Convex relaxation methods for opportunistic localization of a communications satellite

Renato Dias
renato.lourenco.dias@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2021

Abstract

CubeSats are an attractive solution due to their numerous applications and low-cost development, deployment and operation.
However, this reduction in cost impacts the quantity and quality of the equipment aboard the satellite. One of the sensors whose operation is more limited is the GPS receiver, which in some cases may not even be included with the satellite, which requires some alternative methods to obtain position information.
In this work, a study was made on the usage of information of opportunity to obtain a coarse position of a satellite using convex optimization methods. Several convex relaxations for the Doppler-shift cost function were developed to include this information in a formulation that utilizes already known convex relaxations for cost functions based on distance and angular information.
A preliminary theoretical study of the Cramér-Rao Lower Bound strongly suggests that the use of approximate functions for the Doppler-Shift cost function introduces additional information to the estimation problem. However, this study is not sufficient to conclude about the tightness of the developed relaxations which may degrade the estimator performance. This insight was confirmed in the simulations, where no relaxation method could outperform the formulation where no Doppler-shift information is used.
Nonetheless, additional relaxations were proposed to use cost functions based on the known movement laws of satellites. The combination of these regularization terms with the previously studied relaxations constitute good candidates to generate position estimates which may be used to initialize sequential algorithms which can provide refined estimates.
Keywords: Convex optimization, Convex relaxation, Doppler Shift, CRLB, CubeSats

1. Introduction

In the last two decades, GPS became a very useful tool for precise orbit determination of satellites in Low Earth Orbit (LEO) [11]. The usage of GPS measurements in combination with models of the forces (perturbations) acting on the satellite enable the trajectory to be estimated to an accuracy at the centimeter-level [7].
However these tools may not always be available as we are witnessing with the advent of CubeSats, small satellites with dimensions of $10\,\text{cm} \times 10\,\text{cm} \times 10\,\text{cm}$ which are used as a super low-cost alternative to conventional satellites but targeted for very specific tasks.
This new reality requires the study of alternative localization methods which can provide sufficient accuracy for the mission requirement of the satellite.
One source of information which is readily available and does not pose any energetic or computational overhead is the metadata that is inherent to the satellite's operation, such as the time-stamps present on each exchanged message, the pointing direction where the ground antennas find the greatest signal strength, and the received frequency. This free information is called: information of opportunity.

Throughout recent years, several studies were dedicated to the problem of orbit determination using information of opportunity. These first methods used range measurements from powerful radars and angle information based on radio interferometers. With this information, some crude estimates of satellite's position could be obtained and the orbit parameters calculated using previously developed analytical and geometrical methods (ex: Gibbs, Gauss, Laplace, Double-r) which were originally intended to study the orbits of planets. This process is known as preliminary orbit determination [17].
Beyond distance and angular measurements, there exists another valuable source of information that

may be used to help in obtaining correct orbital parameters: the Doppler-shift. This corresponds to the frequency drift of a sent, and subsequently received signal, which is a manifestation of the relative velocity between the emitter and the receiver. This valuable source of information is very cheap due to the modest hardware requirements for receiving the signals.
Once an initial estimate for orbital parameters is available, the common choice to improve the estimation accuracy is to utilize a batch estimator such as Least-Squares or a sequential estimator such as the Extended Kalman filter [10]. Both techniques allow the inclusion of more information sources and are prepared to cope with the measurement and modeling errors to refine the initial estimate in a process which is known as orbit estimation. This produces very good results for moderate measurement uncertainty but may have convergence problems for large noise magnitude. Old studies and application cases may be found in [6], [5], [13], and [4]. Similarly, recent works such as [3], apply a sequential correction algorithm that uses Doppler to refine the position estimates of a CubeSat with previously determined orbital elements.
Although not directly related to orbit determination, [1] follows a conceptually similar strategy to develop a localization algorithm (to estimate the crash site location of the plane involved in a very well-known accident of Malaysian Airlines, Flight MH370) using only range and Doppler-shift information, since there were no angular measurements available on that case. However, the way that Doppler-shift information was included differs from the one that is envisaged here, since it uses a particle filter requiring a heavy computational effort adding to the lack of convergence guarantees.

