respectively of dimensions $p \times p, p \times q, q \times p, q \times q$, we have the following equivalence:

$$
\begin{cases} \mathbf{A} & \succ \mathbf{0} \\ \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \succeq \mathbf{0} \end{cases} \Leftrightarrow \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \succeq \mathbf{0}, \tag{3.24}
$$

therefore since $\alpha^2(\eta) \geq 0$, Equation 3.23 can be rewritten as:

$$
\begin{pmatrix} \alpha^2(\eta) & \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix}^T \\ \begin{pmatrix} \mathbf{x}_k - \hat{\mathbf{c}}_k^l \\ \mathbf{x}_l \end{pmatrix} & \begin{pmatrix} \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \left(-\mathbf{P}_k^l\right) \\ -\mathbf{P}_k^l \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\mathbf{P}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \mathbf{P}_k^l \end{pmatrix} \end{pmatrix} \succeq \mathbf{0}, \tag{3.25}
$$

which is an LMI [13]. Therefore, finding the smallest box satisfying Equation 3.25 amounts to minimizing then maximizing each component x_i of $[\mathbf{x}_l] \times [\mathbf{x}_k]$, which is done by the LMI contractor. □

The reader is referred to Appendix B for a more detailed on contracting optimally a box under LMI constraint.

Remark 14. When the evolution matrix $\mathbf{A} = \mathbf{I}$, the shape matrix

$$\begin{pmatrix} \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \left(-\mathbf{P}_k^l\right) \\ -\mathbf{P}_k^l \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) & \left(\mathbf{P}_k^l\right)^T \cdot \left(\boldsymbol{\Gamma}_c^l\right)^{-1}(k) \cdot \mathbf{P}_k^l \end{pmatrix} \tag{3.26}$$

is singular and Schur complement theorem doesn't apply as is, and Moore-Penrose pseudo-inverse should be used. In this situation, it is better to use Equation 3.22 which is also an LMI and yields to an optimal contractor since the transformation by the identity matrix is box-conservative.

In this work, we refer to "the" Kalman contractor as a Kalman contractor based on the LMI contractor, but any (preferably optimal) contractor satisfying Definition 12 or Definition 8 is a Kalman contractor.

3.4 ALGORITHM

We will now describe the steps necessary to implement the Kalman contractor for state estimation using the Definition 12 derived from Theorem 11. However, it is straightforward to deduce the steps used to apply the version of the Kalman contractor defined at the beginning of Section 3.3.

3.4.1 *Initialization*

At instant $k = 0$ we initialize $\mathbf{P}_0^0, \hat{\mathbf{c}}_0^0, \boldsymbol{\Gamma}_\mathbf{c}^0(0)$ as follows:

$$\begin{aligned} \mathbf{P}_0^0 &= \mathbf{I} \\ \hat{\mathbf{c}}_0^0 &= \mathbf{0} \\ \boldsymbol{\Gamma}_\mathbf{c}^0(0) &= \mathbf{0} \end{aligned} \quad . \tag{3.27}$$

3.4.2 *Prediction*

For every instant k, compute $\mathbf{P}_{k+1}^0, \hat{\mathbf{c}}_{k+1}^0, \boldsymbol{\Gamma}_\mathbf{c}^0(k+1)$ such that:

$$\begin{aligned} \mathbf{P}_{k+1}^0 &= \mathbf{A}_k \cdot \mathbf{P}_k^0 \\ \hat{\mathbf{c}}_{k+1}^0 &= \mathbf{A}_k \cdot \left(\mathbf{u}_k + \hat{\mathbf{c}}_k^0\right) \\ \boldsymbol{\Gamma}_\mathbf{c}^0(k+1) &= \mathbf{A}_k \cdot \left(\boldsymbol{\Gamma}_\boldsymbol{\alpha} + \boldsymbol{\Gamma}_\mathbf{c}^0(k)\right) \cdot \mathbf{A}_k^T \end{aligned} \quad . \tag{3.28}$$

For instants that will be later used for constraint propagation, these three quantities should be memorized. In the case of mobile robot navigation, we will want to store these quantities at instant when an exteroceptive measurement is received.

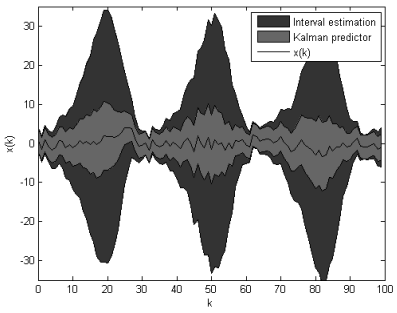


Figure 3.4: The raw signal $x(k)$, its interval tube, the tube given by a Kalman predictor.

3.4.3 *Contraction*

When the current state should be updated according to the measurement history, compute the quantities:

$$\begin{aligned}\hat{\mathbf{c}}_k^l &= \hat{\mathbf{c}}_k^0 - \hat{\mathbf{c}}_l^0 \\ \boldsymbol{\Gamma}_{\mathbf{c}}^l(k) &= \boldsymbol{\Gamma}_{\mathbf{c}}^0(k) - \boldsymbol{\Gamma}_{\mathbf{c}}^0(l) \\ \mathbf{P}_k^l &= \mathbf{P}_k^0 \cdot \left(\mathbf{P}_l^0\right)^{-1}\end{aligned}$$

and apply the Kalman contractor for $\mathbb{E}_k^l(\eta)$ to the box $[\mathbf{x}_l] \times [\mathbf{x}_k]$.

3.5 ILLUSTRATION

To illustrate the use of the Kalman contractor, we will simulate a simple system described by the following state equations:

$$\begin{cases} x_{k+1} &= sin(0.1 * k) \cdot x_k + \omega_k \\ y_k &= cos(x_k) + \alpha_k \end{cases} \tag{3.29}$$

with ω_k, α_k two white, normalized centered Gaussian vectors. For the simulation, we set $x_0 = 0$, and for the Kalman filter we use $\hat{x}_0 = 0, \sigma_0 = 1$.

This system has a linear evolution equation, which makes it a perfect candidate for a Kalman filter. On the other hand, the observation equation is non-linear, and can't be processed reliably as-is by a Kalman filter (we would have to linearize it or to propagate sigma-points to approximate the distributions).

Figure 3.4 illustrates the integration of that system without any observation over 100 steps with a Kalman filter and an Interval estimator. Figure 3.5 and the following figures represent only the width of these tubes.

Now, if after integrating the trajectory we had two observations y_{41}, y_{81} at times $t = 41, t = 81$, the interval observer would be able to contract the trajectory at these instants as illustrated on Figure 3.6.

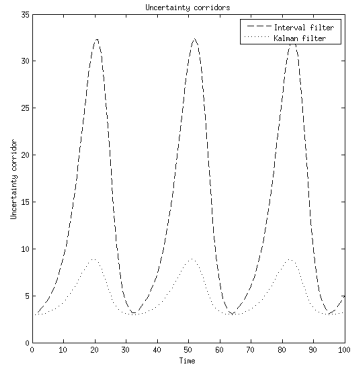


Figure 3.5: Illustration of the system described in Equation 3.29 over a time horizon of 100 steps.

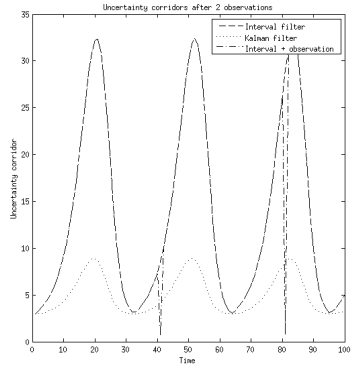
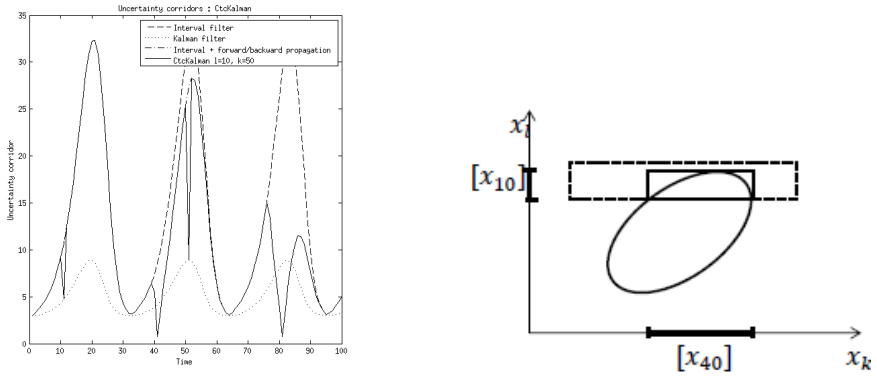
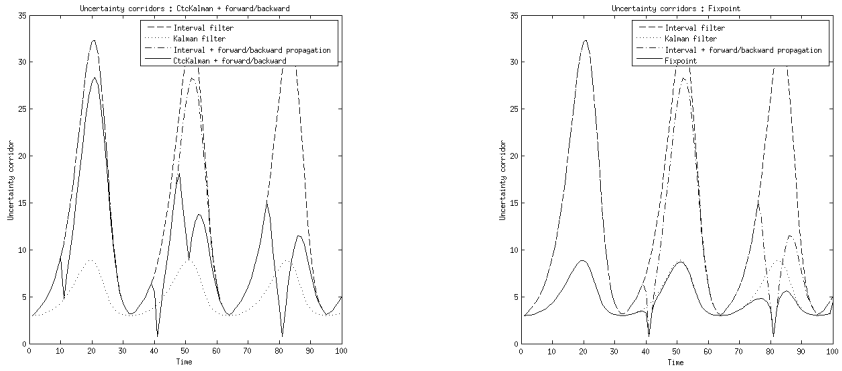


Figure 3.6: Contraction of the trajectory using two observations at times $t = 41, t = 81$



(a) (b)

Figure 3.7: Application of the Kalman contractor on $[x_{10}] \times [x_{50}]$



(a) (b)

Figure 3.8: Forward-Backward propagation after using the Kalman contractor on $[x_{10}] \times [x_{50}]$(a), fix-point after applying the Kalman contractor at several random locations followed by forward-backward propagation

Now, we randomly chose two instants $k = 10, l = 50$, and apply the Kalman contractor on the box $[\boldsymbol{x}] = [x_{10}] \times [x_{50}]$. The result of the contraction is shown on Figure 3.7(a), while $\boldsymbol{\Sigma}_{10,50}$ is shown on Figure 3.7(b).

Figure 3.8(a) shows the trajectory tube when a forward-backward-propagation is performed after having applied the Kalman contractor on $[x_{10}] \times [x_{50}]$, and Figure 3.8(b) shows the trajectory tube after repeating the process for several random instants up to the fix point. We see that the trajectory tube contracted with the Kalman filter is more precise than the one contracted using the forward-backward contractor only and the one using the Kalman filter in dead-reckoning only. The resulting tube can be visualized on Figure

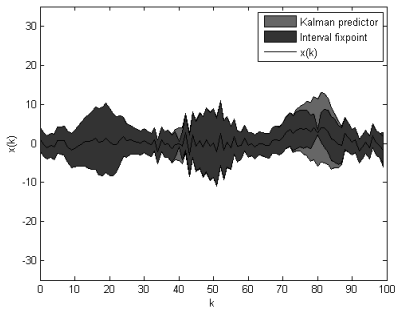


Figure 3.9: Interval tube at the fix-point compared to the Kalman predictor tube

3.6 DISCUSSION

Of course, the example treated in Section 3.5 is both simplistic and naive. Instead of applying the Kalman contractor several times at random locations, it should be applied only to the instants where observations have been made and where we want to benefit from these observations (ex: at the current instant). In the case of this simulation, since only two observations were made at instants 41 and 81, we should have contracted the state x_{100} with only two Kalman contractors: $\mathcal{C}_{41,100}, \mathcal{C}_{81,100}$.

3.7 CONCLUSION

In this Chapter, we have presented a new contractor that enables us to propagate uncertainties on the motion equations in the way a Kalman filter does but in a set-membership framework, without having to re-integrate the motion equations every-time new data have to be processed. This allows to treat systems with a linear evolution equation and a non-linear observation equations, much like the model of our underwater robot motion. In the next chapter, we will focus on the observation part of the Kalman filter equations. Assuming the robot does not move and accumulates measurements, we will see how a Kalman filter incorporates uncertainties in the observation equations. Finally, a method will be proposed which, as in this chapter, enables us to have Kalman-filter-like precision properties on the processing of observations, but in a reliable way.

4

NONLINEAR GAUSSIAN PARAMETER ESTIMATION

4.1 INTRODUCTION

In this chapter, we will zoom on the observation equations of the SLAM problem. Disregarding the dynamic aspect of the problem (the motion equations), and assuming that the poses $\mathbf{x}_k$ at each sampling time is known to belong to a given set $\mathbb{X}_k$ and are *variables* of the problem (i.e.: we are not interested in contracting the $\mathbb{X}_k$'s anymore), we show that finding the set $\mathbb{X}^{b_i}$, the set containing the i^{th} beacon, amounts to solving a classical static parameter estimation problem.

Recall the equations of the SLAM problem:

$$\begin{cases} \mathring{\boldsymbol{x}} &= \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \\ \boldsymbol{y} &= \boldsymbol{g}(\boldsymbol{x}) \end{cases}. \tag{4.1}$$

In this chapter we will completely disregard the evolution equation (the first one), and focus on the observation equation (the second line). Given a vector $\boldsymbol{y}$ of normally distributed measurements available at some instant k, we will want to retrieve $\boldsymbol{x}_k$. While in the previous chapter we were dealing with linear equations subject to additive Gaussian noise, here we have nonlinear equations, also subject to additive Gaussian noise.

Parameter set estimation deals with characterizing a set (preferably small) which encloses the parameter vector $\mathbf{p}$ of a parametric model from data collected on the system. In the context of bounded-error estimation [63, 73, 93], the measurement error is assumed to be bounded and computing the feasible set for $\mathbf{p}$ can be described as a set inversion problem for which interval methods are particularly efficient, even when the model is nonlinear [44]. In a probabilistic context, the error is not anymore described by membership intervals, but by probability density functions (pdf), and the correspondence between the two approaches has been studied by several researchers [8] [57]. For instance, in a Bayesian context, the Bayes rule makes it possible to get the posterior pdf for $\mathbf{p}$ (see, e.g., [30]). The set to estimate becomes the minimal volume credible set [7] and corresponds to the minimal volume set, in the parameter space, which contains $\mathbf{p}$ with a given probability η. This problem cannot be cast into a set inversion problem but existing interval methods can still be used [35]. Now, the approach is limited to few parameters (typically less than 3) and few measurements (typically less than 10).

Recently, an original approach [21] named *Sign-Perturbed Sums* (SPS) has proposed to constructs non-asymptotic confidence regions which is guaranteed to contain the true parameters with a given probability η. This approach has been used for nonlinear models to compute confidence regions which have not a minimal volume (at least in the Gaussian case). Interval analysis has also been considered to deal with the SPS method [56] to compute guaranteed confidence regions. Other methods such as [28] or [36] are also able to compute guaranteed confidence regions using interval analysis, but the computed set is not of minimal volume and it is difficult to evaluate the resulting pessimism.

There exist other approaches that combine bounded-error estimation with probabilistic estimation [2] [67] [34], [96], or that use other frameworks such as random sets [64, 83] or fuzzy-sets [29, 88], but all these methods do not solve a problem which is expressed only in terms of probabilities only and can thus not be used to compute confidence regions.

This chapter considers a problem which can be considered as classic on probabilistic parameter estimation: compute a set which encloses the parameter vector with a fixed probability η. Our main contribution is to be able to solve this problem in a reliable way in the case where the error is Gaussian and the model is nonlinear. To our knowledge, no such method exists in the literature.

This chapter considers the case where the error is additive and Gaussian.

In Section 4.2 we will recall the principles of set-inversion for the specific case where the noises are Gaussian, and will propose different models for the set to invert, by taking into account a confidence threshold. Section 4.3 recalls the principle of the linear Gaussian estimation that will be used for comparison. Section 4.4 will illustrate the proposed approach on simple simulated examples, and compare the discussed approaches as well as a classical linear Gaussian estimator. Section 2.6 will conclude the chapter.

4.2 BOUNDING A WHITE GAUSSIAN RANDOM VECTOR

In this section we will treat the problem of parameter estimation in a set-membership framework. Consider the following static parameter estimation problem

$$\mathbf{y} = \boldsymbol{\psi}(\mathbf{p}) + \mathbf{e}, \tag{4.2}$$

where $\boldsymbol{\psi}$ is the model, $\mathbf{y} \in \mathbb{R}^n$ is the vector of all measurements (which is known) and $\mathbf{e}$ is the error vector. Fact 15 shows that even if only treating the special case of a centered, normalized Gaussian error vector $\mathbf{e}$, the results are applicable with any non-white Gaussian error vector. Therefore, in an effort to maximize readability and without loss of generality we assume that $\mathbf{e}$ follows a Gaussian dis-

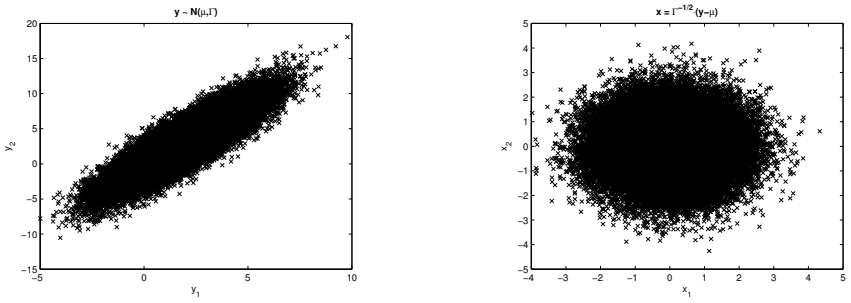


Figure 4.1: Several realizations of the random vector $\mathbf{y} \sim \mathcal{N}\left(\boldsymbol{\mu} = \begin{pmatrix} 2 & 3 \end{pmatrix}^T, \boldsymbol{\Gamma} = \begin{pmatrix} 3 & 5 \\ 5 & 11 \end{pmatrix}\right)$ and their images $\mathbf{x} = \boldsymbol{\Gamma}^{-1/2}(\mathbf{y} - \boldsymbol{\mu})$

tribution with zero mean and a covariance matrix equals to $\mathbf{I}^n$, i.e.: $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$.

Remark 15. As illustrated by Figure 4.1, a random variable $\mathbf{y}$ following a normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ can always be whitened into a random variable $\mathbf{x}$ distributed as $\mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ by the affine transform $\mathbf{x} = \boldsymbol{\Gamma}^{-1/2}(\mathbf{y} - \boldsymbol{\mu})$.

Proof. We have $\boldsymbol{E}[\boldsymbol{y}] = \boldsymbol{E}[\boldsymbol{\mu}] + \boldsymbol{\Gamma}^{1/2}\boldsymbol{E}[\boldsymbol{x}] = \boldsymbol{\mu}$. The covariance $\boldsymbol{\Gamma}_y$ of $\boldsymbol{y}$ is $\boldsymbol{E}\left[(\boldsymbol{y} - \boldsymbol{\mu}) \cdot (\boldsymbol{y} - \boldsymbol{\mu})^T\right] = \boldsymbol{E}\left[\boldsymbol{\Gamma}^{1/2} \cdot \boldsymbol{x} \cdot \left(\boldsymbol{\Gamma}^{1/2} \cdot \boldsymbol{x}\right)^T\right] = \boldsymbol{E}\left[\boldsymbol{\Gamma}^{1/2} \cdot \boldsymbol{x} \cdot \boldsymbol{x}^T \cdot \boldsymbol{\Gamma}^{1/2^T}\right] = \boldsymbol{\Gamma}^{1/2}.\, \boldsymbol{E}[\boldsymbol{x}] \cdot \boldsymbol{E}\left[\boldsymbol{x}^T\right] \cdot \boldsymbol{\Gamma}^{1/2^T} = \boldsymbol{\Gamma}^{1/2} \cdot \boldsymbol{I} \cdot \boldsymbol{I}^T \cdot \boldsymbol{\Gamma}^{1/2^T} = \boldsymbol{\Gamma}$. □

When the y_i's are bounded, finding the set of all feasible parameters compatible with the model and the measurements consist of computing the pre-image of the set $\mathbb{Y} = [y_1] \times [y_2] \times \cdots \times [y_n]$ through the function $\boldsymbol{\psi}(\mathbf{p}) = (\psi_1(\mathbf{p}), \psi_2(\mathbf{p}), \ldots, \psi_n(\mathbf{p}))^T$, and is a classical set-inversion problem:

$$\mathbb{P} = \boldsymbol{\psi}^{-1}(\mathbb{Y}). \tag{4.3}$$

Now, what happens when the observations are not bounded anymore, as is often the case for true systems? In such situation, it is still possible to bound the observations, except that the resulting bounding will not be guaranteed to contain the true measurement, but will contain it with a probability ν_i. Assuming the observations are *independent*, the set $\mathbb{Y}_\eta = [y_1]_{\nu_1} \times [y_2]_{\nu_2} \times \cdots \times [y_n]_{\nu_n}$ will contain $\mathbf{y}$ with a probability $\eta = \prod_{i=1}^n \nu_i$.

Definition 16. Define the function $\mathbf{f}(\mathbf{p}) = \mathbf{y} - \boldsymbol{\psi}(\mathbf{p})$ corresponding to the error $\mathbf{e}$ and a set $\mathbb{E}_\eta$ containing $\mathbf{e}$ with a probability η. The *probabilistic set* associated to $\mathbb{E}_\eta$ is defined as

$$\hat{\mathbb{P}}_{\mathbb{E}_\eta} = \mathbf{f}^{-1}(\mathbb{E}_\eta). \tag{4.4}$$

$\hat{\mathbb{P}}_{\mathbb{E}_\eta}$ contains $\mathbf{p}$ with a prior probability of η [36]. As a consequence, a probabilistic set estimation can be viewed as a set inversion problem

n	$\alpha(\eta)$
1	$\alpha = \sqrt{2}\text{erf}^{-1}(\eta)$
2	$\alpha = \sqrt{-2 \cdot \log(1-\eta)}$
$n \gg 1$	$\alpha \simeq \sqrt{n + 2\sqrt{n} \cdot \text{erf}^{-1}\left(2\eta + \text{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}$

Table 4.1: $\alpha(\eta)$ for $n = 1, 2$ and $n \gg 1$

for which guaranteed set inversion techniques could be used. Now, there exists several manners to choose such a set $\mathbb{E}_\eta$. The purpose of this section is to study different bounding shapes for the case of a white, Gaussian measurement vector.

4.2.1 *Bounding with a sphere*

Recall from [75] some results useful to get a set which encloses the normal error $\mathbf{e}$ with a given probability η.

Theorem 17. *The minimal volume confidence region of probability η associated with $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ is the centered n-dimensional sphere $\mathbb{S}_\eta$ of radius α, where (α, η) are linked by the relation*

$$\eta = \int_0^{\alpha^2} \frac{z^{\left(\frac{n}{2}-1\right)} e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)} \cdot dz \tag{4.5}$$

where Γ_e the Euler function. For $n \in \mathbb{N}$ the Euler function satisfies

$$\Gamma(n) = (n-1)! \tag{4.6}$$

Proof. The random variable $z = \mathbf{e}^T \cdot \mathbf{e}$ follows a χ^2 distribution with n degrees of freedom whose probability density function is

$$\pi(z, n) = \frac{z^{\left(\frac{n}{2}-1\right)} \cdot e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)}. \tag{4.7}$$

The minimal volume confidence region $\mathbb{S}_\eta$ is the set of all $\mathbf{e}$ such that

$$z = \mathbf{e}^T \cdot \mathbf{e} \leq \alpha^2(\eta) \tag{4.8}$$

and the probability η to have $\mathbf{e} \in \mathbb{S}_\mathbf{J}$ is

$$\eta = \int_0^{\alpha^2} \pi(z, n) \cdot dz = \int_0^{\alpha^2} \frac{z^{\left(\frac{n}{2}-1\right)} e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)} \cdot dz. \tag{4.9}$$

□

For $n = 1$, $n = 2$ or n large, from the integral in Equation (4.5), we can have an expression of the radius $\alpha(\eta)$ [7] as recalled in Table 4.1.

In our context, the dimension of $\mathbf{e}$ is large and we can consider that the formula corresponding to $n \gg 1$ is correct.

Thus, from Theorem 17 we know that $\hat{\mathbb{P}}_{S_\eta}$, the invert of the η-confidence sphere is the smallest feasible set of parameters. Now, inverting a sphere is is not particularly well suited for interval methods. Indeed, each parameter appears many times in the expression of the sphere equation, and we know from [48] that an inclusion function is minimal if each variable appears only once in the function's expression.

Example 18. Consider two functions $\mathbf{f}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, and g from $\mathbb{R}^m$ to $\mathbb{R}$. With a measurement vector $\mathbf{y} \in \mathbb{R}^m$, expressions for these functions could be:

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x_1 + cos^2(x_2) - y_1 \\ \dots \\ x_1 + cos^2(x_2) - y_n \end{pmatrix} \tag{4.10}$$

$$g(\mathbf{z}) = \sum_{i=1}^{n=\dim(\mathbf{z})} z_i^2. \tag{4.11}$$

Both the f_i's and g expressions are made up of a finite composition of the operators $+, -, *, /$ and elementary functions (cos, sqr in our case). Moreover, the variables appear only once in their expression. Therefore, their inclusion functions $[\mathbf{f}]([\mathbf{x}]), [g]([\mathbf{z}])$ are *minimal*, i.e. $[\mathbf{f}]([\mathbf{x}])$ and $[g]([\mathbf{z}])$ are the smallest box and interval that contain $\mathbf{f}([\mathbf{x}])$ and $g([\mathbf{z}])$ [48]. If, however, we look at the function $\mathbf{h}$ from $\mathbb{R}^n$ to $\mathbb{R}$ with $\mathbf{h}(\mathbf{x}) = g \circ \mathbf{f}(\mathbf{x})$ defined by:

$$g \circ \mathbf{f}(\mathbf{x}) = \left(x_1 + cos^2(x_2) - y_1\right)^2 + \dots + \left(x_1 + cos^2(x_2) - y_n\right)^2 \tag{4.12}$$

its expression contains many occurrences of the x_i's, and consequently its natural inclusion function *may* not be minimal.

Experience shows that an inclusion function for the sphere involved in the equation describing $\mathbb{S}_\eta$ is indeed very pessimistic, and linear methods such as the centered inclusion function [48] or affine arithmetic [71] should be used to decrease this pessimism.

4.2.2 *Bounding with boxes*

Inverting a box is well suited for interval methods. We show here some theoretical results of inverting the box-hull of the η-sphere and the η-box, the box that contains $\mathbf{e}$ with a probability η.

4.2.2.1 *Hull of the $\eta-$sphere*

Definition 19. We define the box-hull $[\mathbb{S}_\eta]$ as the smallest box that contains $\mathbb{S}_\eta$.

It might be tempting to approximate $\mathbb{S}_\eta$ by its box-hull $[\mathbb{S}_\eta]$. However, as shown in Theorem 20, it is not a good idea since even with a low value of η, when the dimension of the sphere is high, the probability that $[\mathbb{S}_\eta]$ contains $\mathbf{e}$ tends toward 1. This fact is illustrated on Figure 4.2.

Theorem 20. *With $n \gg 1$, the probability ϕ_η to have $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n) \in [\mathbb{S}_\eta]$ is:*

$$\phi_\eta = erf\left(\sqrt{\sqrt{n} \cdot erf^{-1}\left(2 \cdot \eta + erf\left(-\frac{\sqrt{n}}{2}\right)\right) + \frac{n}{2}}\right)^n. \tag{4.13}$$

Proof. From Table 4.1 with $n \gg 1$, for a given confidence η, the radius α of $\mathbb{S}_\eta$ is

$$\alpha = \sqrt{2 \cdot \sqrt{n}\left[\text{erf}^{-1}\left(2\eta + \text{erf}\left(-\frac{\sqrt{n}}{2}\right)\right)\right] + n}. \tag{4.14}$$

Now, $[\mathbb{S}_\eta]$ is the Cartesian product of n intervals $[e_i]$ of length 2α:

$$[\mathbb{S}_\eta] = [e_1] \times [e_2] \times \cdots \times [e_n]. \tag{4.15}$$

From Table 4.1 with $n = 1$, we know that the probability to have $e_i \in [e_i]$ is

$$Pr([e_i]) = \text{erf}\left(\frac{\alpha}{\sqrt{2}}\right). \tag{4.16}$$

Therefore

$$Pr([\mathbf{e}]) = \prod_{i=1}^{n} Pr([e_i]) = \text{erf}\left(\frac{\alpha}{\sqrt{2}}\right)^n. \tag{4.17}$$

By combining (4.14) with (4.17), we get (4.13). □

4.2.2.2 *Bounding with the η-box*

When bounding with $[\mathbb{S}_\eta]$, the fact that ϕ_η grows with the number of measurements even when η is fixed is not desirable. Since $\mathbf{e}$ is a Gaussian vector and the e_i's are *independent*, an other obvious bounding shape for $\mathbf{e}$ is the box whose radius is derived from the n^{th} root of η.

Theorem 21. *The minimal volume box $\mathbb{B}_\eta$ which encloses $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ with a probability η is the centered cube with half-width*

$$\alpha = \sqrt{2} erf^{-1}(\sqrt[n]{\eta}). \tag{4.18}$$

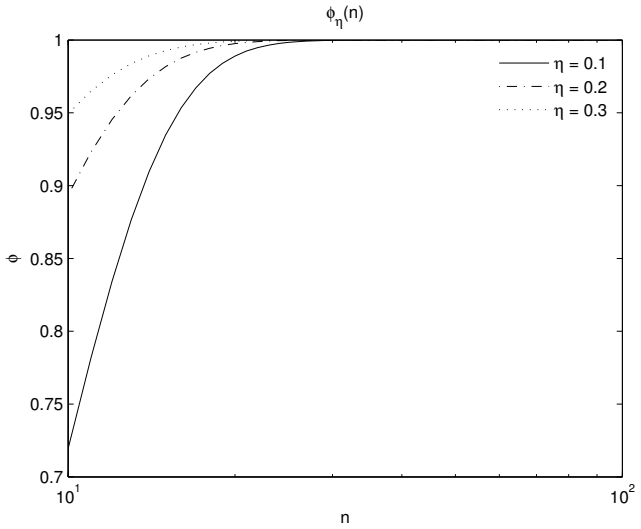


Figure 4.2: As the dimension n of $\mathbf{e}$ (or the number of measurements) grows, the probability $\phi_\eta(n)$ to have $\mathbf{e} \in [\mathbb{S}_\eta]$ grows as well.

Proof. The symmetry of the problem implies that $\mathbb{B}_\eta$ should be centered. Since the e_i are independent, we have:

$$\begin{aligned} \eta &= \Pr(\forall i, e_i \in [-\alpha, \alpha]) \\ &= \prod_{i=1}^{n} \Pr(e_i \in [-\alpha, \alpha]) \\ &= \prod_{i=1}^{n} \operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right) \\ &= \operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right)^n \end{aligned} \tag{4.19}$$

$$\alpha = \sqrt{2}\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right). \tag{4.20}$$

□

Despite its simplicity, Theorem 21 is often overlooked when choosing lower and upper bounds for experimental data. A measurement $\tilde{y}$ often comes with a variance σ (or covariance matrix $\mathbf{\Gamma}$ when $\tilde{\mathbf{y}}$ is not a scalar) given by the sensor, its data-sheet or a pre-treatment, and an interval representation for y is often chosen of the form

$$[y] = \tilde{y} + n \cdot \sigma \cdot [-1, 1] \tag{4.21}$$

with n an integer. This representation comes from the fact that when y is Gaussian, centered and normalized (i.e.: $y \sim \mathcal{N}(0,1)$), some typical values for its inverse cumulative distribution $\Phi^{-1}(p)$ are:

Using Fact 15 we easily get Expression 4.21. For example, $[y]_1 = \tilde{y} + \sigma \cdot [-1, 1]$ contains y with a probability of 68%, while $[y]_2 = \tilde{y} + 2 \cdot \sigma \cdot [-1, 1]$ contains y with a probability of 95%. This choice for bounding y is wrong in two frequent situations:

1. when y is not scalar Equation 4.21 is not valid anymore and is an underestimation, i.e.:

$$\Pr(\mathbf{y}) \in \mathbf{y} + \Phi^{-1}(p) \cdot \mathbf{\Gamma}^{1/2} \cdot [-1, 1]^{\times \dim(\mathbf{y})} < p \tag{4.22}$$

p	$\Phi^{-1}(p)$
68%	1
95%	2
99.7%	3

Table 4.2: Typical integer values for the probit function.

2. when N measurements $\tilde{y}_i$ are available, the probability that $\mathbf{f}^{-1}([\mathbf{y}])$ contains the true parameter $\mathbf{p}$ decreases with N as is easily shown from Equation 4.18:

$$\eta = \operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right)^N \tag{4.23}$$

with $\alpha = n \cdot \sigma$.

This observation is trivial, but is often overlooked, and leads to think for example that from $N = 100$ measurements each bounded with a probability 0.99 we compute $\hat{\mathbb{P}}$ that encloses $\mathbf{p}$ with the same probability, while the actual probability is actually only $0.99^N = 0.37$.

Now, bounding $\mathbf{e}$ with $\mathbb{S}_\eta$, $[\mathbb{S}_\eta]$ or $\mathbb{B}_\eta$ wont tolerate the cases where some data are outliers, i.e.: some of the e_i's do not belong to $\operatorname{proj}_i(\mathbb{S}_\eta)$, $\operatorname{proj}_i([\mathbb{S}_\eta])$ or $\operatorname{proj}_i(\mathbb{B}_\eta)$. In that case, of course, it simply means that either the event "$\mathbf{e}$ is not in the bounding" with its associated probability $1-\eta$ happened, or that our Gaussian assumption doesn't hold. In the latter case, other shapes enabling the measurement to occasionally be far from the origin should be chosen. Such a shape is the cross-shape, and interval methods are able to manipulate those shapes.

4.2.3 *Bounding with a relaxed box*

Interval methods allows to tolerate the presence of faulty measurements when their number is bounded by a known number, i.e.: there are at most a fixed number q of them. The robustness to outliers is obtained by relying on the Q-intersection operator [14, 47], which is more thoroughly discussed in Appendix A.

Solving the Q-relaxed CSP:

$$H^{\{q\}} : \begin{cases} \wedge^q_{i=\{1,\dots,N\}} & f_i(\mathbf{p}) = e_i \\ \mathbf{p} \in [\mathbf{p}], e_i \in [e_i] \end{cases} \tag{4.24}$$

where $\wedge^q$ denotes the logical q-relaxed AND operator, i.e. the operator that combine all the constraints except q at most, amounts to computing $\mathbf{f}^{-1}\left([\mathbf{e}]^{\{q\}}\right)$, where $[\mathbf{e}]^{\{q\}}$ is the q-relaxed box defined below in Definition 22.

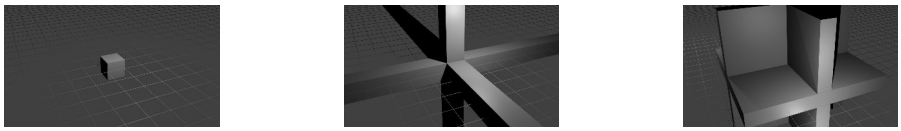


Figure 4.3: Illustration of the q-relaxed box $[\mathbf{x}]^{\{q\}} = [-1,1]^{\times 3}$ with $q = 0$ (left), $q = 1$ (middle), $q = 2$ (right, truncated in $[-10,10]^{\times 3}$ for visibility purpose)

Definition 22. A q-relaxed box $[\mathbf{x}]^{\{q\}}$ associated to a box $[\mathbf{x}] \subset \mathbb{R}^n$ is a union of boxes of the form:

$$[a_1] \times [a_2] \cdots \times [a_n] \tag{4.25}$$

where at least q of the $[a_i]$'s are equal to $[-\infty, \infty]$ and the others correspond to the $[x_i]$'s.

We will write

$$[\mathbf{x}]^{\{q\}} = \prod_{i \in \{1,\ldots,n\}}^{\{q\}} [x_i] \tag{4.26}$$

which defines the q-relaxed Cartesian product operator.

Example 23. For a box $[\mathbf{x}] = [x_1] \times [x_2] \times [x_3]$, we have:

$$\begin{aligned}
[\mathbf{x}]^{\{0\}} &= [x_1] \times [x_2] \times [x_3] \\
[\mathbf{x}]^{\{1\}} &= [-\infty,\infty] \times [x_2] \times [x_3] \cup [x_1] \times [-\infty,\infty] \times [x_3] \cup [x_1] \times [x_2] \times [-\infty,\infty] \\
[\mathbf{x}]^{\{2\}} &= [-\infty,\infty]^{\times 2} \times [x_3] \cup [x_1] \times [-\infty,\infty]^{\times 2} \cup [-\infty,\infty] \times [x_2] \times [-\infty,\infty] \\
[\mathbf{x}]^{\{3\}} &= [-\infty,\infty]^{\times 3}
\end{aligned} \tag{4.27}$$

This example is illustrated on Figure 4.3 for the case where $[x_1] = [x_2] = [x_3] = [-1,1]$.

We see on Figure 4.3 that the q-relaxed box is simply an unmodified box when $q = 0$, an infinite axis-aligned cross-shape when $q = 1$, and its extension to higher dimension when $q > 1$. Finally, when $q = \dim(\mathbf{x})$ we have $[\mathbf{x}]^{\{q\}} = \mathbb{R}^n$. It is obvious that we have the following relation:

$$[\mathbf{x}]^{\{0\}} \subset [\mathbf{x}]^{\{1\}} \subset \cdots \subset [\mathbf{x}]^{\{\dim(\mathbf{x})\}}.$$

As such, when we increase the number q of relaxed measurement, the measurement space is less and less constrained. It would be tempting to say that the measurement space covered by the q-relaxed box increases with q, but it would have no sense since the volume of the q-relaxed box is infinite as soon as q is greater than 0. A consequence of this decrease of constraint is that $\Pr\left(\mathbf{e} \in (\mathbb{B}_\eta)^{\{q\}}\right) \geq \Pr\left(\mathbf{e} \in \mathbb{B}_\eta\right)$, i.e.: the q-relaxation of $\mathbb{B}_\eta$ is of course an over-approximation of $\mathbb{B}_\eta$. We define the domain $\mathbb{B}_\eta^q$ as follows:

Definition 24. With either $\nu = \Pr(e_i) \in \text{proj}_i\left(\mathbb{B}_\eta^q\right)$ or q fixed, the set $\mathbb{B}_\eta^q$ is the smallest q-relaxed box containing $\mathbf{e}$ with a probability η.

The goal of the following statements is to provide a way to compute ν, q such that $\Pr(\mathbf{e}) \in \mathbb{B}_\eta^q = \eta$ when the dimension of $\mathbf{e}$ is large.

Theorem 25. *With ν the probability of having $e_i \in [e_i]$, $N = dim(\mathbf{e})$, γ the probability to have strictly more than q outliers we have the following relation:*

$$\gamma(q, N, \nu) = \frac{1}{2} \cdot \left(1 + erf\left(\frac{N \cdot (1-\nu) - q - 1}{\sqrt{2 \cdot N \cdot \nu (1-\nu)}}\right)\right) \tag{4.28}$$

Proof. The number k of inliers among the N normalized measurements follows a binomial distribution, and the probability of having exactly k inliers among N measurements is

$$\beta(k, N, \nu) = \frac{N!}{k!(N-k)!} \cdot \nu^k \cdot (1-\nu)^{N-k}. \tag{4.29}$$

Thus, the probability of having strictly more than q outliers is

$$\gamma(q, N, \nu) = \sum_{k=0}^{N-q-1} \beta(k, N, \nu). \tag{4.30}$$

The central limit theorem states that when N is large, which is the case for us, $\gamma(q, N, \nu)$ can be approximated by a normal distribution with mean $N \cdot \nu$ and variance $N \cdot \nu \cdot (1-\nu)$.

The probability that there are more than q outliers among the N $\nu-$confidence intervals is therefore

$$\gamma(q, N, \nu) = \frac{1}{2} \cdot \left(1 + \text{erf}\left(\frac{N \cdot (1-\nu) - q - 1}{\sqrt{2 \cdot N \cdot \nu (1-\nu)}}\right)\right) \tag{4.31}$$

from which we deduce for a given γ, ν, N:

$$q(N, \nu, \gamma) = N(1-\nu) - 1 - \sqrt{2 \cdot N \cdot \nu (1-\nu)} \cdot \text{erf}^{-1}(2\gamma - 1). \tag{4.32}$$

□

It naturally follows that the probability η of having more than q outliers is:

$$\eta(q, N, \nu) = 1 - \frac{1}{2} \cdot \left(1 + \text{erf}\left(\frac{N \cdot (1-\nu) - q - 1}{\sqrt{2 \cdot N \cdot \nu (1-\nu)}}\right)\right). \tag{4.33}$$

Example 26. With $N = 1000$, $\nu = 0.9$, $\eta = \nu^N$, Figure 4.4 shows the probability γ as a function of q, the number of relaxed measurements. For each value of ν, we see γ decreases abruptly from almost 1 to approximately 0 when q becomes higher than some threshold.

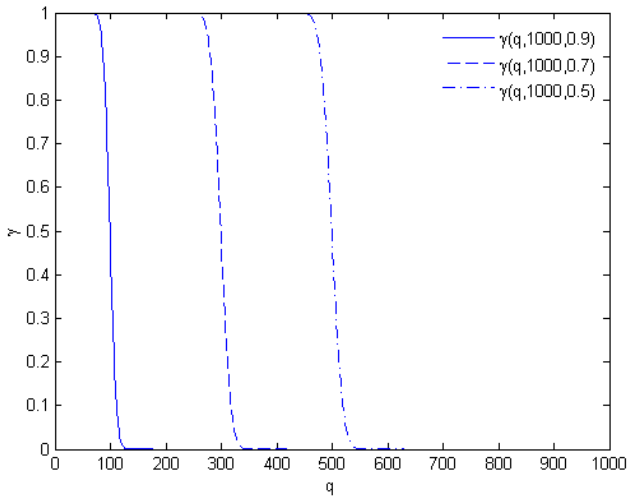


Figure 4.4: Risk of having more than q outliers as a function of q with $\nu \in \{0.5, 0.7, 0.9\}$. When q grows, the the risk diminishes.

Corollary 27. *With N, γ fixed, an expression for q as a function of ν is:*

$$q_{N,\gamma}(\nu) = N(1-\nu) - 1 - \sqrt{2N\nu(1-\nu)} \cdot erf^{-1}(2\gamma - 1). \tag{4.34}$$

Proof. From Equation 4.28 it is trivial to obtain Equation 4.34. □

Corollary 28. *With N, γ, q fixed, an expression for ν is:*

$$\nu_{N,\gamma}(q) = \begin{cases} \alpha_{N,\gamma}(q) + \beta_{N,\gamma}(q) & \text{when } \gamma < 0.5 \\ \alpha_{N,\gamma}(q) - \beta_{N,\gamma}(q) & \text{when } 0.5 \leq \gamma < 1 \end{cases} \tag{4.35}$$

with

$$\begin{cases} \alpha_{N,\gamma}(q) = & \frac{erf^{-1}(2\gamma-1)^2 + N - q - 1}{2erf^{-1}(2\gamma-1)^2 + N} \\ \beta_{N,\gamma}(q) = & \frac{erf^{-1}(2\gamma-1) \cdot \sqrt{\left(2Nq + Nerf^{-1}(2\gamma-1)^2 + 2N - 2q^2 - 4q - 2\right)/n}}{2erf^{-1}(2\gamma-1)^2 + n} \end{cases} \tag{4.36}$$

Proof. From Theorem 25 it is straightforward to get a polynomial $P_{N,\gamma,q}$ in ν, which is easily solved by a formal mathematical solver such as Mathematica [95] or Maple [4] yielding the results in Equations 4.35 and 4.36. □

Corollary 29. *Given q, the smallest q-relaxed box $\mathbb{B}_\eta^q$ of n intervals which contains $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$, with a probability η is:*

$$[\mathbf{x}]^{\{q\}} = \prod_{i \in \{1,\ldots,n\}}^{\{q\}} \varepsilon_{n,1-\eta}(q) \cdot [-1, 1] \tag{4.37}$$

with

$$\varepsilon_{n,1-\eta}(q) = \begin{cases} \sqrt{2} erf^{-1}\left(1 - \alpha_{n,1-\eta}(q) - \beta_{n,1-\eta}(q)\right) & \text{when } \eta < 0.5 \\ \sqrt{2} erf^{-1}\left(1 - \alpha_{n,1-\eta}(q) + \beta_{n,1-\eta}(q)\right) & \text{when } 0.5 \leq \eta < 1 \end{cases} \tag{4.38}$$

with $\alpha_{N,\gamma}(q), \beta_{N,\gamma}(q)$ *defined in Equation 4.36 and where* $\mathbb{B}_\eta^q$ *is the smallest in the sense that there exists no other q-relaxed box* $\mathbb{B}_\eta^{'q}$ *such that* $\mathbb{B}_\eta^{'q} \subset \mathbb{B}_\eta^q$.

Proof. From Table 4.1, we know that the radius r of a confidence interval containing the correct value with a probability ν is

$$r(\nu) = \sqrt{2} \cdot \text{erf}^{-1}(\nu) \tag{4.39}$$

and by inserting Equation 4.35 in 4.39 we get 4.37. □

Now, For computing $\hat{\mathbb{P}}_{\mathbb{B}_\eta^q}$ from a set of N measurements, the couple of parameters (q, ν) has to be chosen: fixing a value for q amounts to fixing ν and reverse, so there is one degree of freedom left in the problem. One might ask whether there is a choice for (q, ν) that minimizes the volume of $\hat{\mathbb{P}}_{\mathbb{B}_\nu^q}$. We define here two criterion that could be further investigated in order to constraint this degree of freedom surplus.