Based on the problems of the previously cited works, this thesis searches for an alternative approach to include Doppler-shift information in target localization. This study follows a line of work on convex algorithms for cooperative localization which have shown very good results in scenarios of sensor networks [16] although they solely used angular and distance information. Since the Doppler-shift effect has a clear manifestation in satellite communications, this constitutes a good candidate scenario to test the extension of the previously developed algorithms to include this additional source of information.
However, the results of this work are not intended to outperform the state-of-the-art algorithms in performance but may be useful when there is large uncertainty about the initial satellite's position, providing better estimates for initialization.
This is an innovative work on this field since to the best of our knowledge there is no published literature that uses convex optimization tools to perform preliminary orbit determination or orbit estimation.

The goal of this thesis is to formulate a fast and robust method for initial orbit determination and orbit estimation using information of opportunity associated with a satellite telecommunication link.
The problem scenario is composed of two agents: the satellite, in particular a CubeSat, and a ground station that observes it. It is assumed that the observer has few resources (which is close to the reality in an academic context) and so the measurements may present moderate to large error.
The estimation follows the classical Maximum Likelihood framework but adopts convex relaxation and optimization techniques rather than EKF-based tracking or local search algorithms as a way of attaining improved robustness to model uncertainties which may cause the algorithm to diverge. This builds on previous work [16] which developed an efficient convex relaxation approach for collaborative localization of autonomous vehicles combining range and angular information. The novelty introduced in this thesis is the incorporation of Doppler-shift information that is present in a satellite communication context, as well as the integration of motion models into the problem that are more pertinent for satellite orbits than the generic ones adopted in [16].
Extending the approach to a collaborative context with more than one target and one observer is envisaged as a future development.

2. Background

2.1. Coordinate Systems

The motion of Earth-orbiting satellites is usually described in an Earth-Centered Inertial (ECI) frame. However, the majority of observations are made at ground sites and the output of sensors are relative to the *Local Tangent* (LT) coordinate system or *Topocentric Horizon* (TH) coordinate system.

In order to convert an observation taken at the LT frame to the ECI frame, an intermediary frame must be used to account for Earth's rotation. This coordinate system is called *Earth-Centered Earth Fixed* (ECEF) frame. This frame rotates along with Earth such that a point at Earth's surface always has the same coordinates. A comparison is given in Fig. 1.

3. Celestial Mechanics

In order to study the relations between some observations and the true position of a satellite, its dynamics must be addressed in first place.

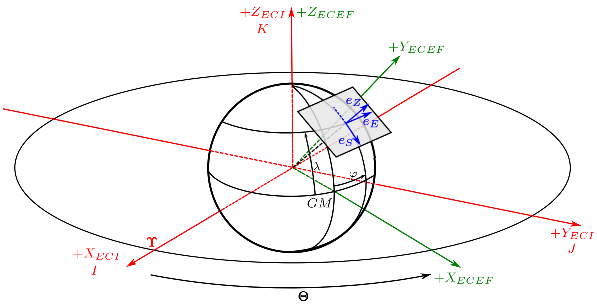

Figure 1: Representation of ECI frame (Red), ECEF frame (Green), *Local Tangent* frame (Blue). Θ is the Greenwich Apparent Siderial Time, a time-varying angle between the Vernal Equinox (Υ) and the Greenwich meridian (GM) that increases by 360° every 24 h. Ground station's latitude and longitude are represented by λ and φ respectively. (Figure based on [10]).

3.1. Two Body Problem

The gravitational interaction between two spherically symmetric bodies with reference to an inertial frame is described by Newton's law of universal gravitation. Denoting the mass of each body by M and m, the distance between them by r and the universal constant of gravitation by G, the magnitude of the force acting on each body (assuming no other forces than gravity) is given by the *inverse-square* law

$$\boldsymbol{F_{grav}} = \frac{GMm}{r^2}\frac{\boldsymbol{r}}{r}. \tag{1}$$

Since $\boldsymbol{F_{grav}}$ is a vector quantity which is colliniar with $\boldsymbol{r}$, the direction is captured by the unitary vector $\frac{\boldsymbol{r}}{r}$.