Definition 30. Given a set $\mathbb{E}_\eta$ which contains $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ with a probability η. We define the pessimism of $\mathbb{E}_\eta$ as:

$$\text{pess}(\mathbb{E}_\eta) = \text{prob}(\mathbf{e} \in \mathbb{E}_\eta \setminus \mathbb{S}_\eta) - \text{prob}(\mathbf{e} \in \mathbb{S}_\eta \setminus \mathbb{E}_\eta) \tag{4.40}$$

The set that minimizes the pessimism is of course the η-sphere $\mathbb{S}_\eta$. It is generally difficult to compute analytically this pessimism. It however easy to evaluate it using a Monte-Carlo method. A choice for q would be one that minimizes the pessimism.

Definition 31. Given a set $\mathbb{E}_\eta$ which contains $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ with a probability η. We define the accuracy of $\mathbb{E}_\eta$ as:

$$\text{acc}(\mathbb{E}_\eta) = \int_{\mathbb{E}_\eta} \pi^2(\mathbf{e}) \cdot d\mathbf{e} \tag{4.41}$$

This quantity corresponds to the mathematical esperance of the likelihood inside $\mathbb{E}_\eta$, and can be seen as the complementary of the pessimism. Again, the set that maximizes the accuracy is the η-sphere $\mathbb{S}_\eta$. There is still no easy way to generally compute the accuracy of a set analytically, but it should be easy to evaluate it using Monte-Carlo methods.

In the frame of this work, we will arbitrarily fix a value for q or ν when computing $\mathbb{B}_\eta^q$.

It should be noted that $\mathbb{B}_\eta^q$ is more suited to deal with situation where rare events such as a multi-path occur. Of course, in those situation the Gaussian assumption simply doesn't hold anymore, but while there is a high risk that computing $\hat{\mathbb{P}}_{\mathbb{S}_\eta}, \hat{\mathbb{P}}_{[\mathbb{B}_\eta]}, \hat{\mathbb{P}}_{\mathbb{B}_\eta}$ will return the empty set $\emptyset$ even for high values of η, it is less likely to also be the case for $\hat{\mathbb{P}}_{\mathbb{B}_\eta^q}$. Indeed, cross-shapes are more suited to represent

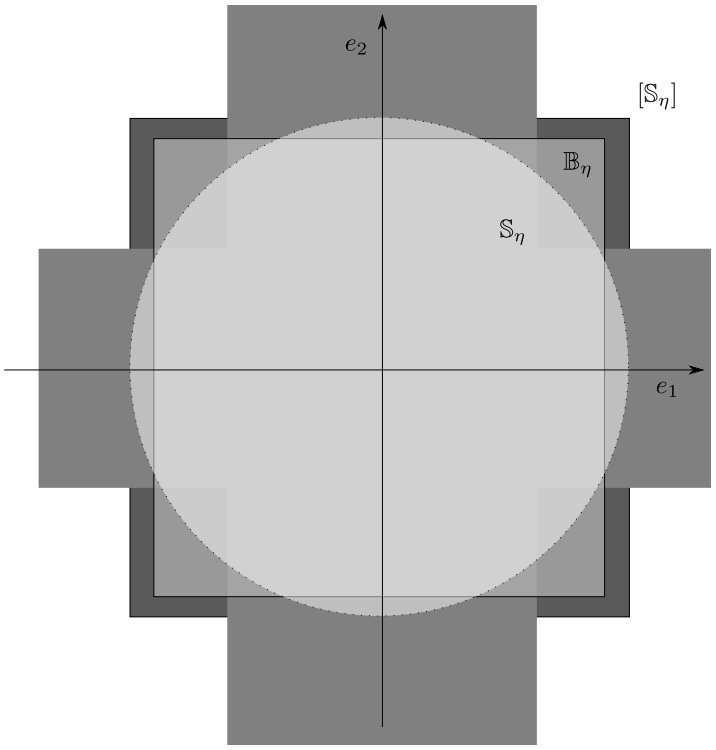


Figure 4.5: Idealized representation for (darkest to lightest): $[\mathbb{S}_\eta]$, $\mathbb{B}_\eta^q$, $\mathbb{B}_\eta\mathbb{S}_\eta$

heavy-tailed distributions, where most of the events are distributed around the mean, but where the likelihood to be far away from the mean doesn't decrease exponentially with the distance from the mean. When events that should be rare occur frequently, the error is heavy-tailed and $\mathbb{B}_\eta^q$ should be chosen.

4.3 LINEARIZING METHOD

A classical probabilistic approach to estimate the set $\hat{\mathbb{P}}_{\mathbb{E}_\eta}$ is the *Maximum Likelihood Estimator* (MLE) that uses a linear approximation of the model. Unfortunately, the linearization error cannot be quantified in a reliable way, and therefore the result of the MLE cannot be trusted. Additionally, we will show that a Maximum Likelihood Estimation for the case of additive Gaussian noise is only able to represent ellipsoidal sets even if $\mathbb{P}_{\mathbb{E}_\eta}$ can be an arbitrary set.

The linearization method searches for the parameter vector $\hat{\mathbf{p}}$ which maximizes the likelihood function

$$L\left(y_i|\mathbf{p}\right) = \prod_i \pi\left(y_i|\mathbf{p}\right). \tag{4.42}$$

In our case, recall that the noise follows a centered, normal distribution such that $\pi\left(y_i|\mathbf{p}\right) \propto e^{-\left(\psi_i(\mathbf{p})-y_i\right)^2}$ and Equation 4.42 becomes:

$$L\left(y_i|\mathbf{p}\right) \propto \prod_i e^{-\left(\psi_i(\mathbf{p})-y_i\right)^2}. \tag{4.43}$$

Now, as the logarithm function is monotonic increasing, this is equivalent to minimizing:

$$\lambda(\mathbf{p}) = -log(L(y_i|\mathbf{p})) = \sum_i (\psi_i(\mathbf{p}) - y_i)^2 \tag{4.44}$$

which corresponds to a non-linear least-square minimization problem.

It seems reasonable to assume that the true value for $\mathbf{p}$ is close to the minimizer $\hat{\mathbf{p}}$ and that $\lambda(\mathbf{p})$ can be approximated by a second order Taylor development of Equation 4.44 around $\hat{\mathbf{p}}$. Since $\hat{\mathbf{p}}$ minimizes $\lambda(\mathbf{p})$, the gradient of λ at $\hat{\mathbf{p}}$ is zero and we get:

$$\lambda(\mathbf{p}) \simeq \lambda(\hat{\mathbf{p}}) + \frac{1}{2} \cdot (\mathbf{p} - \hat{\mathbf{p}})^T \cdot \mathbf{H}_\lambda \cdot (\mathbf{p} - \hat{\mathbf{p}}) \tag{4.45}$$

where $\mathbf{H}_\lambda$ is the Hessian matrix of of $\lambda(\mathbf{p})$ in $\hat{\mathbf{p}}$.

With $\mathbf{e}^T \cdot \mathbf{e} \leq \alpha^2(\eta)$ and $\mathbf{e}^T \cdot \mathbf{e} = \sum_i (\psi_i(\mathbf{p}) - y_i)^2$ a confidence region that contains $\mathbf{p}$ with a probability η is:

$$\lambda(\hat{\mathbf{p}}) + \frac{1}{2} \cdot (\mathbf{p} - \hat{\mathbf{p}})^T \cdot \mathbf{H}_\lambda(\hat{\mathbf{p}}) \cdot (\mathbf{p} - \hat{\mathbf{p}}) \leq \alpha^2(\eta) \tag{4.46}$$

which is an ellipsoid of radius $2 \cdot (\alpha^2(\eta) - \lambda(\hat{\mathbf{p}}))$ and shape matrix $\mathbf{H}_\lambda(\hat{\mathbf{p}})$. Note that $\mathbf{H}_\lambda(\hat{\mathbf{p}})$ corresponds to the observed Fisher information matrix at $\hat{\mathbf{p}}$ which is the inverse of the covariance matrix $\mathbf{\Sigma}_{\hat{\mathbf{p}}}$ for the estimated maximum likelihood parameter $\hat{\mathbf{p}}$ from which the Cramér-Rao bound is derived [82].

4.4 COMPARISON

We will now illustrate our method with three illustrative test-cases involving parameter estimation under white, additive Gaussian noise. In these test-cases, the method based on the q-relaxed intersection discussed in Section 4.2.3 will not be tested, and will instead be more thoroughly covered in Appendix C to conduct robust, reliable Maximum Likelihood Estimation (MLE) on real-data.

4.4.1 *Test-case 1*

Consider the following model

$$y(t) = p_2 \cdot e^{-p_1 \cdot t} + p_1 \cdot e^{-p_2 \cdot t} + w(t) \tag{4.47}$$

where $t \in \{0, 0.01, 0.02, \ldots, 12\}$ and $w(t)$ is a white centered Gaussian noise with a variance $\sigma^2 = 1$. Figure 4.6 represents the collected data $y(t)$.

Figure 4.7 represents the three sets $\mathbb{P}_{0.99}$ obtained by an inversion of $[\mathbb{S}_{0.99}]$, $\mathbb{B}_{0.99}$ and $\mathbb{S}_{0.99}$. This comparison confirms that the box-hull inversion $\mathbb{P}_{[\mathbb{S}_{0.99}]}$ is too pessimistic. Figure 4.8 illustrates a situation

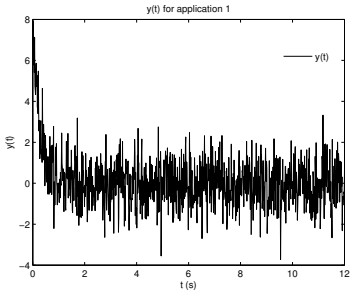


Figure 4.6: Measurements $y(t)$ for Test-case 1

where $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}} \not\subset \hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$. From Theorem 34 we could have expected an inclusion. Now, this example is quite atypical: the parametric model is not globally identifiable, *i.e.*, p_1 and p_2 can be interchanged without any effect on the output. Figure 4.8 also represents the confidence ellipsoid generated by the linear estimator. Due to the non identifiability problem, we have two global minimizers. We chose to draw the ellipsoid centered around the minimizer corresponding to the true parameter vector $\mathbf{p}^*$. Otherwise, the 0.99 ellipsoid would not contains $\mathbf{p}^*$.

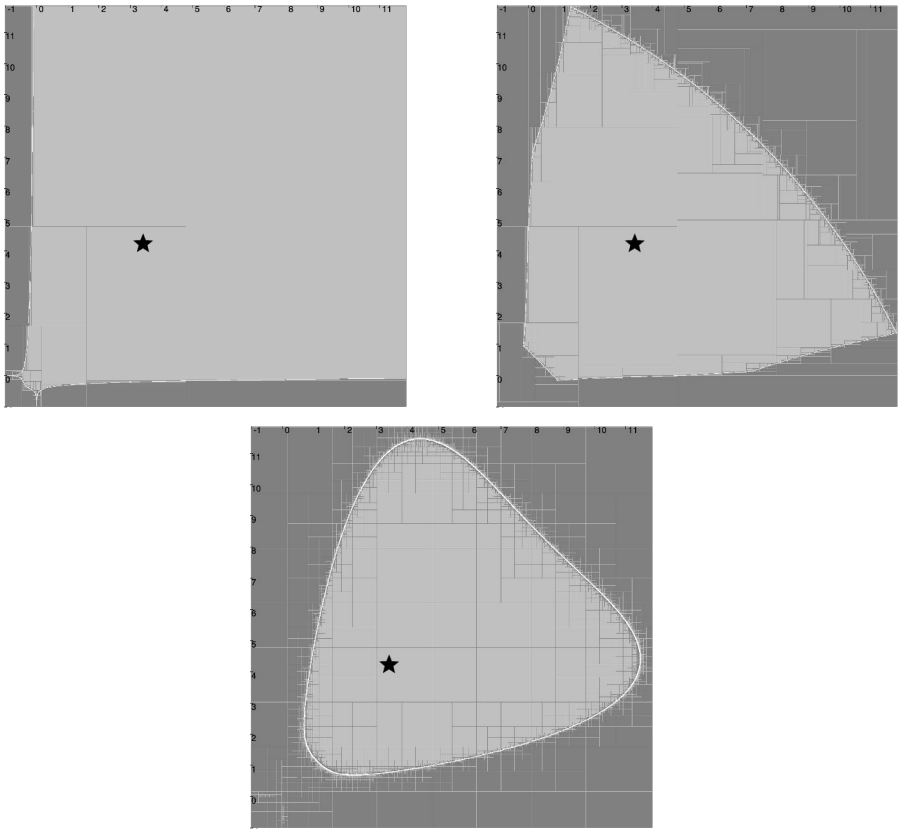


Figure 4.7: $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ (top left), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$(top right), $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (bottom). The black star is the true parameters vector.

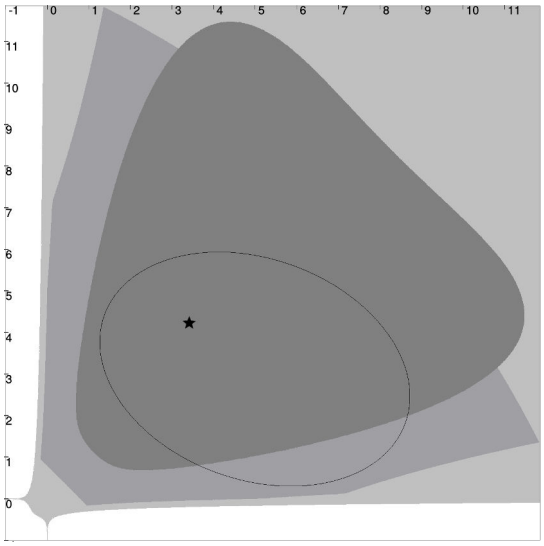


Figure 4.8: Superposition of $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$(gray), $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (dark gray), and the 0.99 confidence ellipse obtained with a linear estimator. The black star is the true parameter vector $\mathbf{p}^*$

4.4.2 *Test-case 2*

Consider the following model studied in [45]

$$y(t) = 20 \cdot e^{-p_1 \cdot t} - 8 \cdot e^{-p_2 \cdot t} + w \tag{4.48}$$

which is similar to the model of Test-case 1 but the model is now identifiable. Again, $w(t)$ is a centered normal noise with a unit variance. We collected 1000 measurements for $y(t)$ at different times $t \in [0, 25]$ as represented on Figure 4.9.

Figure 4.10 shows that the inversion $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ of the confidence sphere $\mathbb{S}_{0.99}$ is more precise than the inversion of $[\mathbb{S}_{0.99}]$ and $\mathbb{B}_{0.99}$. The set $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ has two disjoint components at a confidence level $\eta = 0.99$. Figure 4.11 shows that the linear estimator was able to capture the correct parameters vector.

Remark 32. Figure 4.10 shows that the proposed approach suffers from an important pessimism: the border of the computed set is quite thick, and the generated sub-paving is not minimal. This is due to the multiple-occurrences in the parameter variables in the expression of the inequalities describing $\mathbb{S}_\eta$. Interval methods are sensitive to this type of situation which adds pessimism in the propagation of uncertainties [41]. To limit this phenomena, linear approximations such as the centered or affine forms of the constraints could be used.

4.4.3 *Test-case 3*

In this example, a lost underwater vehicle tries to get its position by gathering range-only measurements to three beacons. The position

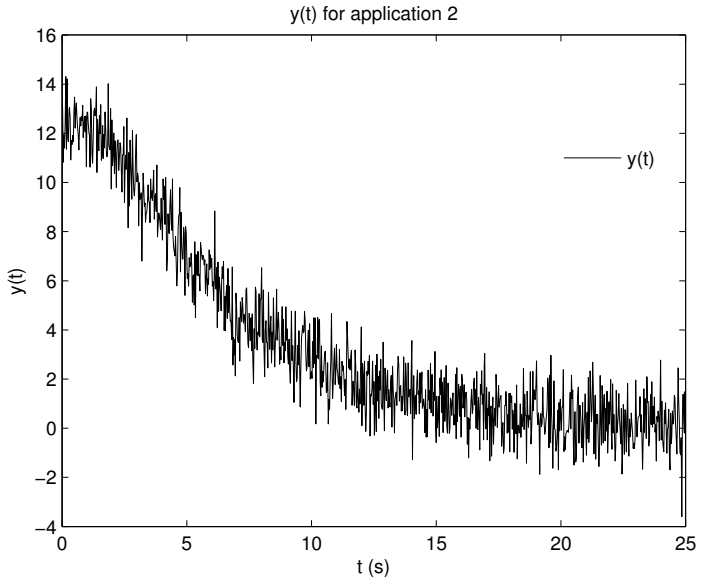


Figure 4.9: Collected data $y(t)$ for Test-case 2

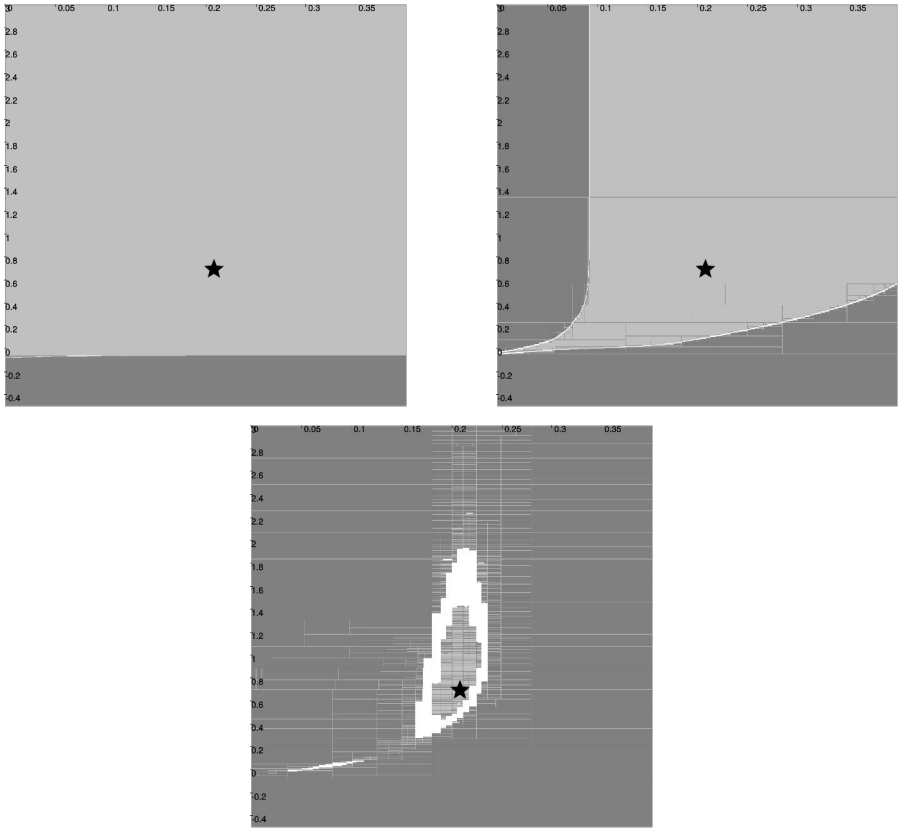


Figure 4.10: $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ (top left), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (top right), $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (bottom). The black star is the true parameter vector $\mathbf{p}^*$

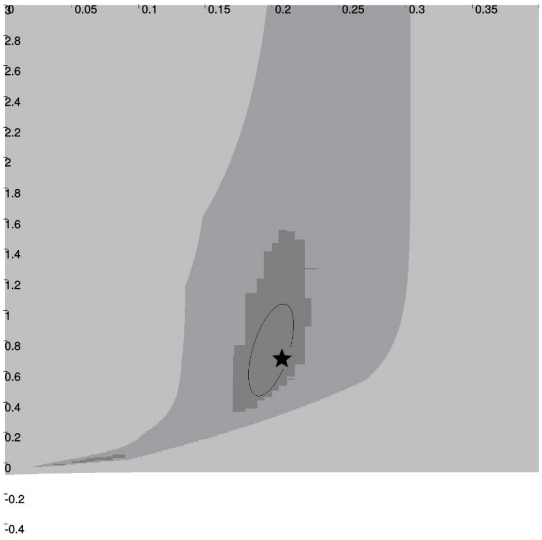


Figure 4.11: Superposition of $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$(gray), $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (dark gray), and the 0.99 confidence ellipse obtained with a linear estimator.

$\mathbf{x}_j = \begin{pmatrix} x_j & y_j & z_j \end{pmatrix}$ of the j^{th} beacon is precisely known from a previous survey of the area, as well as the altitude z_m of the robot, thanks to a pressure sensor.

The three beacons are almost aligned, which causes a bad conditioning. The robot is assumed to be static during the acquisition. For each measurement $\tilde{d}_i$ to the beacon j we have

$$\tilde{d}_{ij} = \sqrt{(x_j - x_m)^2 + (y_j - y_m)^2 + (z_j - z_m)^2} + w \tag{4.49}$$

where w is a white centered Gaussian noise, whose variance is given by the sensor for each measurement. The signals associated to the three beacons are pictured in Figure 4.12.

From Figure 4.14, we observe that $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}} \subset \hat{\mathbb{P}}_{\mathbb{B}_{0.99}} \subset \hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$, which confirms that the $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ is more precise than the two other confidence regions. Figure 4.15 is the superposition of $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}, \hat{\mathbb{P}}_{\mathbb{B}_{0.99}}, \hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ and the 0.99 confidence ellipse of a linear estimator. While the linear estimator gives an estimate that is consistent (it contains the true solution), it is obvious that it doesn't fully capture the underlying banana-shaped probability density function, which is more accurately seized by our nonlinear methods.

Table 4.3 compares the time it takes to compute $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$, $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$, $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ on a classical laptop for the three test-cases. As it could have been anticipated, it is clear that inverting boxes, which are convenient representations for interval methods, takes much less time than inverting a sphere.

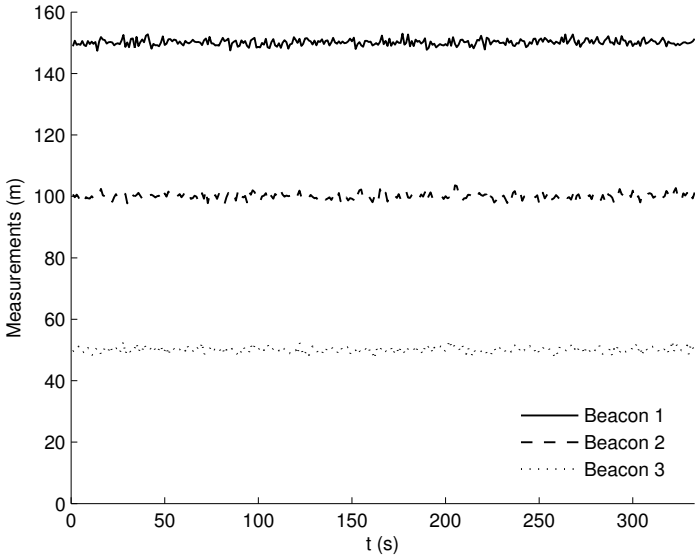


Figure 4.12: Range signals received from the three beacons

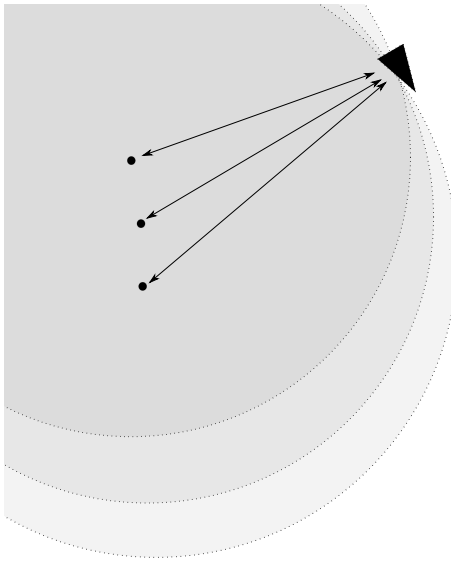


Figure 4.13: An underwater robot stays fixed on the seafloor and gathers range measurements from 3 beacons whose positions are known, in order to estimate its position

Computation time (s)	Test-case 1	Test-case 2	Test-case 3
$\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$	35 sec	1sec	26sec
$\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$	62sec	6sec	45sec
$\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$	839sec	89sec	510sec

Table 4.3: Computation times for Test-cases 1, 2 and 3

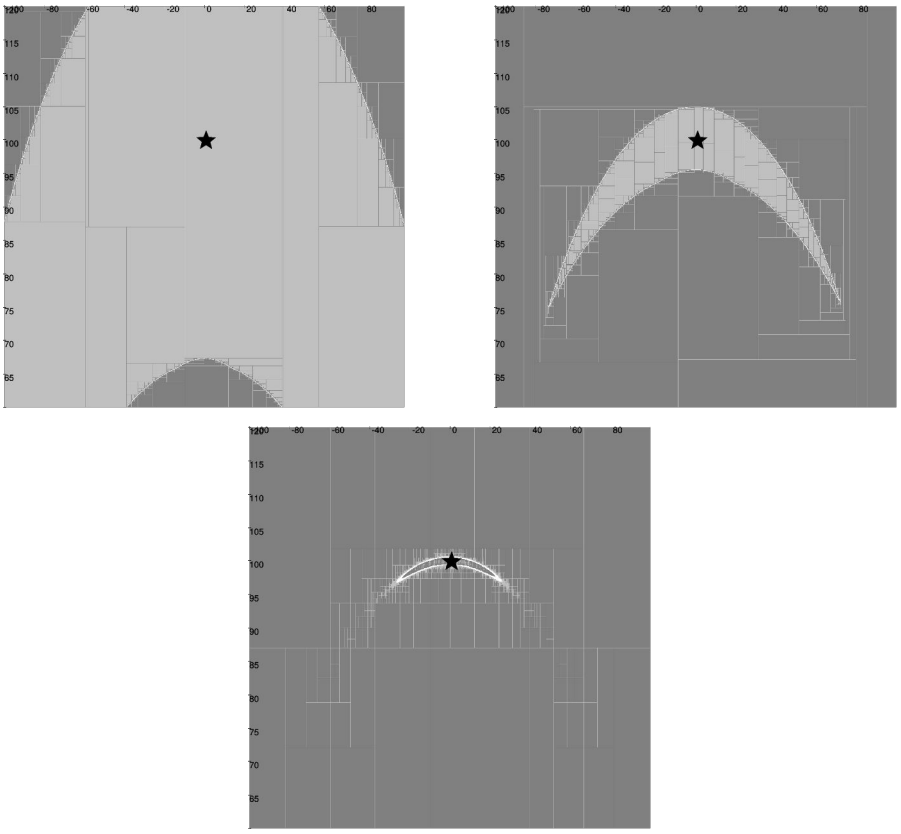


Figure 4.14: $\hat{\mathbb{P}}_{[S_{0.99}]}$ (top left), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (top right), $\hat{\mathbb{P}}_{S_{0.99}}$ (bottom). The black star represents $\mathbf{p}^*$

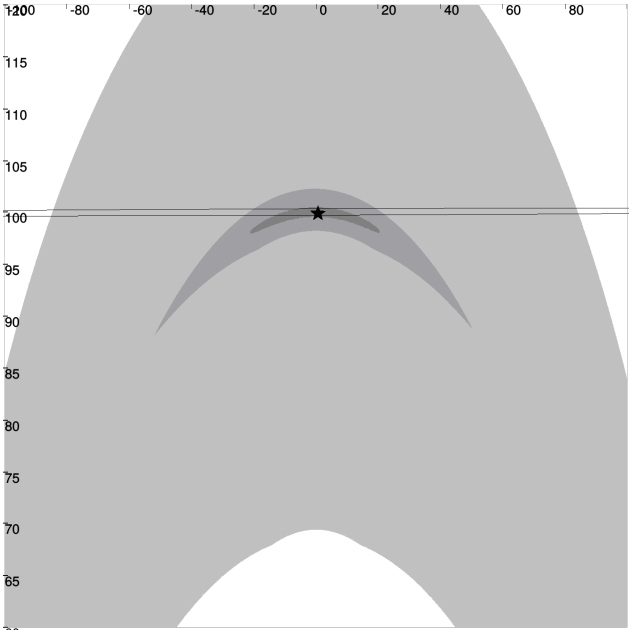


Figure 4.15: Superposition of $\hat{\mathbb{P}}_{[S_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (gray), $\hat{\mathbb{P}}_{S_{0.99}}$ (dark gray), and the 0.99 confidence ellipse of the linear estimator (black). The black star represents $\mathbf{p}^*$

4.5 DISCUSSION

We will now want to see if we could have anticipated some of the results of Section 4.4. More specifically, are there properties of the proposed shapes that we could use to prove that one is more precise than the other, where a bounding shape $\mathbb{S}_a$ is *more precise* than $\mathbb{S}_b$ if $\hat{\mathbb{P}}_{\mathbb{S}_a} \subset \hat{\mathbb{P}}_{\mathbb{S}_b}$.

A trivial case happens when $\mathbb{S}_a \subset \mathbb{S}_b$: in that case, $\mathbb{S}_a$ is obviously more precise than $\mathbb{S}_b$ since $\hat{\mathbb{P}}_{\mathbb{S}_a} \subset \hat{\mathbb{P}}_{\mathbb{S}_b}$. This is the case for the two non-relaxed box bounding: as shown in Theorem 33, since $\mathbb{B}_\eta \subset \left[\mathbb{S}_\eta\right]$, bounding $\mathbf{e}$ with $\mathbb{B}_\eta$ will result in more precise estimation than bounding it with $\left[\mathbb{S}_\eta\right]$.

Theorem 33. *Bounding* $\mathbf{e}$ *with* $\mathbb{B}_\eta$ *is more precise than bounding with the box-hull* $\left[\mathbb{S}_\eta\right]$ *of* $\mathbb{S}_\eta$.

Proof. $\mathbb{B}_\eta$ and $\left[\mathbb{S}_\eta\right]$ are both centered boxes. It then suffices to show that $\text{rad}\left(\mathbb{B}_\eta\right) < \text{rad}\left(\mathbb{S}_\eta\right)$ to prove that $\mathbb{B}_\eta \subset \left[\mathbb{S}_\eta\right]$. From Equation 4.18 we have $\text{rad}\left(\mathbb{B}_\eta\right) = \sqrt{2} \cdot \text{erf}^{-1}\left(\sqrt[n]{\eta}\right)$ and from Equation 4.14 we have $\text{rad}\left(\mathbb{S}_\eta\right) = \sqrt{2 \cdot \sqrt{n}\left(\text{erf}^{-1}\left(2\eta + \text{erf}\left(-\frac{\sqrt{n}}{2}\right)\right)\right) + n}$. □

Now, is it preferable to use $\mathbb{B}_\eta$ or $\mathbb{S}_\eta$? We don't have an inclusion relation for these two representation, as their intersection $\mathbb{B}_\eta \cap \mathbb{S}_\eta$ is non-empty. From Theorem 17 we know that $\mathbb{S}_\eta$ is the minimal volume set that contains $\mathbf{e}$ with a probability η, and is therefore *de facto* the most precise shape. However, we propose a reflection based on the *ratio of the volumes* of the two sets when the dimension d of $\mathbf{e}$ grows. In that sense, Theorem 34 shows that the ratio of the volume of $\mathbb{S}_\eta$ over $\mathbb{B}_\eta$ tends towards zero when d tends towards infinity.

Theorem 34. *We have:*

$$\lim_{n\to\infty} \frac{vol\left(\mathbb{S}_\eta\right)}{vol\left(\mathbb{B}_\eta\right)} = 0 \tag{4.50}$$

Proof. Since the volume of an n-dimensional sphere $\mathbb{S}_\eta$ of radius α is

$$V_n = \frac{\pi^{n/2} \cdot \alpha^n}{\Gamma_e\left(n/2 + 1\right)}, \tag{4.51}$$

we have:

$$\rho_\eta\left(n\right) = \frac{\text{vol}\left(\mathbb{S}_\eta\right)}{\text{vol}\left(\mathbb{B}_\eta\right)} = \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{n} \cdot \text{erf}^{-1}\left(2\eta + \text{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}^n}{\Gamma_e(n/2+1)} \Big/ \left(2\sqrt{2}\text{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^n. \tag{4.52}$$

The Stirling formula $\Gamma_e\left(n+1\right) = n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ implies that:

$$\rho_\eta\left(n\right) \sim \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{2} \cdot \text{erf}^{-1}\left(2\eta + \text{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}^n}{\sqrt{2\pi\frac{n}{2}}\left(\frac{n}{2e}\right)^{n/2} \cdot \left(2\sqrt{2}\text{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^n}. \tag{4.53}$$

Now, $2\eta + \operatorname{erf}\left(\frac{-\sqrt{n}}{2}\right) \sim 2\eta - 1$. Therefore:

$$\rho_\eta(n) \sim \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{n} \cdot \operatorname{erf}^{-1}(2\eta - 1)}^{n}}{\sqrt{\pi n}\left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot \left(2\sqrt{2}\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^{n}}. \tag{4.54}$$

Since $n + 2\sqrt{n} \cdot \operatorname{erf}^{-1}(2\eta - 1) \sim n$, we get:

$$\begin{aligned} \rho_\eta(n) &\sim \frac{\pi^{n/2} \cdot \sqrt{n}^{n}}{\sqrt{\pi n}\left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot \left(2\sqrt{2}\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^{n}} \\ &= \frac{\pi^{n/2} \cdot n^{\frac{n}{2}} \cdot (2e)^{\frac{n}{2}}}{\sqrt{n}\left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot \left(2\sqrt{2}\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^{n}} \\ &\sim \frac{(e\pi)^{\frac{n}{2}}}{\left(2\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right)\right)^{n}} \\ &= \left(\frac{\sqrt{e\pi}}{2\operatorname{erf}^{-1}\left(\sqrt[n]{\eta}\right)}\right)^{n}. \end{aligned} \tag{4.55}$$

which converges to zero as n increases. □

Theorem 34 doesn't serve as a proof that $\mathbb{S}_\eta$ is more precise than $\mathbb{B}_\eta$, but it gives an intuition on how pessimist each set becomes when we add more measurements: as the number of measurements grows, the volume of $\mathbb{S}_\eta$ becomes more and more negligible in comparison of the volume of $\mathbb{B}_\eta$.

Such intuition based on the volume can not be extended when treating the case of relaxed boxes, because we know from Definition 22 that it is obvious that $\operatorname{vol}\left(\mathbb{B}_\eta^q\right)$ is infinite when q is positive. In the framework of this thesis we therefore do not further try to discuss the pessimism of $\mathbb{B}_\eta^q$, and its precision relatively to $\mathbb{B}_\eta$ or $\mathbb{S}_\eta$. That notion of precision may even depend on the model considered. Experimentation shows however that this representation is precise for our problem, and that it can naturally accommodate the presence of outliers in an elegant way, as will be demonstrated in Appendix C.

4.6 CONCLUSION

In this chapter, we have presented a new approach for parameter estimation of nonlinear models with additive Gaussian noise. The resulting method makes it possible to compute a set which contains the parameter vector with a given probability. The main contribution of this chapter is that the results are guaranteed, which is not the case for existing approaches. Indeed, although existing methods are also able to provide an estimation of such a confidence region of probability η, they perform some linearization without quantifying the corresponding approximation error.

Three simulated test-cases were presented and compared to existing and linear methods. The computing time is clearly a limitation and our method cannot be considered yet as suitable for online estimation. A future extension would be to adapt this work to be able to

run in real time, so it could be embedded in a sensor's software for example.

5 CONCLUSIONS AND PROSPECTS

5.1 CONCLUSION

In this thesis, we tackled the problem of simultaneously localizing an underwater vehicle while mapping a set of beacons lying on the seafloor using range-only measurements. Chapter 1 was dedicated to the modeling of the SLAM problem, as well as the most commonly used sensors in the subsea domain and the sensors used in the scope of this thesis. We highlighted the fact that for our scenario, using a high grade INS coupled with a DVL, the motion equation can be considered linear. On the other hand, the observation equations, the equations that link the measurement gathered by the vehicle with its state variables, are non-linear and non bijective. In Chapter 2 we reviewed two of the most common approaches used to solve this problem, the Extended Kalman Filter, a probabilistic filter, and the Robust State Observer which is based on set-membership methods using interval analysis, and compared them on real data. From that comparison, we observed that the reliability and the robustness of the interval approach came at the cost of a very coarse precision. We made the hypothesis that the Gaussianity of the motion and observation noises were not exploited extensively. Additionally, since the motion equation of our SLAM problem are linear and the observation equations nonlinear, we observed that the Kalman filter was the most appropriate tool to treat the motion equations, while an interval approach would be best suited to treat the observation. The following two chapters were dedicated to finding a way to combine both approaches. In Chapter 3, we focused on the propagation of the motion model noises. We observed that for a random walk model, a Kalman filter's uncertainty grows as a square root of time, while an interval based filter's uncertainty grows linearly with time, leading, in part, to its huge pessimism. In addition, contrarily to a Kalman filter, a Robust State Observer is not recursive and is based on a fix-point method on a sliding horizon. Therefore, its computational complexity is dependent on the sampling rate of the inertial sensors. By reformulating the Kalman filter's prediction equations in an integral form, we showed that the joint probability distribution of two states $\boldsymbol{x}_k, \boldsymbol{x}_l$ at different instants k, l is Gaussian, and that the covariance matrix $\boldsymbol{\Gamma}_{k,l}$ of the vector $(\boldsymbol{x}_k, \boldsymbol{x}_l)^T$ can be easily computed. A contractor on the confidence ellipsoid for $(\boldsymbol{x}_k, \boldsymbol{x}_l)^T$ was defined, and enables us to benefit from the square root of time propagation of uncertainties of the Kalman filter, but in a set-membership framework. This reformulation not only also

enables us to pre-integrate the motions equation, removing the computational dependency on the motion sensors sampling rate, but the fact that we can stay in a set-membership framework means we can naturally treat our observation equations with interval methods, staying in a coherent, unique structure. Now Chapter 4 was dedicated to the processing of the observations, namely the range measurements. Assuming the noise on the observations $\boldsymbol{y} = (y_1, y_2, \ldots, y_i)^T$ is white, additive and Gaussian, we were interested in the possible boundings of such such a vector, and their effects on the confidence associated with the resulting estimation. We concluded that bounding with a sphere is the most precise option, but also the most computationally expensive. An other option, based on the q-relaxed intersection was explored more in depth in Appendix C. Treating $\boldsymbol{y}$ with q-relaxed methods not only allows us to compute precise confidence regions when the noise is not Gaussian, but we also explored the situation when a confidence region is not enough, and a close-to-punctual estimation is needed. For that situation there exist an estimator, GOMNE, which is best suited. We have shown that in our context, GOMNE is a maximum likelihood estimator, and that it can compute a precise, close to punctual estimation of the 2D position of an acoustic beacon on real data. In Appendix B, we presented a contractor based on linear matrix inequalities, which enables us to contract optimally sets for general quadratic constraints, for which simple forward-backward methods are not optimal. The development of this contractor was motivated by the fact that for the method developed in Chapter 3, we need a contractor on possibly high dimensional ellipsoid, which are described by quadratic constraints. Finally, Appendix A quickly introduced interval methods to the unfamiliar reader.

5.2 PROSPECTS

A future development of this thesis is the combination of the Kalman contractor with the methods developed in Chapter 4 and in Appendix C to get the desired real-time, reliable, precise and robust estimator for the SLAM problem. Indeed, in this thesis we have shown why interval methods are pessimistic with additive Gaussian noises, and proposed tools to reduce this pessimism, but we haven't combined them to obtain a complete estimator. An other possible development would be to study the use of the Kalman contractor in the case where the attitude provided by the INS is noisy. Indeed, in our scenario, the attitude is provided with an extreme precision, allowing us to neglect the associated uncertainty, but such a precision is not available with cheaper INS based, for example, on MEMS sensors. In that sense, the relative formulation of the 6DOF equations of motion proposed in [31] could be the way to our the Kalman contractor for the 6DOF SLAM problem with range-only measurement and low-cost INS.

Part III

APPENDICES

In the Part iii of this manuscript, we will present some of the research conducted during this thesis that were not directly related to the subject of localization and mapping. But first, in Appendix A, we will briefly recall the main concepts behind set-theoretical tools used in this thesis, and their implementation using interval analysis. Appendix B will present a contractor based on Linear Matrix Inequalities, that is used among other things to implement the Kalman Contractor introduced in Chapter 3 for the general case. Finally, in Appendix C, as a continuation of Chapter 4, we will show how to retrieve a punctual, statistically meaningful estimation from a set of measurements based on probabilistic considerations, for where estimating a confidence set is not precise enough for practical applications.

A

INTRODUCTION TO INTERVAL ANALYSIS

This section briefly introduces the basics of interval analysis, which is the main tool used in this thesis for representing uncertainties. For an in-depth presentation of interval analysis the reader is referred to [40, 78]. In the first section, the motivation for developping interval methods in first place is presented. The second section will present interval arithmetic.

A.1 INTRODUCTION

To represent numbers, computers use a binary system whose precision and extent are limited by the number of bits on which they are coded. The binary coding makes it impossible to exactly represent the set of real numbers. The accumulation of imprecisions can become, along the computations, a real problem.

An illustration of the problems that can occur with rounding error can be found in [71] and is recalled in Example

Example 35. Consider the following equation:

$$f(x,y) = 33.75y^6 + x^2\left(11x^2y^2 - y^6 - 121y^4 - 2\right) + 5.5y^8 + \frac{x}{2y} \quad \text{(A.1)}$$

By evaluating f at the point $x = 77617$ and $y = 33096$ in python on a 32 bits platform, we get:

$$f(77617, 33096) = -3.9425235 \cdots * 10^{29} \quad \text{(A.2)}$$

while the true result is:

$$f(77617, 33096) = -\frac{54767}{66192} = -0.8273960599. \quad \text{(A.3)}$$

As can be seen, the true result has nothing to do with the value computed by the computer.

Interval analysis was developped by Moore to obtain a range of values instead of an approximated value. A real number is then represented, not by an approximated value anymore, but by an interval containing the exact value. It is for example possible to express the π number with a precision to the tenth by using the interval $[3.14, 3.15]$.

Definition 36. An interval $[x]$ is defined as the set of real numbers x between its lower bound $\underline{x}$ and its upper bound $\bar{x}$:

$$[x] = [\underline{x}, \bar{x}] = \{x \in \mathbb{R} | \underline{x} \leq x \leq \bar{x}\}. \quad \text{(A.4)}$$

.

Definition 37. The center of a non-empty interval is given by:

$$\text{mid}\left([x]\right) = \frac{\underline{x} + \bar{x}}{2}. \tag{A.5}$$

.

Definition 38. The width of a non-empty interval is defined by:

$$w\left([x]\right) = \bar{x} - \underline{x}. \tag{A.6}$$

We write $\{x\} = [x, x]$ a punctual interval.

An interval being a set, set-membership operations can be defined over intervals.

Definition 39. The union of two non-empty intervals $[x]$ and $[y]$ satisfies:

$$[x] \sqcup [y] = \left[min\left(\underline{x}, \underline{y}\right), max\left(\bar{x}, \bar{y}\right)\right]. \tag{A.7}$$

Example 40. Example:

$$[-1, 7] \sqcup [12, 19] = [-1, 19]. \tag{A.8}$$

It should be noted that $[x] \cup [y]$, not to be confused with $[x] \sqcup [y]$, is not considered as an interval in the framework of this thesis.

Definition 41. The intersection of two non-empty intervals $[x]$ and $[y]$ satisfies:

$$[x] \cap [y] = \begin{cases} \left[max\left\{\underline{x}, \underline{y}\right\}, min\left\{\bar{x}, \bar{y}\right\}\right] & \text{if } max\left\{\underline{x}, \underline{y}\right\} \leq min\left\{\bar{x}, \bar{y}\right\} \\ \emptyset & \text{otherwise} \end{cases} \tag{A.9}$$

A.2 INTERVAL ARITHMETICS

Operations over real numbers, such as the division, multiplication, addition, subtraction, can be extended to intervals.

For two intervals $[x]$ and $[y]$ and an operator $\diamond \in \{+, -, *, /\}$, we define $[x] \diamond [y]$ as the smallest interval containing all feasible values for $x \diamond y$ when $x \in [x]$ and $y \in [y]$, or:

$$[x] \diamond [y] = \left[\left\{x \diamond y \in \mathbb{R} | x \in [x], y \in [y]\right\}\right] \tag{A.10}$$

In the case of closed intervals, we have:

$$\begin{cases} [x] + [y] &= [\underline{x} + \underline{y}, \bar{x}, \bar{y}] \\ [x] - [y] &= [\underline{x} - \bar{y}, \bar{x} - \underline{y}] \\ [x] * [y] &= \left[min\left(\underline{xy}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\right), max\left(\underline{xy}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\right)\right] \end{cases}. \tag{A.11}$$

The inversion is given by:

$$\frac{1}{[y]} = \begin{cases} \emptyset & \text{if } [y] = [0,0] \\ [1/\bar{y}, 1/\underline{y}] & \text{if } 0 \notin [y] \\ [1/\bar{y}, \infty] & \text{if } \underline{y}{=}0 \text{ and } \bar{y}{>}0 \\ [-\infty, 1/\bar{y}] & \text{if } \underline{y} < 0 \text{ and } \bar{y} = 0 \\ [-\infty, \infty] & \text{if } \underline{y} < 0 \text{ and } \bar{y} > 0 \end{cases} . \tag{A.12}$$

Note that in the case of punctual intervals, these operations are identical to their equivalent for real numbers.

A.3 EXTENSION TO HIGHER DIMENSION

A vectorial interval, also called interval vector or box $[\boldsymbol{x}]$ of dimension n is the Cartesian product of n intervals. The set of all boxes of $\mathbb{R}^n$ is denoted by $\mathbb{IR}^n$.

$$[\boldsymbol{x}] = [x_1] \times [x_2] \times \cdots \times [x_n] \tag{A.13}$$

An interval vector can be represented by an axis-aligned box.

Most of the operations defined for intervals can be extended to their vectorial counterparts. The upper and lower bounds for a box are:

$$\begin{aligned} \underline{\boldsymbol{x}} &= (\underline{x}_1, \ldots, \underline{x}_n)^T \\ \bar{\boldsymbol{x}} &= (\bar{x_1}, \ldots, \bar{x_n})^T \end{aligned} . \tag{A.14}$$

The width of a box is

$$w([\boldsymbol{x}]) = \max_{1 \le i \le n} w([x_i]) . \tag{A.15}$$

And operations on intervals can be extended to operations on boxes by considering interval computation on each component of the boxes.

A.4 INCLUSION FUNCTION

The image of $\boldsymbol{f}([\boldsymbol{x}])$ of a box is not a box in general:

$$\boldsymbol{f}([\boldsymbol{x}]) = \{\boldsymbol{f}(\boldsymbol{x}) \,|\boldsymbol{x} \in [\boldsymbol{x}]\} . \tag{A.16}$$

There exist no general representation to manipulate the image set of a box through a function. We will then use interval functions, called inclusion functions, which is defined as the function returning the interval hull of $\boldsymbol{f}([\boldsymbol{x}])$:

$$[\boldsymbol{f}]([\boldsymbol{x}]) = [\{\boldsymbol{f}(\boldsymbol{x}) \,|\boldsymbol{x} \in [\boldsymbol{x}]\}] . \tag{A.17}$$

The interval function $[\boldsymbol{f}] : \mathbb{IR}^n \to \mathbb{IR}^m$ is an inclusion function for $\boldsymbol{f} \to \mathbb{R}^n \to \mathbb{R}^m$ if and only if

$$\forall [\boldsymbol{x}] \in \mathbb{IR}^n, \boldsymbol{f}([\boldsymbol{x}]) \subset [\boldsymbol{f}]([\boldsymbol{x}]). \tag{A.18}$$

An inclusion function is minimal if, for all $[\boldsymbol{x}]$, $[\boldsymbol{f}]([\boldsymbol{x}])$ is the smallest box enclosing $\boldsymbol{f}([\boldsymbol{x}])$. This inclusion function is unique and is noted $[\boldsymbol{f}]^*$. Minimality is a desired property for an inclusion function, in order to reduce the pessimism.

An inclusion function is thin if the image of a punctual interval $\{x\}$ is a punctual interval, i.e. if $[f](\{x\}) = f(\{x\})$.

An inclusion function is convergent if for a series of boxes $[\boldsymbol{x}](k)$, we have:

$$\lim_{k\to\infty} w([\boldsymbol{x}](k)) = 0 \Rightarrow \lim_{k\to\infty} w([\boldsymbol{f}]([\boldsymbol{x}](k))) = 0. \tag{A.19}$$

The convergence property is generally required for the convergence of algorithms based on interval methods.

An inclusion function is said to be monotonic with respect to the inclusion if

$$[\boldsymbol{x}] \subset [\boldsymbol{y}] \Rightarrow [\boldsymbol{f}]([\boldsymbol{x}]) \subset [\boldsymbol{f}]([\boldsymbol{y}]). \tag{A.20}$$

A function f built as a finite number of composition of elementary functions such as sin, cos, exp, $\sqrt{.}$, $max \ldots$ and operators $+, -, *, /$. The simplest method to build an inclusion function for f is by replacing the scalar variables by their interval counterpart and replacing the elementary functions by their interval extension. The function obtained is called the natural inclusion function, and is denoted $[f]_n$. It is finite and monotonic with respect to the inclusion. Moreover, if f is solely made of continuous elementary functions, the $[f]_n$ is convergent.

Natural inclusion functions are usually not minimal because of dependencies between variables and the wrapping effect [58, 65]. However, the natural inclusion function will be minimal if each variable appears only once in the definition of f and if all operators and elementary functions used are continuous.

A.5 CONTRACTORS

Consider n_x real variables $x_i \in \mathbb{R}, i \in \{1, \ldots, n_f\}$ linked by n_f relations (or constraints) of the form:

$$f_j(x_1, x_2, \ldots, x_{n_x}) = 0, j \in \{1, \ldots, n_f\} \tag{A.21}$$

whereby f_j denotes the function for each coordinate j. We know that each variable x_i belongs to a domain $\mathbb{X}_i$. To simplify, we consider the domains as intervals noted $[x_i]$. We define $\boldsymbol{x} = (x_1, x_2, \ldots, x_{n_x})^T$ and the domain fo $\boldsymbol{x}$ as $[\boldsymbol{x}] = [x_1] \times [x_2] \times \cdots \times [x_{n_x}]$. We also note

that f is a function with coordinate functions f_j. We can therefore re-write Equation A.21 in a vector form $f(x) = 0$ and this represents a constraint satisfaction problem (CSP) that we can call $\mathscr{H}$ and write

$$\mathscr{H} : (f(x) = 0, x \in [x]). \tag{A.22}$$

Therefore, a CSP is composed of a set of variables, domains containing these variables and constraints. The solution set $\mathbb{S}$ of $\mathscr{H}$ is defined as

$$\mathbb{S} = \{x \in [x] \,|\, f(x) = 0\}. \tag{A.23}$$

Contracting a CSP $\mathscr{H}$ consists in replacing the domain $[x]$ with a smalled domain $[x'] \subset [x]$ without changing the solution set. We have $\mathbb{S} \subset [x'] \subset [x]$. We define the minimal contractor as the contractor replacing $[x]$ by the smallest box containing $\mathbb{S}$. An operator that allows the contraction of $\mathscr{H}$ is called a contractor.

Different existing basic contractors enable the contraction of CSP [16]. Depending on the class of the problem to solve, several methods can be used to optimally contract the domains: Gauss elimination, Gauss-Seidel Algorithm, Krawczyk method, Newton algorithm, etc... [40]. A minimal contractor dedicated to problems involving quadratic constraints is developped in this thesis in Appendix B. One of the most commonly used contractor in interval robotics is the forward-backward contractor [5], which contracts the domains of the CSP $\mathscr{H} : (f(x) = 0, x \in [x])$ by isolating each constraint separately. We suppose that each constraint has the form $f_j(x_1, x_2, \ldots, x_{n_x}) = 0$, and that the function f_j can be broken down into a series of operations involving operators and elementary functions such as $+, -, *, /$, etc. The the constraint is broken down into primary constraints.

A.6 SET INVERSION VIA INTERVAL ANALYSIS (SIVIA)

Intervals and boxes, thanks to interval arithmetics and inclusion functions, are easy to manipulate. Contractors allow us to reduce a box according to one or several constraints.

However, generally the solutions of the problems treated by set-membership approach are not boxes. Approximating the solution set by a box is, most of the time, not satisfying, particularily when the solution set is not connected.

A.6.1 *Subpaving*

In order to represent an arbitrary set $\mathbb{X}$ while being able to make use of the computation tools specific to intervals and boxes, we will use subpavings of $\mathbb{R}^n$.

A subpaving of a box $[x] \subset \mathbb{R}^n$ is the union of non-empty boxes of $[x]$ that do not overlap. Two boxes of a single subpaving can have

a non-empty intersection if they share a common border, but their interior must be disjoint. A subpaving of $[\boldsymbol{x}]$ that covers $[\boldsymbol{x}]$ is called a paving of $[\boldsymbol{x}]$.

We can enclose a set $\mathbb{X}$ between two sub-pavings: an inner approximation $\underline{\mathbb{X}}$ and an outer approximation $\bar{\mathbb{X}} = \underline{\mathbb{X}} \cup \Delta\mathbb{X}$ (where $\Delta\mathbb{X}$ is the subpaving of the border), such that:

$$\underline{\mathbb{X}} \subset \mathbb{X} \subset \bar{\mathbb{X}} \tag{A.24}$$

Knowing $\underline{\mathbb{X}}$ and $\bar{\mathbb{X}}$ brings us much insight on the set $\mathbb{X}$. If $\underline{\mathbb{X}}$ is not empty then $\mathbb{X}$ is not empty. If $\bar{\mathbb{X}}$ is empty, then $\mathbb{X}$ is empty as well. The volume of the subpaving of the border enables us to characterize the precision of the approximation.

When each box of a subpaving can be obtained by a finite number of bissections of the original box, the subpaving is called regular. Regular subpavings have several advantages, including their representability as a binary tree which limits their storage requirement and makes them easy to manipulate for operations like the union, intersection, the inclusion etc... [55].

A.6.2 *Set Inversion Via Interval Analysis*

We want to characterize the set $\mathbb{X}$ such that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}$ with $\boldsymbol{f} : \mathbb{X} \to \mathbb{Y}$, i.e. we want to characterize the pre-image $\mathbb{X} = \boldsymbol{f}^{-1}(\mathbb{Y})$. Given an inclusion function for $\boldsymbol{f}$, the Set Inversion Via Interval Analysis (SIVIA for short) algorithm computes an inner approximation $\underline{\mathrm{X}}$ and an outer approximation $\bar{\mathbb{X}}$ for $\mathbb{X}$. SIVIA can also be used together with contractors for a more efficient computation of the subpaving. The reader should refer to [41] for a detailed presentation of the algorithm.

A.6.3 *Q-relaxed intersection*

The q-relaxed intersection is a set membership operator. Given m sets $\mathbb{X}_1, \mathbb{X}_2 \ldots, \mathbb{X}_m$, their q-relaxed intersection, noted $\cap^q \mathbb{X}_i$ is the set of all $x \in \mathbb{R}^n$ that belongs to the $\mathbb{X}_i's$ except q at most. This operator is deeply covered in [15, 20, 62].

A.7 CONCLUSION

In this Appendix, we briefly reviewed the interval set-membership tools that are commonly used for parameter estimation and that have been used in this thesis. For a complete introduction on the topic, the reader should read [41].

B

THE LMI CONTRACTOR

B.1 INTRODUCTION

Linear Matrix Inequalities (LMI for short) is a large class of convex constraints. Boxes, Ellipsoids, linear constraints, can be represented by LMIs. Intersections of LMIs are also LMIs. Interior point methods are able to minimize or maximize any criterion of LMIs with a complexity which is polynomial regarding to the number of variables. As a consequence, as shown in this appendix, it is possible to build optimal contractors for sets represented by LMIs. When solving a set of nonlinear constraints, one may extract from all constraints those that are LMIs in order to build a single optimal LMI contractor for a subset of the original CSP. A combination with all contractors obtained for other non LMI constraints can thus be performed up to the fixed point. The resulting propagation is shown to be more efficient than other conventional contractor-based approaches.

Research on this contractor has been motivated in Chapter 3: in the general case, the Kalman contractor consists of contracting the state vectors at two instants k, l with respect to some ellipsoid. An ellipsoid is described by quadratic constraints potentially containing crossed terms, thus shows multi-occurrences causing classical constraint propagation methods to be non-optimal [40].

Using well-known convex optimization methods, which have already been used to contract optimally interval matrices for the semi-definite positive constraint [39], this appendix will introduce a new optimal contractor for the LMI constraint for which, to the best of my knowledge, no optimal algorithms have been presented yet. This work has been published in [69].

This appendix is organized as follows. We shall first recall what an LMI is and present a few problems that can be represented as LMIs. In the third section we will present the problem of convex optimization under LMI constraints, and see how it can be used in the context of Interval Analysis. In the fourth section we will make a brief recall on Interval Constraints Propagation and contractor programming, more deeply covered in Appendix A, and show how to leverage optimization under LMI constraints methods to create a minimal contractor for the LMI constraint. In Section B.5 we will present some examples of application, compare the efficiency of our contractor with existing ones, and illustrate the power of contractor programming by mixing convex and non-convex constraints.

B.2 LINEAR MATRIX INEQUALITIES

In this section, we will recall some properties about LMIs. For an in depth review on the topic, see [12] and [13].

A Linear Matrix Inequality (LMI) has the form:

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + x_1\boldsymbol{F}_1 + x_2\boldsymbol{F}_2 + \cdots + x_n\boldsymbol{F}_n \succeq 0 \tag{B.1}$$

with $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ the unknown, $\boldsymbol{F}_{i=0\ldots,n} \in \mathbb{R}^{n\times n}$ the set of symmetric matrices.

The inequality $\succeq$ means that $\boldsymbol{F}(\boldsymbol{x})$ is a positive definite matrix, i.e.:

$$\forall \boldsymbol{u} \in \mathbb{R}^n,\ \boldsymbol{u}^T \cdot \boldsymbol{F} \cdot \boldsymbol{u} > 0,\ \boldsymbol{u} \neq \boldsymbol{0}. \tag{B.2}$$

Since the set

$$\mathbb{S} = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{F}(\boldsymbol{x}) \succeq 0\} \tag{B.3}$$

is convex, an LMI is a convex constraint on $\boldsymbol{x}$.

A SYSTEM DESCRIBED BY SEVERAL LMIS CAN BE DESCRIBED BY A SINGLE LMI. The set $\mathbb{S}_{12} = \mathbb{S}_1 \cap \mathbb{S}_2$ with $\mathbb{S}_1, \mathbb{S}_2$ two convex sets described by the LMIs $\boldsymbol{S}_1(\boldsymbol{x}) \succeq \boldsymbol{0}, \boldsymbol{S}_2(\boldsymbol{x}) \succeq \boldsymbol{0}$ is convex and is also described as an LMI $\boldsymbol{S}_{12}(\boldsymbol{x}) \succeq \boldsymbol{0}$, obtained by concatenation of $\boldsymbol{S}_1(\boldsymbol{x})$ and $\boldsymbol{S}_2(\boldsymbol{x})$:

$$\boldsymbol{S}_{12}(\boldsymbol{x}) = \operatorname{diag}(\boldsymbol{S}_1(\boldsymbol{x}), \boldsymbol{S}_2(\boldsymbol{x})) = \begin{pmatrix} \boldsymbol{S}_1(\boldsymbol{x}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_2(\boldsymbol{x}) \end{pmatrix} \succeq \boldsymbol{0}. \tag{B.4}$$

A SET OF LINEAR INEQUALITIES IS AN LMI. For example:

$$\begin{cases} I_1(\boldsymbol{x}) = & a_{11}x_1 + a_{12}x_2 + b_1 \geq 0 \\ I_2(\boldsymbol{x}) = & a_{21}x_1 + a_{22}x_2 + b_2 \geq 0 \end{cases} \tag{B.5}$$

is equivalent to the following LMI:

$$\begin{pmatrix} I_1(\boldsymbol{x}) & 0 \\ 0 & I_2(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + b_1 & 0 \\ 0 & a_{21}x_1 + a_{22}x_2 + b_2 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.6}$$

i.e.,

$$\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} + x_1 \begin{pmatrix} a_{11} & 0 \\ 0 & a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} & 0 \\ 0 & a_{22} \end{pmatrix} \succeq \boldsymbol{0}. \tag{B.7}$$

A BOX IS DESCRIBED BY AN LMI. $\boldsymbol{x} \in [\boldsymbol{x}]$, with $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $[\boldsymbol{x}] = [\underline{x_1}, \overline{x_1}] \times [\underline{x_2}, \overline{x_2}] \times \cdots \times [\underline{x_n}, \overline{x_n}] \in \mathbb{IR}^n$ is the set of all the intervals of $\mathbb{R}^n$, is an LMI. Indeed, it can be decomposed as a set of $2 \times n$ linear inequalities:

$$\begin{cases} \boldsymbol{I}_1^- (\boldsymbol{x}) = & -\underline{x}_1 + x_1 \geq 0 \\ \boldsymbol{I}_1^+ (\boldsymbol{x}) = & \overline{x}_1 - x_1 \geq 0 \\ \boldsymbol{I}_2^- (\boldsymbol{x}) = & -\underline{x}_2 + x_2 \geq 0 \\ \boldsymbol{I}_2^+ (\boldsymbol{x}) = & \overline{x}_2 - x_2 \geq 0 \\ & \vdots \\ \boldsymbol{I}_n^- (\boldsymbol{x}) = & -\underline{x}_n + x_n \geq 0 \\ \boldsymbol{I}_n^+ (\boldsymbol{x}) = & \overline{x}_n - x_n \geq 0 \end{cases} \tag{B.8}$$

which in turn can be formulated as an LMI. For example, the box:

$$[\boldsymbol{x}] = [-10, 15] \times [3, 7] \tag{B.9}$$

is an LMI and is written as:

$$\boldsymbol{B} (\boldsymbol{x}) = \begin{pmatrix} x_1 + 10 & 0 & 0 & 0 \\ 0 & 15 - x_1 & 0 & 0 \\ 0 & 0 & x_2 - 3 & 0 \\ 0 & 0 & 0 & 7 - x_2 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.10}$$

i.e.,

$$\boldsymbol{B} (\boldsymbol{x}) = \begin{pmatrix} -\underline{x}_1 & 0 & 0 & 0 \\ 0 & \overline{x}_1 & 0 & 0 \\ 0 & 0 & -\underline{x}_2 & 0 \\ 0 & 0 & 0 & \overline{x}_2 \end{pmatrix} + x_1 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.11}$$

AN ELLIPSOID IN $\mathbb{R}^d$ IS DESCRIBED BY AN LMI. An ellipsoid $\mathbb{E} \in \mathbb{R}^d$ is described as:

$$\mathbb{E} = \left\{ \boldsymbol{x} \in \mathbb{R}^d | r - (\boldsymbol{x} - \boldsymbol{c})^T \boldsymbol{P}^{-1} (\boldsymbol{x} - \boldsymbol{c}) \geq 0 \right\} \tag{B.12}$$

with $r > 0$ its radius, $\boldsymbol{c} \in \mathbb{R}^d$ its center and $\boldsymbol{P}$ its characteristic matrix.

Using Schur complement theorem, which states that for a set of matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ respectively of dimensions $p \times p, p \times q, q \times p, q \times q$,

$$\begin{cases} \boldsymbol{A} & \succ \boldsymbol{0} \\ \boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \succeq \boldsymbol{0} \end{cases} \Leftrightarrow \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \succeq \boldsymbol{0} \tag{B.13}$$

we show that $\mathbb{E}$ can be described as the LMI:

$$E (x) = \begin{pmatrix} r & \boldsymbol{x} - \boldsymbol{c}^T \\ \boldsymbol{x} - \boldsymbol{c} & \boldsymbol{P} \end{pmatrix} \succeq \boldsymbol{0}. \tag{B.14}$$

For example, the unit disk $\mathbb{D} \in \mathbb{R}^2$ is an ellipsoid with a characteristic matrix $\boldsymbol{P} = \boldsymbol{Id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, a radius $r = 1$, centered on the origin $\boldsymbol{c} = (0,0)^T$:

$$\mathbb{D} = \left\{ \boldsymbol{x} \in \mathbb{R}^2 | 1 - \boldsymbol{x}^T \boldsymbol{Id} \boldsymbol{x} \geq 0 \right\} \tag{B.15}$$

i.e.,

$$1 - x_1^2 - x_2^2 \geq 0. \tag{B.16}$$

Using the Schur complement theorem with

$$\boldsymbol{A} = 1, \boldsymbol{B} = \boldsymbol{C}^T = (x_1 x_2), \boldsymbol{D} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{B.17}$$

we get the LMI:

$$\begin{pmatrix} 0 & x_1 & x_2 \\ x_1 & 1 & 0 \\ x_2 & 0 & 1 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.18}$$

i.e.,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + x_1 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.19}$$

Many other convex sets can be represented by LMIs. See [13][6] for a presentation of a consequent amount of LMI-representable sets.

B.3 OPTIMIZATION UNDER LMI CONSTRAINTS

Given a vector of variable $\boldsymbol{x} \in \mathbb{R}^n$, a cost vector $\boldsymbol{c} \in \mathbb{R}^n$ and a matrix $\boldsymbol{F}$, an optimization problem under the LMI constraint $\boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{0}$ is stated as follows:

$$P_{\boldsymbol{c},\boldsymbol{F}(\boldsymbol{x})} = \begin{cases} \text{minimize :} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{under the constraint :} & \boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{0} \end{cases}. \tag{B.20}$$

Many engineering problems, for example in the field of control theory [**Arzelier02multi-objectiveh2/h=00221E/impulse-to-peak**], can be framed as an optimization problem under LMI constraints. The book [13] presents a consequent amount of problem that can be formulated as optimization problems under LMI constraints. Only trivial cases of optimization under LMI constraints can be solved analytically. On the other hand, numerical methods have been developed to solve this type of problem efficiently.

Among such methods, the most efficient are based on interior point methods. Interior point methods were introduced in 1984 by Karmarkar for solving linear programming problems with a polynomial-time algorithm [53]. A lot of research activities followed, leading to [66], where general interior-point methods solving LMI problems are introduced. We refer the reader to [12] for a more detailed review on the algorithms for solving LMI optimization problems and their implementation.

Unlike simplex method commonly used in linear programming, which moves along the edges of the polytope defining the feasible set until it reaches the vertex of optimum solution, interior point methods start from a feasible solution and iterates inside this set until the optimal solution is found.

COMPUTING THE BOX-HULL OF A SET IN $\mathbb{R}^d$ DESCRIBED BY AN LMI REDUCES TO SOLVING $2d$ OPTIMIZATION PROBLEMS UNDER LMI CONSTRAINTS. The box-hull, or the bounding-box $[\mathbb{X}]$ of a set $\mathbb{X}$ in $\mathbb{R}^d$ is the smallest axis-aligned box enclosing $\mathbb{X}$.

Indeed, for each dimension d, we are looking for the maximal and minimal values $\overline{x}_d, \underline{x}_d$ of x_d, giving us the interval $[x_d] = [\underline{x}_d, \overline{x}_d]$, their Cartesian product $[\boldsymbol{x}] = [x_1] \times [x_2] \times \cdots \times [x_d]$ being the box-hull of $\mathbb{X}$.

Since finding these $2 \times d$ extremas reduces to solving an optimization problem under LMI constraints for each of them, each minimization problem having a polynomial-time complexity, the box-hull problem also has a polynomial-time complexity.

ILLUSTRATION. Given two sets of $\mathbb{R}^2$: an ellipsoid $\mathbb{S}_1$ and a triangle $\mathbb{S}_2$ described by the LMIs $\boldsymbol{S}_1(\boldsymbol{x}) \succeq \boldsymbol{0}$, $\boldsymbol{S}_2(\boldsymbol{x}) \succeq \boldsymbol{0}$, we want to find the box-hull $[\mathbb{S}_\cap]$ for their intersection $\mathbb{S}_\cap = \mathbb{S}_1 \cap \mathbb{S}_2$, described by the LMI $\boldsymbol{S}(\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{S}_1(\boldsymbol{x}) & 0 \\ 0 & \boldsymbol{S}_2(\boldsymbol{x}) \end{pmatrix} \succeq \boldsymbol{0}$. $\mathbb{S}_\cap$, $\mathbb{S}_1$ and $\mathbb{S}_2$ are illustrated on Figure B.1.

First, we are looking for an upper-bound $\overline{x}_1$ for x_1. Therefore we solve the optimization problem:

$$P_{\boldsymbol{c},\boldsymbol{S}(\boldsymbol{x})} = \begin{cases} \text{minimize:} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{under the constraint:} & \boldsymbol{S}(\boldsymbol{x}) \succeq \boldsymbol{0} \end{cases} \tag{B.21}$$

with $\boldsymbol{c}^T = (-1, 0)$. The dashed line (a) illustrates the solution found by the algorithm while searching for $\overline{x_1}$. We are then looking for the lower-bound $\underline{x}_1$ for x_1, which is done by solving $P_{\boldsymbol{c},\boldsymbol{S}(\boldsymbol{x})}$ with $\boldsymbol{c}^T = (1, 0)$. The solution is the dashed-line (b). The same process is then repeated for x_2 as illustrated on the dashed lines (c), (d).

The box-hull $[\mathbb{S}]$ is then given by the Cartesian product $[x_1] \times [x_2]$ of the intervals $[x_1] = [\underline{x}_1, \overline{x}_1]$, $[x_2] = [\underline{x}_2, \overline{x}_2]$.

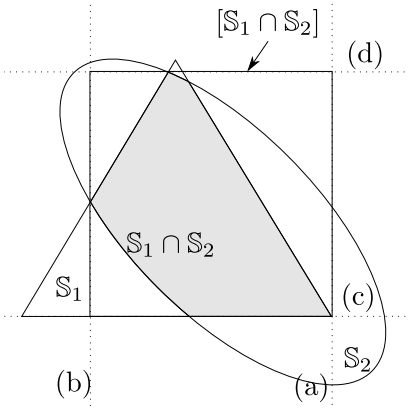


Figure B.1: Characterization of $[\mathbb{S}_1 \cap \mathbb{S}_2]$(a) maximization of x_1, (b) minimization of x_1, (c) minimization of x_2, (d) maximization of x_2

B.4 OPTIMAL CONTRACTOR UNDER LMI CONSTRAINTS

The basic definitions and properties of contractors have been given in Chapter A. We simply define here the optimal contractor for the LMI constraint. Let $\mathbb{S}$ be a convex set described by the LMI $\boldsymbol{S}(\boldsymbol{x}) \succ \boldsymbol{0}$ and $[\boldsymbol{x}] \in \mathbb{IR}^n$ a box described by the LMI $\boldsymbol{B}(\boldsymbol{x}) \succ \boldsymbol{0}$. The operator $\boldsymbol{C}_{LMI}$, which maps $[\boldsymbol{x}]$ to the box-hull of the set $\mathbb{S}^{\cap}$ described by $\begin{pmatrix} \boldsymbol{B}(\boldsymbol{x}) & 0 \\ 0 & \boldsymbol{S}(\boldsymbol{x}) \end{pmatrix} \succeq \boldsymbol{0}$ is a contractor for $\mathbb{S}$, and it is minimal.

B.5 EXAMPLES

Using the C++ library SDPA [33], we implemented the LMI contractor for the IBEX library [3, 16, 72]. Given a set of matrices $\boldsymbol{F}_0, \boldsymbol{F}_1, \ldots \boldsymbol{F}_n$ and a box $[\boldsymbol{x}] \in \mathbb{IR}^n$ as inputs, it contracts $[\boldsymbol{x}]$ with respect to the LMI constraint:

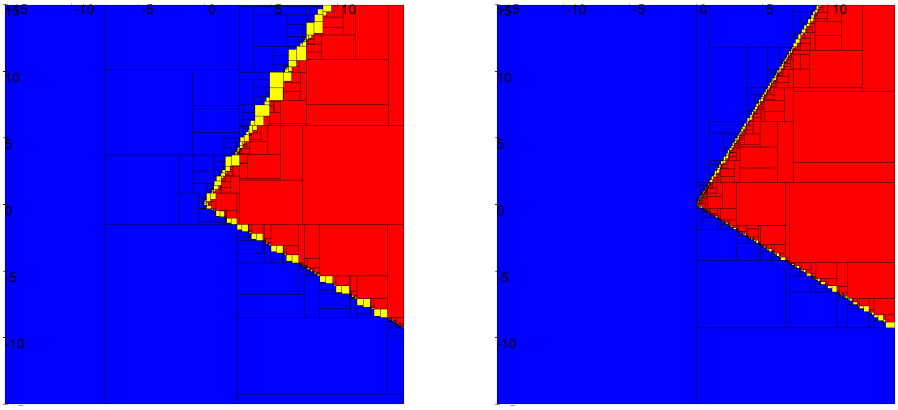
$$\boldsymbol{C}_{LMI}(\boldsymbol{x}) = \text{diag}\left(\boldsymbol{B}(\boldsymbol{x}), \boldsymbol{F}_0 + \sum_{i=1}^{n} \boldsymbol{F}_i(\boldsymbol{x}_i)\right) \succeq \boldsymbol{0} \tag{B.22}$$

where $\boldsymbol{B}(\boldsymbol{x}) \succeq \boldsymbol{0}$ is the LMI constraint for $\boldsymbol{x} \in [\boldsymbol{x}]$ as presented in Section B.2. All the illustrations shown here were made using the VIBEs [26] drawer and its C++ API.

B.5.1 *Representation of a simple LMI-set*

Consider the following LMI:

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 + x_2 \end{pmatrix} \succeq \boldsymbol{0} \tag{B.23}$$



(a) (b)

Figure B.2: Sub-paving of $\mathbb{F}$ (a) with forward-backward contractors only, (b) with a forward-backward and a LMI contractor

which means that the two eigenvalues λ_1, λ_2 of $\boldsymbol{F}$ are real positive, i.e.:

$$\begin{array}{rl} \lambda_1 = & \frac{2x_1+x_2+\sqrt{x_2^2+4x_2^2}}{2} \geq 0 \\ \lambda_2 = & \frac{2x_1+x_2-\sqrt{x_2^2+4x_2^2}}{2} \geq 0 \end{array} . \tag{B.24}$$

Using set-inversion and contractor programming methods, we want to approximate this LMI-set $\mathbb{F}$ with a sub-paving. For this we need two contractors:

$$\boldsymbol{C}^{in}_{\uparrow\downarrow} = \boldsymbol{C}^{in}_{\lambda_1\uparrow\downarrow} \cup \boldsymbol{C}^{in}_{\lambda_2\uparrow\downarrow} \tag{B.25}$$

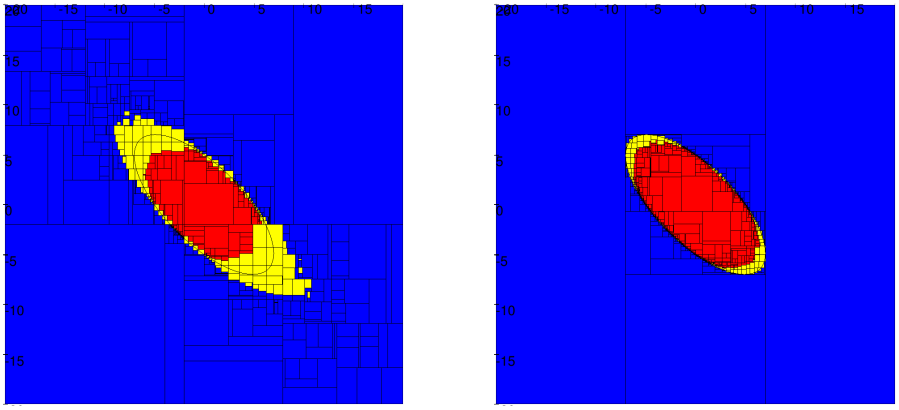
which sill remove part of the search space that are consistent with the constraints $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ and

$$\boldsymbol{C}^{out}_{\uparrow\downarrow} = \boldsymbol{C}^{out}_{\lambda_1\uparrow\downarrow} \cap \boldsymbol{C}^{out}_{\lambda_2\uparrow\downarrow} \tag{B.26}$$

which will remove part of the search space that are inconsistent with the constraints $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. The boxes removed by $\boldsymbol{C}^{in}$ are drawn in red, the boxes removed by $\boldsymbol{C}^{out}$ are drawn in blue while the ones undetermined will be drawn in yellow.

Since the constraint $\boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{0}$ is an LMI constraint, we can replace the outer contractor $\boldsymbol{C}^{out}_{\uparrow\downarrow}$ with our LMI contractor $\boldsymbol{C}_{LMI}$ introduced in Section B.4. On the other hand, the constraints defining $\boldsymbol{C}^{out}$ are non-convex, and a forward-backward approach is appropriated.

Figure B.2 compares the two approaches, with the same allowed number of bisections. As can be seen, the sub-paving computed by the combination of the forward-backward and the LMI contractor is much finer than the one computed with the forward-backward approach only. This is due to the fact that since $\boldsymbol{C}_{LMI}$ is minimal, less time is spent bisecting parts of the search space that do not satisfy the positive definite constraint for sure, and more bisection can be performed elsewhere.



(a) (b)

Figure B.3: Characterization of $\mathbb{E}$ (a) with forward-backward contractions only (b) with forward-backward contractor and the LMI contractor.

B.5.2 *Characterization of ellipsoidal sets*

In this example we want to characterize the set

$$\mathbb{E} = \left\{ \boldsymbol{x} \in \mathbb{R}^2 | \boldsymbol{x}^T \boldsymbol{P}^{-1} \boldsymbol{x} \leq \boldsymbol{r} \right\} \tag{B.27}$$

which is an ellipsoid with $\boldsymbol{P} = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}$, $r = 5$. As in the previous example, we want to approximate $\mathbb{E}$ with a sub-paving. Since it can be described as an LMI-set, we compare the results obtained with the LMI contractor and forward-backward contractor for the outer contractor. The sub-pavings generated are shown in Figure B.3. Again, thanks to the minimality of the LMI contractor, we see that the sub-paving generated in (b) is much more precise than the one generated in (a) using forward-backward contractors only.

B.5.3 *Manipulating the LMI contractor using contractor algebra*

In this example we have two sets: a triangle

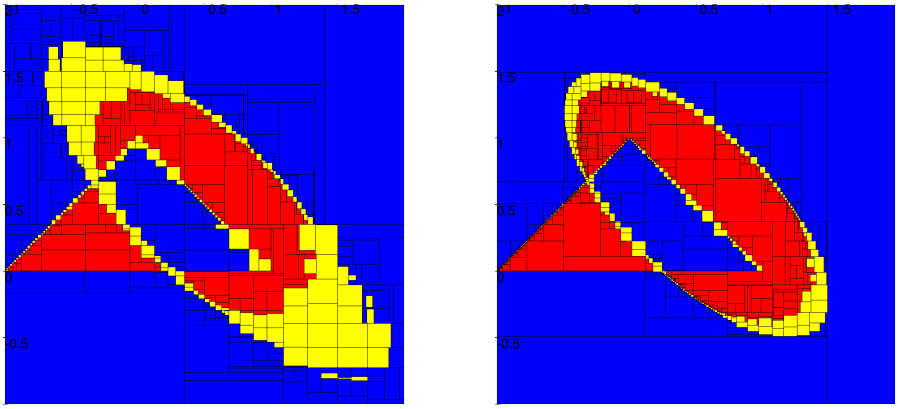
$$\mathbb{T} = \left\{ \boldsymbol{x} \in \mathbb{R}^2 | x_2 + x_1 - 1 \geq 0, x_2 - x_1 - 1 \geq 0, x_2 \geq 0 \right\} \tag{B.28}$$

and an ellipse:

$$\mathbb{E} = \left\{ \boldsymbol{x} \in \mathbb{R}^2 | \left(\boldsymbol{c} - \boldsymbol{x} \right)^T \boldsymbol{P}^{-1} \left(\boldsymbol{x} - \boldsymbol{c} \right) \leq r \right\} \tag{B.29}$$

with $\boldsymbol{P} = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}$, $r = 0.5$, $\boldsymbol{c} = (0.5, 0.5)^T$.

We would like to compute a sub-paving for the set $\mathbb{S} = (\mathbb{T} \cup \mathbb{E}) \setminus (\mathbb{T} \cap \mathbb{E})$.



(a) (b)

Figure B.4: Characterization of $\mathbb{S}$ (a) with forward-backward contractions only (b) with forward-backward contractors and LMI contractors

For this we need the inner and outer contractors $\mathcal{C}_{\mathbb{T}}^{in}, \mathcal{C}_{\mathbb{T}}^{out}$ for $\mathbb{T}$ and the inner and outer contractors $\mathcal{C}_{\mathbb{E}}^{in}, \mathcal{C}_{\mathbb{E}}^{out}$ for $\mathbb{E}$.

Using contractor algebra, we obtain the outer and inner contractors for $\mathbb{S}$:

$$\begin{aligned} \mathcal{C}_{\mathbb{S}}^{out} &= \left(\mathcal{C}_{\mathbb{E}}^{in} \cap \mathcal{C}_{\mathbb{T}}^{out}\right) \cup \left(\mathcal{C}_{\mathbb{E}}^{out} \cap \mathcal{C}_{\mathbb{T}}^{in}\right) \\ \mathcal{C}_{\mathbb{S}}^{in} &= \left(\mathcal{C}_{\mathbb{E}}^{out} \cup \mathcal{C}_{\mathbb{T}}^{in}\right) \cap \left(\mathcal{C}_{\mathbb{E}}^{in} \cup \mathcal{C}_{\mathbb{T}}^{out}\right) \end{aligned} . \tag{B.30}$$

In Figure B.4 we compare the sub-pavings obtained when $\mathcal{C}_{\mathbb{E}}^{out}$ and $\mathcal{C}_{\mathbb{T}}^{out}$ are forward-backward contractors (a) and when they are LMI contractors. Again, we observe the gain in precision when using the LMI-based contractor.

B.6 CONCLUSION AND OUTLOOK

In this appendix, we introduced a new contractor based on convex optimization under LMI constraints. To the best of our knowledge, this kind of constraint is under-exploited in the interval analysis community and has never been implemented in the contractor framework before.

LMI constraints are omnipresent in the context of robotics. For example, range or pseudo-range measurements can be framed as LMI constraints [10, 24, 37], and therefore the use of our contractor could be used in this kind of application.

Furthermore, in many Bayesian estimation techniques, the parameters to be estimated are modeled as Gaussian random variables, for which the confidence domain is an ellipsoid, which can be represented by an LMI. Such is the case for the inter-temporal distribution presented in Chapter 3, whose confidence domain is used to define the Kalman contractor. Therefore, we anticipate that our contractor

could be used to combine probabilistic and non-probabilistic estimations methods in some other possibly interesting ways.

C

GUARANTEED MAXIMUM LIKELIHOOD ESTIMATION

C.1 INTRODUCTION

In our context, detailed in Section 2, as in many operational application, a set of feasible solution guaranteed to be consistent with the definition of the problem is a valuable piece of information, but it does not suffice. Even if interval analysis set-membership based methods can accommodate arbitrary parameter distribution shapes, even multi-modal ones, for practical application (e.g.: sizing of a pipe, computation by a control algorithm of the next control input for an autonomous vehicle based on its current state...), a punctual estimation is often required.

Now there exist many ways to extract a punctual estimation from a set. In Section C.2 we present and discuss the most frequently used methods. As we will see, they are mostly based on Geometrical considerations. As we deal with probabilistic errors, the classical punctual extraction methods discard a lot of statistically relevant information, and therefore are often not precise.

Based on probabilistic considerations, we review in Section C.3 a reliable version of the linearizing Maximum Likelihood Estimator discussed in Section 4.3 that doesn't approximate the model. In Section C.3 we extend that principle to accommodate for faulty measurements to get a reliable and robust Maximum Likelihood Estimator based on the methods developed in Chapter 4.

C.2 CLASSICAL PUNCTUAL ESTIMATORS

CENTROID A widely used estimator that returns a point from a set is the centroid. It is defined as follows.

Definition 42. Let $\mathbb{S} \in \mathbb{R}^n, \mathbf{x} \in \mathbb{R}^n$, and $g(\mathbf{x})$ the characteristic function of $\mathbb{S}$, i.e.:

$$g(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathbb{S} \\ 0 & \text{otherwise} \end{cases}. \tag{C.1}$$

The centroid $\mathbf{x}_c$ of $\mathbb{S}$ is defined by the integral:

$$\mathbf{x}_c = \frac{\int_{\mathbb{R}^n} \mathbf{x} \cdot g(\mathbf{x}) \cdot d\mathbf{x}}{\int_{\mathbb{R}^n} g(\mathbf{x}) \cdot d\mathbf{x}}. \tag{C.2}$$

The centroid is a good feature in the framework of Interval methods: it can be estimated recursively, which is convenient for paving

algorithms, and well-suited for real-time applications where the algorithm may be interrupted as in [27]. It corresponds to the arithmetic mean position of all the points inside $\mathbb{S}$.

From a probabilistic point of view, we can view $\mathbf{x}$ as a uniformly distributed random variable defined over $\mathbb{S}$. Int that case, $\mathbf{x}_c$ is simply the Expected Value of $\mathbf{x}$. There is, however, no justification to assume that $\mathbf{x}$ is uniformly distributed over $\mathbb{S}$. A common interpretation of interval methods is to view the bounding of a random variable $\mathbf{r}$ by the box $[\mathbf{r}]$ as a formulation of the hypothesis that $\mathbf{r}$ is uniformly distributed over $[\mathbf{r}]$. Now even if that hypothesis was acceptable, it still wouldn't justify to view our parameter as uniformly distributed over a set of confidence η.

Indeed, let's assume that in our scenario we do have $\mathbf{e} \sim \mathbf{I}_{\mathbb{E}}$. Then it is generally not correct to assume that $\mathbf{p} = \mathbf{f}^{-1}(\mathbf{e})$ is uniformly distributed over $\hat{\mathbb{P}}_{\mathbb{E}} = \mathbf{f}^{-1}(\mathbf{I}_{\mathbb{E}})$, as the image of an uniformly distributed random variable through a nonlinear function is generally not uniformly distributed. Finding which distribution that random variable follows is hard in the general case, and that difficulty is the one that justifies the use of nonlinear methods.

Therefore, even if this feature is convenient, it is not statistically meaningful.

CHEBYCHEV CENTER The Chebychev center of a set $\mathbb{S}$ is the center of the minimal-radius ball enclosing the entire set. In the framework of parameter estimation, the Chebychev center minimizes the worst feasible estimation error for the true parameter vector given $\mathbb{S}$, the set of all feasible parameters. Not that it can also be defined as the center of the maximal-radius ball enclosed by the set, but in that case it may not be unique and it is ill-defined when $\mathbb{S}$ is not connex.

CENTER OF THE BOX-HULL This representation simply computes the center of the minimum-width box enclosing $\mathbb{S}$. It is easy to compute with a complexity linear with respect to the number of boxed that makes the sub-paving. However, this is not a precise representation: when the main axis of $\mathbb{S}$ are not axis-parallel, $[\mathbb{S}]$ is a gross approximation of $\mathbb{S}$ and may lead to great pessimism.

CENTER OF A SMALLEST ENCLOSING ELLIPSE This feature is an extension of the Chebychev center for the case of Ellipsoids. The criterion according to what the ellipse will be minimized is fundamental. Two examples of criterion commonly used are the trace and the volume of the ellipsoid. In the first case, we are computing the ellipsoid whose shape matrix's trace is minimized, and in the second case we are computing the ellipse whose volume (proportional to the determinant of its shape matrix) is minimized. This approach has the advantage of better capturing the shape of $\mathbb{S}$ than using its box-hull,

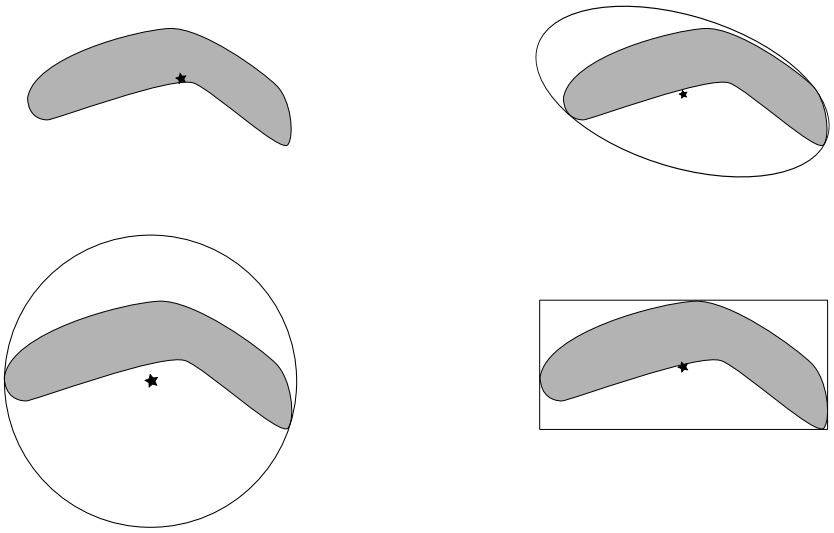


Figure C.1: On this figure, we show several ways to get a punctual estimation (star shape) $\mathbf{x}$ from a set $\mathbb{X}$: (a) Centroid of $\mathbb{X}$, (b) : Center of the smallest volume ellipse $\mathbb{E}_{vol}$ enclosing $\mathbb{X}$, (c) : Chebyshev center of $\mathbb{X}$, (d) : Center of $[\mathbb{X}]$, the box-hull of $\mathbb{X}$.

but has the drawbacks that it may be computationally intensive, and the choice of whether minimizing the trace or the determinant of its shape matrix will greatly influence the results.

Figure C.1 illustrates the centroid, center of the smallest volume enclosing ellipse, Chebychev center and center of the box-hull for an arbitrary set $\mathbb{S}$. Of course there are many other features that are good candidates for providing a punctual estimation, such as the geometric median as discussed in [27]. However, all these methods have the drawback that they are based solely on geometrical considerations, and discard the information we have about the probabilistic nature of the uncertainties. Worse, they can lead to situations where the punctual estimation is known to not belong to the set of feasible solutions, as is illustrated in Figure C.1, which is contradictory with the reliability aspect of the estimator. Also, our methods can naturally estimate multi-modal parameter distributions as illustrated in Figure C.2, but when such situation occurs, how do which chose which mode is more likely to contain the correct one to make a decision?

Classical probabilistic approaches produce a punctual estimate that is *best* according to some criterion. As defined in Section 4.3, a punctual estimation in the sense of Maximum Likelihood estimation will be the parameter vector that minimizes the likelihood function

$$L\left(y_i|\mathbf{p}\right) = \prod_i \pi\left(y_i|\mathbf{p}\right) \tag{C.3}$$

and reduces to a nonlinear least-square parameter estimation when the y_i's are independent, normally distributed as recalled in Section 4.3..

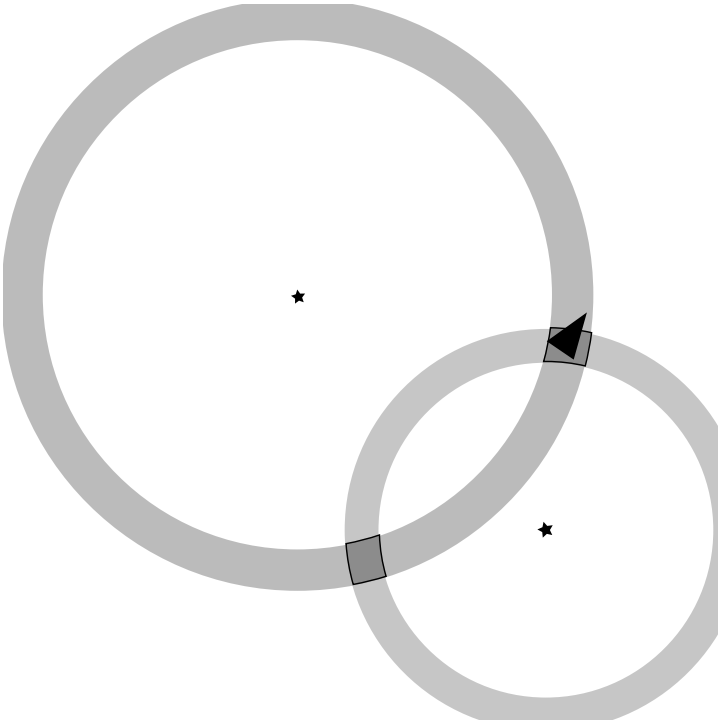


Figure C.2: In case of symmetrical or ill-defined problems, the parameter set may be non connected, as is the case for the underconstrained global localization problem. When this situation occurs, which set should be chosen to make a decision based on the robot position? On this figure, we illustrate this problem for localizing a robot in two dimensions with only two landmarks.

Figure C.3: To get a statistically meaningful, close to punctual estimation, one might be tempted to compute $\mathbb{P}_\eta$ with the smallest value of η such that $\mathbb{P}_\eta \neq \emptyset$. From darkest to clearest: $\mathbb{P}_{\eta_1}, \mathbb{P}_{\eta_2}, \mathbb{P}_{\eta_3}, \mathbb{P}_{\eta_4}, \mathbb{P}_{\eta_5}$ with $\eta_1 > \eta_2 > \eta_3 > \eta_4 > \eta_5$.

Interval methods are perfectly able to solve global, nonlinear optimization problems, so a first step will be to simply translate the Maximum Likelihood Estimation problem as a CSP to be solved by an interval solver. We will remark that this approach cannot easily be made robust to the presence of outliers.

Another, robust approach will be proposed that simply computes $\hat{\mathbb{P}}_{\mathbb{B}^q_{\bar{\eta}}}$, the probabilistic parameter set obtained from the q-relaxed error box with a confidence $\bar{\eta}$ described in Chapter 4, with $\bar{\eta}$ the smallest value of η such that $\hat{\mathbb{P}}_{\mathbb{B}^q_\eta}$ is not empty. Indeed, for any bounding $\mathbb{E}_\eta$ containing the error with a probability η, $\mathbb{P}_{\mathbb{E}_\eta}$ is a level-set of the probability distribution function for the parameter vector $\mathbf{p}$ (i.e.: the set of all parameters whose likelihood is greater than some threshold), for which we don't have an expression. Computing $\hat{\mathbb{P}}_{\mathbb{E}_{\hat{\eta}}}$ therefore amounts to computing $\hat{\mathbf{p}}$, the parameter vector that maximizes the pdf of $\mathbf{p}$, which is nothing more than the MLE for $\mathbf{p}$, with $\hat{\eta}$ its likelihood. Choosing $\mathbb{B}^q_\eta$ as a bounding allows to tolerate faulty measurements, and we will see it is a special application of the OMNE and GOMNE estimators proposed in [42, 59]. Figure C.3 shows an illustration of the principle of computing smaller and smaller level-sets enclosing the Maximum Likelihood Estimator until some precision is reached.

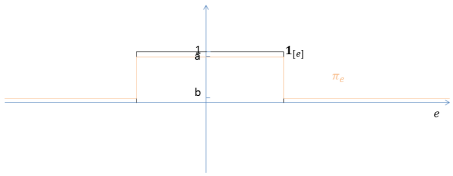


Figure C.4: Illustration of $\pi(e)$ and $\mathbf{1}(e)$

C.3 OMNE AS A MAXIMUM LIKELIHOOD ESTIMATOR

OMNE OMNE (Outlier Minimal Number Estimator) introduced in [60] is a well known bounded-error estimator which is robust with respect to outliers. It returns the set of parameter vectors that are consistent with the maximal number of data bar. OMNE proposes to minimize the number of outliers or equivalently, to maximize the number of inliers. The relaxed intersection made it possible to implement OMNE in a reliable way for nonlinear models [46]. OMNE has been used in several applications such as the localization of robots [23], and has received some probabilistic interpretations [25].

Consider the static parameter estimation mode:

$$\boldsymbol{y} = \boldsymbol{\psi}(\boldsymbol{p}) + \boldsymbol{e} \tag{C.4}$$

where $\boldsymbol{\psi} : \mathbb{R}^n \to \mathbb{R}^m$ is the model, $\boldsymbol{p} \in \mathbb{R}^n$ is the parameter vector to be estimated, $\boldsymbol{y} \in \mathbb{R}^m$ is the measurement vector and $\boldsymbol{e}$ the noise. Define the function $\boldsymbol{f}(\boldsymbol{p}) = \boldsymbol{e} = \boldsymbol{y} - \boldsymbol{\psi}(\boldsymbol{e})$. Given an error interval $[e] \subset \mathbb{R}$ that is supposed to contain the error e_i if the corresponding data y_i is an inlier. In practice, $[e]$ is a small interval which contains 0. The OMNE estimator returns the set of all $\boldsymbol{p}$ such as the property $f_i(\boldsymbol{p}) \in [e]$ is satisfied for a maximal number of data. More precisely, OMNE returns the set

$$\hat{\mathbb{P}} = argmax_{\boldsymbol{p}\in\mathbb{R}^n} \sum_i \mathbf{1}_{[e]}(f_i(\boldsymbol{p})) \tag{C.5}$$

where $\mathbf{1}_{\mathbb{X}}$ denotes the characteristic function of the set $\mathbb{X}$ or equivalently, $\mathbf{1}_{[e]}(f_i(\boldsymbol{p})) = 1$ if and only if $f_i(\boldsymbol{p}) \in [e]$ and $\mathbf{1}_{[e]}(f_i(\boldsymbol{p})) = 0$ otherwise, as illustrated on Figure C.4.

Theorem 43 shows that OMNE corresponds to a maximum likelihood estimator.

Theorem 43. *Assume that the error vector* $\mathbf{e} = (e_1, \ldots, e_n)$ *is white,* i.e., *all* e_i *are independent and identically distributed with the probability density function* π_e *which is half uniform. More precisely,* $\pi_e(e_i) = a$ *if* $e_i \in [e]$ *and* $\pi_e(e_i) = b < a$ *otherwise. Then the maximum likelihood estimator corresponds to OMNE.*

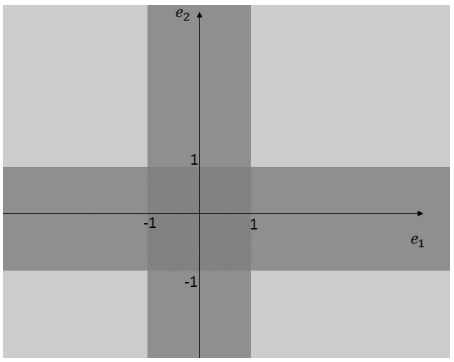


Figure C.5: Illustration of a two-dimensional independent half-uniform random variable vector

Proof. The likelihood is defined by

$$\begin{aligned} \pi(\mathbf{y} \mid \mathbf{p}) &= \prod_i \pi_e(y_i - \psi_i(\mathbf{p})) = \prod_i \pi_e(f_i(\mathbf{p})) \\ &= \prod \left(a \cdot \mathbf{1}_{[e]}(f_i(\mathbf{p})) + b \cdot \mathbf{1}_{\mathbb{R}\setminus[e]}(f_i(\mathbf{p})) \right). \end{aligned} \tag{C.6}$$

It is maximal if the log-likelihood

$$\begin{aligned} \log \pi(\mathbf{y} \mid \mathbf{p}) &= \sum_i \log \left(a \cdot \mathbf{1}_{[e]}(f_i(\mathbf{p})) + b \cdot \mathbf{1}_{\mathbb{R}\setminus[e]}(f_i(\mathbf{p})) \right) \\ &= \sum_i \log \left((a-b) \cdot \mathbf{1}_{[e]}(f_i(\mathbf{p})) + b \right) \end{aligned} \tag{C.7}$$

is also maximal. Now, since the function $\log((a-b) \cdot x + b)$ is increasing with respect to x, we conclude thatlog $\pi(\mathbf{y} \mid \mathbf{p})$ is maximal if $\sum_i \mathbf{1}_{[e]}(f_i(\mathbf{p}))$ is maximal, which corresponds to what OMNE computes. □

Now the "half uniform" distribution has two advantages. First, it is an appropriate approximation for a statistical distribution whose expression is not known, but for which it can be seen that most of its values are concentrated in a single mode whose width can be measured. A probability ν is assigned for events arising in this mode, and there is a probability $1 - \nu$ assigned for events arising outside this main mode, that can be looked at as rare events. The second advantage is that this kind of distribution can be easily represented in an interval framework. Indeed, consider a vector of two independents identically distributed half-uniform variables $\boldsymbol{e} = (e_1, e_2)$, with $\pi_e(e_i) = 0.8$ if $e_i \in [-1, 1]$, 0.2 otherwise. Figure C.5 illustrates their joint probability distribution function.

That probability density function satisfies the probabilities given on the following table.

In that scenario, there is a 64% chance that $\boldsymbol{e}$ will be in the box $\mathbb{K}_0 = [-1, 1] \times [-1, 1]$, a 96% chance that $\boldsymbol{e}$ will be in the cross-shaped domain $\mathbb{K}_1 = [-\infty, \infty] \times [-1, 1] \cup [-1, 1] \times [-\infty, \infty]$ and a 100% chance that $\boldsymbol{e}$ will be in $\mathbb{K}_2 = \mathbb{R}^2$. Now if we recall Chapter 4,

$[e_2] \backslash [e_1]$	$[-\infty, -1]$	$[-1, 1]$	$[1, \infty]$
$[1, \infty]$	0.01	**0.08**	0.01
$[-1, 1]$	**0.08**	**0.64**	**0.08**
$[-\infty, -1]$	0.01	**0.08**	0.01

Table C.1: Probabilities for the random vector described in Figure C.5

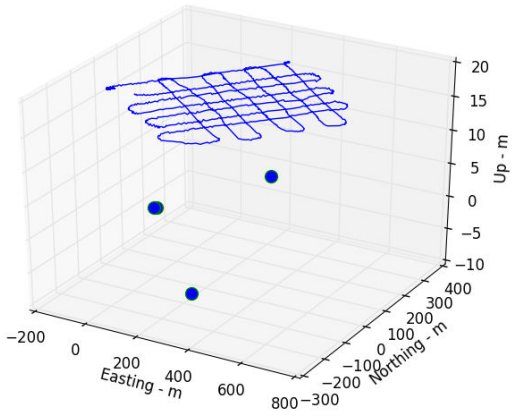


Figure C.6: The surface vessel's trajectory above the set of 4 acoustic beacons

interval representations for such domains are easily defined using q-relaxed boxes. Indeed, if we define the box $[\boldsymbol{x}] = [-1, 1] \times [-1, 1]$, we note that the q-relaxations of $[\boldsymbol{x}]$ corresponds to the $\mathbb{K}_i$'s : $[\boldsymbol{x}]^{\{0\}} = \mathbb{K}_0$, $[\boldsymbol{x}]^{\{1\}} = \mathbb{K}_1$, $[\boldsymbol{x}]^{\{2\}} = \mathbb{K}_2$. From Theorem 25, Corollaries 27 and 28, we have expressions to get the probabilities associated to the $\mathbb{K}_i$'s, and therefore the probabilities associated to the parameter set resulting from the inversion of the $\mathbb{K}_i$'s. Therefore, we will be able to associate a probability to the maximum likelihood set computed by OMNE. The following section illustrates this approach on real data.

C.4 APPLICATION

In this section, we will consider once again the data-set described in Section 2. As a reminder, a surface vessel followed a survey-like grid trajectory above a set of 4 acoustic beacons installed on the seabed. The surface vessel is precisely located by a GPS with RTK correction and its attitude is precisely determined with a PHINS INS. Additionally, the ship is equipped with a DVL and a RAMSES LBL positioning device. Each beacon's position is precisely known thanks to previous calibrations. The setup is illustrated on Figure C.6.

As it flies over the beacons, the RAMSES interrogates all the beacons at an individual rate of about 1Hz. The ranges are then de-

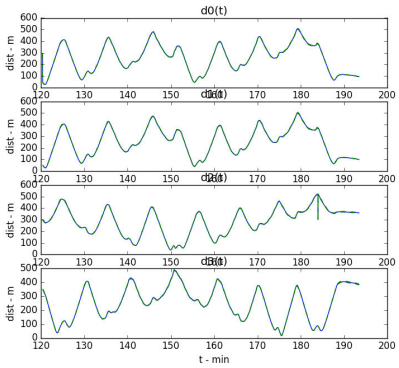


Figure C.7: Ranges measured by the RAMSES during the survey from the center of its acoustic head to the 4 beacons, as a function of time.

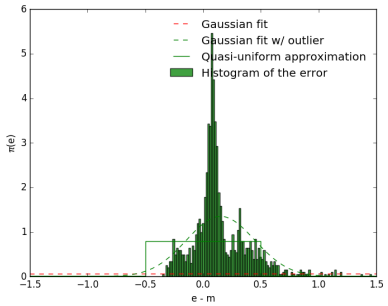


Figure C.8: Histogram of the error between the measured and the actual range for the third beacon, as well as a Gaussian approximation, a Gaussian approximation after removing 1% of the biggest errors, and a half-uniform approximation of the histogram.

duced from the measured time of flights, producing the signals on Figure C.7.

We can see that the signal for the third beacon gets a huge outlier toward the end of the survey. Therefore, estimating this beacon's position using the methods developed in Chapter 4 would likely fail unless a huge confidence threshold is used. Since we know the ship's and the beacon's positions with a centimetric precision, we are able to compute the histogram of the error between the measured range and the actual range to each beacon for the survey. For this illustration, we will work with the ranges to the third beacon, and the histogram of the error is depicted in Figure C.8.

We can see from Figure C.8 that the distribution seems to have three modes, and a tail on the positive side of the error. Even if an expression for such distribution could be found, it would be far from trivial to work with it. Therefore, a Gaussian modelization of the error seems like a good candidate. An other model is the previously discussed half-uniform distribution, represented by the green rectangle on Figure C.8. The green rectangle is the centered rectangle which

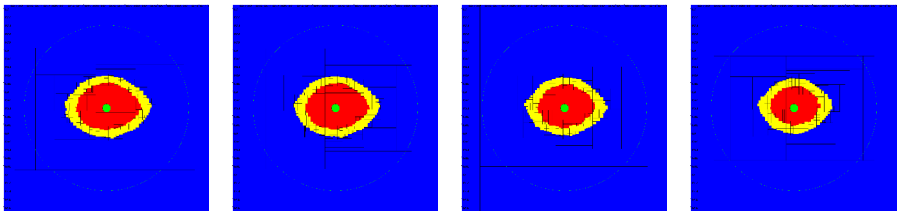


Figure C.9: From left to right: the 2d position of the beacon with 50% confidence, with 0.25% confidence, with $\sim$ 0.8% confidence, with $\sim 10^{-7}$% confidence. The green dot represents the true beacon position.

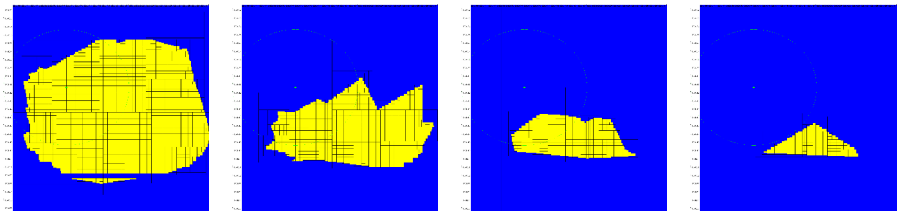


Figure C.10: From left to right: $q = 20, q = 17, q = 15, q = 14$. The green dot represents the true beacon position.

contains 80% of the errors, while its tail, not visible on the figure, contains the remaining 20%. By using the knowledge that the error is biased positively, we could have reduced the width of its main mode by shifting its center towards the right. It should be noted, however, that this bad approximation will only lead to a less precise estimation, but will not affect the probability associated with the estimated set. This is not the case for a mis-centered Gaussian approximation.

The next figures represent the confidence sets computed by GOMNE on the signal from the third beacon. The z-component of the beacon is assumed to be known for easier representation. The signals contains 1338 range measurements, and the sets are computed by dichotomy on q, the number of relaxations. The signal is assumed to be perturbed by a half-uniform distribution whose radius is 0.3% of the measured range, centered on the measured range.

Figure C.9 shows 4 level-sets with confidence thresholds ranging from 0.5 to 10^{-9}. Figure C.10 then represents 4 level sets on a zoomed region with confidence thresholds close to zero. From the last set computed by GOMNE, we have an error of approximately 10 centimeters which is well suited for practical application.

C.5 CONCLUSION

In this chapter, we have shown that OMNE and its nonlinear version, GOMNE, are Maximum Linear Estimators. By recalling observations made in Chapter 4, we have shown that a probability can be computed for the sets return by GOMNE, and that the smallest

non-empty set it returns is distant from about 8 centimeters from the real beacon's position. Now, we should note that there is no interior for the sets pictured on Figure C.10, which could mean that those sets may actually be empty, and that a more precise Maximum Likelihood estimate could be obtained by running GOMNE with a smaller ϵ.

A limitation of this method is that it is computationally expensive to run. It could be a useful tool as a double-checker for least-squares methods traditionally used in the industry. Also, this kind of algorithm is perfectly suited for used on massively parallel architecture, and could benefit from an implementation on a GPU [86] or on a FPGA [91] for example.

BIBLIOGRAPHY

[1] F. Abdallah, A. Gning, and P. Bonnifait. "Box particle filtering for nonlinear state estimation using interval analysis." In: *Automatica* 44.3 (2008), pp. 807–815 (cit. on p. 7).

[2] Fahed Abdallah, Amadou Gning, and Philippe Bonnifait. "Box particle filtering for nonlinear state estimation using interval analysis." In: *Automatica* 44.3 (2008), pp. 807–815. DOI: http://dx.doi.org/10.1016/j.automatica.2007.07.024. URL: http://www.sciencedirect.com/science/article/pii/S0005109807003731 (cit. on p. 58).

[3] Ignacio Araya, Gilles Trombettoni, Bertrand Neveu, and Gilles Chabert. "Upper bounding in inner regions for global optimization under inequality constraints." In: *J. Global Optimization* 60.2 (2014), pp. 145–164. DOI: 10.1007/s10898-014-0145-7. URL: http://dx.doi.org/10.1007/s10898-014-0145-7 (cit. on p. 96).

[4] Michael B. Monagan, Keith O. Geddes, K. Michael Heal, George Labahn, Stefan M. Vorkoetter, James McCarron, and Paul DeMarco. *Maple 10 Programming Guide*. Waterloo ON, Canada: Maplesoft, 2005 (cit. on p. 67).

[5] F. Benhamou, F. Goualard, L. Granvilliers, and J. F. Puget. "Revising Hull and Box Consistency." In: *Proceedings of the International Conference on Logic Programming*. Las Cruces, NM, 1999, pp. 230–244 (cit. on p. 89).

[6] A. Ben-Tal and A. Nemirovskii. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Philadelphia, PA: SIAM, 2001 (cit. on p. 94).

[7] J. Berger. *Statistical Decision Theory and Bayesian Analysis*, 2nd edition. New York, NY: Springer-Verlag, 1985 (cit. on pp. 57, 60).

[8] D. Berleant, L. Xie, and J. Zhang. "Statool: a tool for Distribution Envelope Determination (DEnv), an interval-based algorithm for arithmetic on random variables." In: *Reliable Computing* 9.2 (2003), pp. 91–108 (cit. on p. 57).

[9] D. Bertsekas and I. Rhodes. "Recursive state estimation for a set-membership description of uncertainty." In: *IEEE Transactions on Automatic Control* 16.2 (Apr. 1971), pp. 117–128. DOI: 10.1109/TAC.1971.1099674 (cit. on pp. 29, 39).

[10] J.L. Blanco, J. Gonzalez, and J.A. Fernández-Madrigal. "A pure probabilistic approach to range-only SLAM." In: *Proceedings of the IEEE International Conference on Robotics and Automation*. 2008, pp. 1436–1441 (cit. on p. 99).

[11] D. Bouvet and G. Garcia. "Guaranteed 3-D mobile robot localization using an odometer, an automatic theodolite and indistinguishable landmarks." In: *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*. Vol. 4. 2001, 3612–3617 vol.4. DOI: 10.1109/ROBOT.2001.933178 (cit. on p. 39).

[12] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Vol. 15. Studies in Applied Mathematics. Philadelphia, PA: SIAM, 1994 (cit. on pp. 92, 95).

[13] S. Boyd and L. Vanbenberghe. *Convex optimization*. Available at www.stanford.edu/∼boyd. Lecture notes at Stanford University and University of California at Los Angeles, 2003 (cit. on pp. 51, 92, 94).

[14] Clément Carbonnel, Gilles Trombettoni, Philippe Vismara, and Gilles Chabert. "Q-Intersection Algorithms for Constraint-Based Robust Parameter Estimation." In: *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada*. 2014, pp. 2630–2636. URL: http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8290 (cit. on p. 64).

[15] Clément Carbonnel, Gilles Trombettoni, Philippe Vismara, and Gilles Chabert. "Q-intersection Algorithms for Constraint-Based Robust Parameter Estimation." In: *AAAI'14 - Twenty-Eighth Conference on Artificial Intelligence*. Québec City, Canada, July 2014, pp. 2630-2636. URL: https://hal.archives-ouvertes.fr/hal-01084606 (cit. on p. 90).

[16] G. Chabert and L. Jaulin. "Contractor Programming." In: *Artificial Intelligence* 173 (2009), pp. 1079–1100 (cit. on pp. 29, 89, 96).

[17] Chen Tung Chen and Frank J. Millero. "Speed of sound in seawater at high pressures." In: *The Journal of the Acoustical Society of America* 62.5 (1977), pp. 1129–1135. DOI: http://dx.doi.org/10.1121/1.381646. URL: http://scitation.aip.org/content/asa/journal/jasa/62/5/10.1121/1.381646 (cit. on p. 21).

[18] Guanrong Chen, Jianrong Wang, and L. S. Shieh. "Interval Kalman filtering." In: *IEEE Transactions on Aerospace and Electronic Systems* 33.1 (Jan. 1997), pp. 250–259. DOI: 10.1109/7.570759 (cit. on p. 7).

[19] F.L. Chernousko. "Ellipsoidal state estimation for dynamical systems." In: *Nonlinear Analysis: Theory, Methods & Applications* 63.5-7 (2005). Invited Talks from the Fourth World Congress of Nonlinear Analysts (WCNA 2004), pp. 872–879. DOI: http://dx.doi.org/10.1016/j.na.2005.01.009. URL: http://www.sciencedirect.com/science/article/pii/S0362546X05000210 (cit. on p. 39).

[20] P. Chew and K. Marzullo. "Masking Failures of Multidimensional Sensors." In: *Tenth Symposium on Reliable Distributed Systems (SRDS)*. 1991, pp. 32–41 (cit. on p. 90).

[21] B. C. Csáji, M. C. Campi, and E. Weyer. "Sign-Perturbed Sums: A New System Identification Approach for Constructing Exact Non-Asymptotic Confidence Regions in Linear Regression Models." In: *IEEE Transactions on Signal Processing* 63.1 (Jan. 2015), pp. 169–181. DOI: 10.1109/TSP.2014.2369000 (cit. on p. 58).

[22] Andrew J. Davison, Ian D. Reid, Nicholas D. Molton, and Olivier Stasse. "MonoSLAM: Real-Time Single Camera SLAM." In: *IEEE Trans. Pattern Anal. Mach. Intell.* 29.6 (June 2007), pp. 1052–1067. DOI: 10.1109/TPAMI.2007.1049. URL: http://dx.doi.org/10.1109/TPAMI.2007.1049 (cit. on p. 2).

[23] Benoît Desrochers, Simon Lacroix, and Luc Jaulin. "Set-Membership Approach to the Kidnapped Robot Problem." In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015)*. Hamburg, Germany, Sept. 2015, 9p. URL: https://hal.archives-ouvertes.fr/hal-01182607 (cit. on pp. 2, 106).

[24] V. Drevelle and P. Bonnifait. "High integrity GNSS location zone characterization using interval analysis." In: *ION GNSS*. 2009 (cit. on pp. 3, 99).

[25] V. Drevelle and P. Bonnifait. "A set-membership approach for high integrity height-aided satellite positioning." In: *GPS Solutions* 15.4 (2011), pp. 357–368 (cit. on p. 106).

[26] V. Drevelle and J. Nicola. "VIBes: A Visualizer for Intervals and Boxes." In: *Mathematics in Computer Science* 8.3 (2014), pp. 563–572 (cit. on p. 96).

[27] Vincent Drevelle. "Study of robust set estimation methods for a high integrity multi-sensor localization. Application to navigation in urban areas." Theses. Université de Technologie de Compiègne, Dec. 2011. URL: https://tel.archives-ouvertes.fr/tel-00679502 (cit. on pp. 2, 102, 103).

[28] Vincent Drevelle and Philippe Bonnifait. "Localization Confidence Domains via Set Inversion on Short-Term Trajectory." In: *IEEE Transactions on Robotics* 29.5 (2013), pp. 1244–1256 (cit. on p. 58).

[29] Didier DuBois. *Fuzzy Sets and Systems: Theory and Applications*. Orlando, FL, USA: Academic Press, Inc., 1997 (cit. on p. 58).

[30] P. Eykhoff. *System Identification, Parameter and State Estimation*. London: John Wiley, 1974 (cit. on p. 57).

[31] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza. "IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation." In: Georgia Institute of Technology. 2015 (cit. on p. 82).

[32] Thor I Fossen. *Guidance and control of ocean vehicles*. John Wiley & Sons, Ltd, 1994 (cit. on p. 19).

[33] Katsuki Fujisawa, Mituhiro Fukuda, Kazuhiro Kobayashi, Masakazu Kojima, Kazuhide Nakata, Maho Nakata, and Makoto Yamashita. *SDPA (SemiDefinite Programming Algorithm) User's Manual – Version 7.0.5*. Tech. rep. 2008 (cit. on p. 96).

[34] C. Jauberthie, N. Verdière, and L. Travé-Massuyès. "Set-membership identifiability: definitions and analysis." In: *In Proceedings of the 18th IFAC World Congress, Milan, Italie*. Vol. 00. aaa, 2011, pp. 12024–12029 (cit. on pp. 43, 58).

[35] L. Jaulin. "Computing minimal-volume credible sets using interval analysis; Application to Bayesian estimation." In: *IEEE Trans. on Signal Processing* 54.9 (2006), pp. 3632–3636 (cit. on p. 57).

[36] L. Jaulin. "Probabilistic set-membership approach for robust regression." In: *Journal of Statistical Theory and Practice* 4.1 (2010) (cit. on pp. 58, 59).

[37] L. Jaulin. "Range-only SLAM with occupancy maps; A set-membership approach." In: *IEEE Transaction on Robotics* 27.5 (2011), pp. 1004–1010 (cit. on p. 99).

[38] L. Jaulin. "Set-membership localization with probabilistic errors." In: *Robotics and Autonomous Systems* 59.6 (2011), pp. 489–495 (cit. on p. 30).

[39] L. Jaulin and D. Henrion. "Contracting optimally an interval matrix without loosing any positive semi-definite matrix is a tractable problem." In: *Reliable Computing* 11.1 (2005), pp. 1–17 (cit. on p. 91).

[40] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis*. Springer, 2001 (cit. on pp. 49, 85, 89, 91).

[41] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. London: Springer-Verlag, 2001 (cit. on pp. 72, 90).

[42] L. Jaulin, M. Kieffer, E. Walter, and D. Meizel. "Guaranteed Robust Nonlinear Estimation with Application to Robot Localization." In: *IEEE Transactions on systems, man and cybernetics; Part C Applications and Reviews* 32.4 (2002), pp. 374–382 (cit. on p. 105).

[43] L. Jaulin and E. Walter. "Guaranteed nonlinear estimation and robust stability analysis via set inversion." In: *Proceedings of the 2nd European Control Conference*. 1993, pp. 818–821 (cit. on p. 25).

[44] L. Jaulin and E. Walter. "Guaranteed Nonlinear Parameter Estimation Via Interval Computations." In: *Conference on Numerical Analysis with Automatic Result Verification, (Lafayette)*. 1993 (cit. on p. 57).

[45] L. Jaulin and E. Walter. "Bounding Approaches to System Identification." In: ed. by M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter. Plenum, 1996. Chap. Guaranteed nonlinear set estimation via interval analysis, pp. 363–382 (cit. on p. 72).

[46] L. Jaulin and E. Walter. "Guaranteed robust nonlinear minimax estimation." In: *IEEE Transaction on Automatic Control* 47.11 (2002), pp. 1857–1864 (cit. on p. 106).

[47] Luc Jaulin. "Robust set-membership state estimation; application to underwater robotics." In: *Automatica* 45.1 (2009), pp. 202–206. DOI: http://dx.doi.org/10.1016/j.automatica.2008.06.013. URL: http://www.sciencedirect.com/science/article/pii/S0005109808003853 (cit. on p. 64).

[48] Luc Jaulin, Michel Kieffer, and Olivier Didrit. *Applied interval analysis : with examples in parameter and state estimation, robust control and robotics*. Pas de CD pour l'exemplaire acheté en 2008. London: Springer, 2001. URL: http://opac.inria.fr/record=b1097265 (cit. on p. 61).

[49] Simon J. Julier and Jeffrey K. Uhlmann. "A New Extension of the Kalman Filter to Nonlinear Systems." In: 1997, pp. 182–193 (cit. on p. 27).

[50] R. E. Kalman. "Contributions to the theory of optimal control." In: *Bol. Soc. Mat. Mex.* 5 (1960), pp. 102–119 (cit. on p. 25).

[51] R.E. Kalman. "Contributions to the Theory of Optimal Control." In: *Bol. Soc. Mat. Mex.* 5 (1960), pp. 102–119 (cit. on p. 26).

[52] George A Kantor and Sanjiv Singh. "Preliminary Results in Range-only Localization and Mapping." In: *Proceedings of the IEEE Conference on Robotics and Automation (ICRA '02)*. Vol. 2. May 2002, pp. 1818–1823 (cit. on p. 2).

[53] M. Karmarkar. "A new polynomial-time algorithm for linear programming." In: *Combinatorica* 4 (1984), pp. 373–395 (cit. on p. 95).

[54] M. Kieffer, L. Jaulin, and E. Walter. "Guaranteed Recursive Nonlinear State Estimation Using Interval Analysis." In: *Proceedings of the 37th IEEE Conference on Decision and Control*. Tampa, FL, 1998, pp. 3966–3971 (cit. on p. 29).

[55] M. Kieffer, L. Jaulin, and E. Walter. "Guaranteed recursive nonlinear state estimation using interval analysis." In: *Decision and Control, 1998. Proceedings of the 37th IEEE Conference on*. Vol. 4. Dec. 1998, 3966–3971 vol.4. DOI: 10.1109/CDC.1998.761917 (cit. on pp. 29, 90).

[56] M. Kieffer and E. Walter. "Guaranteed Characterization of Exact non-Asymptotic Confidence Regions as Defined by LSCR and SPS." In: *Automatica* 50.2 (2014), pp. 507–512 (cit. on p. 58).

[57] V. Kreinovich, G.P. Dimuro, and A. Carlos da Rocha Costa. "Probabilities, intervals, what next ? Extension of interval computations to situations with partial information about probabilities." In: *10th IMEKO TC7 International symposium*. 2004 (cit. on p. 57).

[58] W. Kühn. "Rigorously Computed Orbits of Dynamical Systems Without the Wrapping Effect." In: *Computing* 61 (1998), pp. 47–67 (cit. on p. 88).

[59] H. Lahanier, E. Walter, and R. Gomeni. "OMNE: a new robust membership-set estimator for the parameters of nonlinear models." In: *Journal of Pharmacokinetics and Biopharmaceutics* 15 (1987), pp. 203–219 (cit. on p. 105).

[60] Hélène Lahanier, Eric Walter, and Roberto Gomeni. "OMNE: A new robust membership-set estimator for the parameters of nonlinear models." In: *Journal of Pharmacokinetics and Biopharmaceutics* 15.2 (1987), pp. 203–219. DOI: 10.1007/BF01062344. URL: http://dx.doi.org/10.1007/BF01062344 (cit. on p. 106).

[61] Moussa Maiga, Nacim Ramdani, Louise Travé-Massuyès, and Christophe Combastel. "A CSP Versus a Zonotope-Based Method for Solving Guard Set Intersection in Nonlinear Hybrid Reachability." English. In: *Mathematics in Computer Science* 8.3-4 (2014), pp. 407–423. DOI: 10.1007/s11786-014-0204-y. URL: http://dx.doi.org/10.1007/s11786-014-0204-y (cit. on pp. 37, 39).

[62] K. A. Marzullo. "Maintaining the Time in a Distributed System: An Example of a Loosely-Coupled Distributed Service." PhD dissertation. Stanford University, 1984 (cit. on p. 90).

[63] M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, eds. *Bounding Approaches to System Identification*. New York, NY: Plenum Press, 1996 (cit. on p. 57).

[64] Ilya S. Molchanov. *Theory of random sets*. Probability and its applications. New York: Springer, 2005. URL: http://opac.inria.fr/record=b1102003 (cit. on p. 58).

[65] Nedialko S. Nedialkov and Kenneth R. Jackson. "A New Perspective on the Wrapping Effect in Interval Methods for Initial Value Problems for Ordinary Differential Equations." In: *in: Perspectives on Enclosure Methods*. Springer-Verlag, 1999, pp. 219–264 (cit. on pp. 37, 88).

[66] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial methods in convex programming*. Philadelphia, PA: SIAM, 1994 (cit. on p. 95).

[67] R. Neuland, J. Nicola, R. Maffei, L. Jaulin, E. Prestes, and M. Kolberg. "Hybridization of Monte Carlo and set-membership methods for the global localization of underwater robots." In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sept. 2014, pp. 199–204. DOI: 10.1109/IROS.2014.6942561 (cit. on pp. 7, 58).

[68] P Newman and J J Leonard. "Pure range-only sub-sea SLAM." In: *icra*. Vol. 2. 2003, pp. 1921–1926 (cit. on p. 2).

[69] J. Nicola. "Contractors and Linear Matrix Inequalities." In: *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B. Mechanical Engineering* 1.3 (2015) (cit. on p. 91).

[70] J Nicola and L Jaulin. "Comparison of Kalman and interval approaches for the simultaneous localization and mapping of an underwater vehicle." In: *Special Issue on Ocean Engineering and Oceanography, Springer* (2017) (cit. on pp. 25, 43).

[71] Jordan Ninin. "Global Optimization based on Interval Analysis: Affine Relaxation and Limited Memory." Theses. Institut National Polytechnique de Toulouse - INPT, Dec. 2010. URL: https://tel.archives-ouvertes.fr/tel-00580651 (cit. on pp. 61, 85).

[72] Jordan Ninin, Frédéric Messine, and Pierre Hansen. "An Automatic Linear Reformulation Technique Based on Affine Arithmetic." anglais. In: *International Symposium on Mathematical Programming (ISMP), Chicago, 23/08/2009-28/08/2009*. http://www.journals.uchicago.edu/: University of Chicago Press, 2009, p. 52 (cit. on p. 96).

[73] J. P. Norton, ed. Special Issue on Bounded-Error Estimation: Issue 1. *International Journal of Adaptive Control and Signal Processing* **8**(1):1–118. 1994 (cit. on p. 57).

[74] E. Olson, J.J. Leonard, and S. Teller. "Robust Range-Only Beacon Localization." In: *Oceanic Engineering, IEEE Journal of* 31.4 (Oct. 2006), pp. 949–958. DOI: 10.1109/JOE.2006.880386 (cit. on p. 25).

[75] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984 (cit. on p. 60).

[76] *PHINS Product Page*. https://www.ixblue.com/products/phins. Accessed: 2015-12-07 (cit. on p. 30).

[77] Boris T. Polyak, Sergey A. Nazin, Cécile Durieu, and Eric Walter. "Ellipsoidal parameter or state estimation under model uncertainty." In: *Automatica* 40.7 (2004), pp. 1171–1179. DOI: http://dx.doi.org/10.1016/j.automatica.2004.02.014. URL: http://www.sciencedirect.com/science/article/pii/S0005109804000676 (cit. on p. 39).

[78] R. E. Moore. *Interval Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1966 (cit. on p. 85).

[79] T. Raissi, N. Ramdani, and Y. Candau. "Set membership state and parameter estimation for systems described by nonlinear differential equations." In: *Automatica* 40 (2004), pp. 1771–1777 (cit. on p. 25).

[80] N. Ramdani and N. Nedialkov. "Computing Reachable Sets for Uncertain Nonlinear Hybrid Systems using Interval Constraint Propagation Techniques." In: *Nonlinear Analysis: Hybrid Systems* 5.2 (2011), pp. 149–162 (cit. on p. 37).

[81] *RAMSES Product Page*. https://www.ixblue.com/products/ramses. Accessed: 2015-12-07 (cit. on p. 30).

[82] Radhakrishna C. Rao. "Information and the accuracy attainable in the estimation of statistical parameters." In: *Bull. Calcutta Math. Soc.* 37 (1945), pp. 81–91. URL: http://www.ams.org/mathscinet-getitem?mr=0015748 (cit. on p. 70).

[83] B. Ristic. "Bayesian Estimation With Imprecise Likelihoods: Random Set Approach." In: *IEEE Signal Processing Letters* 18.7 (July 2011), pp. 395–398. DOI: 10.1109/LSP.2011.2152392 (cit. on p. 58).

[84] Vincent CREUZE. "Robots marins et sous-marins Perception, modélisation, commande." In: *Techniques de l'ingénieur Applications en robotique* base documentaire : TIB623DUO.ref. article : s7783 (2016). fre. eprint: basedocumentaire:TIB623DUO.. URL: http://www.techniques-ingenieur.fr/base-documentaire/automatique-robotique-th16/applications-en-robotique-42623210/robots-marins-et-sous-marins-s7783/ (cit. on pp. 13, 23).

[85] O. C. Rodriguez. *General description of the BELLHOP ray tracing program - Version 1.0*. Tech. rep. Universidade do Algarve - Physics Department, June 2008. URL: http://oalib.hlsresearch.com/Rays/index.html (cit. on p. 22).

[86] J. M. Singh and P. J. Narayanan. "Real-Time Ray Tracing of Implicit Surfaces on the GPU." In: *IEEE Transactions on Visualization and Computer Graphics* 16.2 (Mar. 2010), pp. 261–272. DOI: 10.1109/TVCG.2009.41 (cit. on p. 111).

[87] R. Smith, M. Self, and P. Cheeseman. "Autonomous Robot Vehicles." In: ed. by Ingemar J. Cox and Gordon T. Wilfong. New York, NY, USA: Springer-Verlag New York, Inc., 1990. Chap. Estimating Uncertain Spatial Relationships in Robotics, pp. 167–193. URL: http://dl.acm.org/citation.cfm?id=93002.93291 (cit. on p. 3).

[88] Olivier Strauss. "Quasi-continuous histograms." In: *Fuzzy Sets and Systems* 160.17 (2009). Theme: Learning, pp. 2442–2465. DOI: http://dx.doi.org/10.1016/j.fss.2009.01.013. URL: http://www.sciencedirect.com/science/article/pii/S0165011409000244 (cit. on p. 58).

[89] S. Thrun, W. Bugard, and D. Fox. *Probabilistic Robotics*. Cambridge, M.A.: MIT Press, 2005 (cit. on pp. 5, 28).

[90] Tuan Anh Tran, Carine Jauberthie, Françoise Le Gall, and Louise Travé-Massuyès. "Interval Kalman filter enhanced by positive definite upper bounds." In: *IFAC World Congrès*. Toulouse, France, July 2017, 6p. URL: https://hal.laas.fr/hal-01561951 (cit. on pp. 7, 43).

[91] Bastien Vincke. "Architectures pour des systèmes de localisation et de cartographie simultanées." 2012PA112332. PhD thesis. 2012. URL: http://www.theses.fr/2012PA112332/document (cit. on p. 111).

[92] Bastien Vincke and Alain Lambert. "Experimental comparison of Bounded-Error State Estimation and Constraints Propagation." In: *IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, 9-13 May 2011*. IEEE, 2011, pp. 4724–4729. DOI: 10.1109/ICRA.2011.5980313. URL: http://dx.doi.org/10.1109/ICRA.2011.5980313 (cit. on p. 29).

[93] E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. London, UK: Springer-Verlag, 1997 (cit. on p. 57).

[94] Eric A. Wan and Rudolph Van Der Merwe. "The Unscented Kalman Filter." In: *Kalman Filtering and Neural Networks*. Wiley, 2001, pp. 221–280 (cit. on pp. 25, 27).

[95] Wolfram Research Inc. *Mathematica 8.0*. 2010. URL: http://www.wolfram.com (cit. on p. 67).

[96] Jun Xiong, Carine Jauberthie, Louise Travé-Massuyes, and Françoise Le Gall. "Fault detection using interval Kalman filtering enhanced by constraint propagation." In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE. 2013, pp. 490–495 (cit. on pp. 43, 58).

Localisation et cartographie en simultané fiable, précise et robuste d'un robot sous-marin.

Mots-clés: SLAM, Filtre de Kalman, Méthodes ensemblistes, Analyse par intervalles, Méthodes probabilistes

Dans cette thèse on s'intéresse au problème de la localisation d'un robot sous-marin et de la cartographie en simultané d'un jeu de balises acoustiques installées sur le fond marin, en utilisant un distance-mètre acoustique et une centrale inertielle. Nous nous focalisons sur les deux approches principales utilisées pour résoudre ce type de problème: le filtrage de Kalman et le filtrage ensembliste basé sur l'analyse par intervalles. Le filtre de Kalman est optimal quand les équations d'état du robot sont linéaires et les bruits sont additifs, Gaussiens. Le filtrage par intervalles ne modélise pas les incertitudes dans un cadre probabiliste, et ne fait qu'une seule hypothèse sur leur nature: elles sont bornées. De plus, l'approche utilisant les intervalles permet la propagation rigoureuse des incertitudes, même quand les équations sont non linéaires. Cela résulte en une estimation hautement fiable, au prix d'une précision réduite.

Nous montrons que dans un contexte sous-marin, quand le robot est équipé avec une centrale inertielle de haute précision, une partie des équations du SLAM peut raisonnablement être considérée comme linéaire avec un bruit Gaussien additif, en faisant le terrain de jeu idéal d'un filtre de Kalman. De l'autre côté, les équations liées aux observations du distance-mètre acoustique sont bien plus problématiques: le système n'est pas observable, les équations sont non linéaires, et les outliers sont fréquents. Ces conditions sont idéales pour une approche à erreur bornées basée sur l'analyse par intervalles.

En prenant avantage des propriétés des bruits Gaussiens, cette thèse réconcilie le traitement probabiliste et ensembliste des incertitudes pour les systèmes aussi bien linéaires que non linéaires sujets à des bruits Gaussiens additifs. En raisonnant de manière géométrique, nous sommes capables d'exprimer la partie des équations du filtre de Kalman modélisant la dynamique du véhicule dans un cadre ensembliste. De la même manière, un traitement plus rigoureux et précis des incertitudes est décrit pour la parties des équations du filtre de Kalman liée aux mesures de distances.

Ces outils peuvent ensuite être combinés pour obtenir un algorithme de SLAM qui est fiable, précis et robuste. Certaines des méthodes développées dans cette thèse sont illustrées sur des données réelles.

Robust, precise and reliable simultaneous localization and mapping for and underwater robot.

Keywords: SLAM, Kalman filter, Set-membership methods, Interval analysis, Probabilistic methods

In this thesis, we work on the problem of simultaneously localizing an underwater robot while mapping a set of acoustic beacons lying on the seafloor, using an acoustic range-meter and an inertial navigation system. We focus on the two main approaches classically used to solve this type of problem: Kalman filtering and set-membership filtering using interval analysis. The Kalman filter is optimal when the state equations of the robot are linear, and the noises are additive, white and Gaussian. The interval-based filter do not model uncertainties in a probabilistic framework, and makes only one assumption about their nature: they are bounded. Moreover, the interval-based approach allows to rigorously propagate the uncertainties, even when the equations are non-linear. This results in a high reliability in the set estimate, at the cost of a reduced precision.

We show that in a subsea context, when the robot is equipped with a high precision inertial navigation system, a part of the SLAM equations can reasonably be seen as linear with additive Gaussian noise, making it the ideal playground of a Kalman filter. On the other hand, the equations related to the acoustic range-meter are much more problematic: the system is not observable, the equations are non-linear, and the outliers are frequent. These conditions are ideal for a set-based approach using interval analysis.

By taking advantage of the properties of Gaussian noises, this thesis reconciles the probabilistic and set-membership processing of uncertainties for both linear and non-linear systems with additive Gaussian noises. By reasoning geometrically, we are able to express the part of the Kalman filter equations linked to the dynamics of the vehicle in a set-membership context. In the same way, a more rigorous and precise treatment of uncertainties is described for the part of the Kalman filter linked to the range-measurements. These two tools can then be combined to obtain a SLAM algorithm that is reliable, precise and robust. Some of the methods developed during this thesis are demonstrated on real data.