Due to the difference of nearly 22 orders of magnitude between the mass of the bodies, the term m may be neglected leading to

$$\ddot{\boldsymbol{r}} \approx -\frac{\mu}{\|\boldsymbol{r}\|^3}\boldsymbol{r} \tag{2}$$

where $\mu = GM$ is called *gravitational parameter*. Physically, this means Earth's center is approximately coincident with the center of the inertial frame and in this case, $\boldsymbol{r}$ becomes the position vector of the satellite with $\boldsymbol{r} = [x\ y\ z]$.
From (2) it is possible to show that a satellite's position and velocity are contained on a plane that does not change over time which is encoded in the constant of motion, *angular momentum per unit mass*, $\boldsymbol{h} = \boldsymbol{r} \times \boldsymbol{v}$.

3.2. Keplarian Elements

There is an alternative description of a body's movement in the conditions verified by the two body problem, developed by Johannes Kepler.

According to Kepler's formulation, an orbit may be described with a set of 6 parameters: (e, a, i, Ω, ω, θ). The five first parameters define the shape and orientation of the orbit in space and the last one accounts for satellite's position along the trajectory. These are known as *orbital elements*.

Given a set of values for the orbital parameters at some epoch t_1, Kepler equations provide a convenient way to determine the orbital elements at some future time t_2.

3.3. Orbit Perturbations

Given the increased complexity of the additional perturbing terms and observing the weak effect of these forces on the satellite, for the problem at hand, the two body problem assumptions are deemed sufficiently good to meet the accuracy expected for the position estimates.

4. Mathematical Background

4.1. Cramér-Rao Lower Bound

In order to determine the best performance achievable by an estimator and to benchmark it against other candidate estimators it is necessary to have a statistical tool for efficiency evaluation. For an unbiased estimator, the easiest way to determine the lower bound for the variance of the estimator is the *Cramér-Rao Lower Bound* (CRLB)[8].

When several parameters are to be determined, the variance for the estimator of each parameter verifies

$$\mathrm{Var}[\hat{\boldsymbol{\theta}}_i] \geq [\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{ii} \tag{3}$$

where $\boldsymbol{\theta}$ is the vector of parameters and $\boldsymbol{I(\theta)}$ is the *Fisher Information Matrix.*

Each entry of $\boldsymbol{I(\theta)}$ is given by

$$\boldsymbol{I(\theta)}_{ij} = -E\left[\frac{\partial^2 \ln p(\boldsymbol{x};\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right] \tag{4}$$

where $p(\boldsymbol{x};\boldsymbol{\theta})$ is the *pdf* of $\boldsymbol{x}$ which has a dependence on the vector parameter $\boldsymbol{\theta}$.

4.2. MLE Framework

The objective of the Maximum Likelihood estimation is to find $\hat{\boldsymbol{\theta}}_{ML}$ that maximizes $p(\boldsymbol{x};\boldsymbol{\theta})$ *i.e* that makes the observed data most probable.

The problem can be formulated as

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{x};\boldsymbol{\theta}) \tag{5}$$

Here, the *pdf* is viewed as a function of the unknown parameter for an observed $\boldsymbol{x}$ and so is called the *likelihood-function* $L(\boldsymbol{x};\boldsymbol{\theta}) = p(\boldsymbol{x};\boldsymbol{\theta})$

If the observations are independent and identically distributed (IID), the joint probability function can be written as the product of the *pdf*'s for each observation leading to

$$L(\boldsymbol{x},\boldsymbol{\theta}) = p(x(1),\dots,x(N);\boldsymbol{\theta}) = \prod_{n=1}^{N} p(x(n);\boldsymbol{\theta}). \tag{6}$$

If $L(\boldsymbol{x},\boldsymbol{\theta})$ is strictly positive (which certainly is for Gaussian distributions) the logarithm is applied to ease the computational implementation of the maximization problem, leading to

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} \ \ell(\boldsymbol{\theta};\boldsymbol{x}) = \arg\min_{\boldsymbol{\theta}} \ -\ell(\boldsymbol{\theta};\boldsymbol{x}) \tag{7}$$

where $\ell(\boldsymbol{\theta};\boldsymbol{x}) = \log L(\boldsymbol{x};\boldsymbol{\theta})$ is called the *log-likelihood function* which is related with (6) by

$$\ell(\boldsymbol{\theta};\boldsymbol{x}) = \log L(\boldsymbol{x};\boldsymbol{\theta}) = \sum_{n=1}^{N} \log p(x(n);\boldsymbol{\theta}). \tag{8}$$

4.3. Convex Optimization

When trying to find the global minima of (7) using an iterative algorithm (*e.g.*, gradient method), it's not surprising that the algorithm may converge to a sub-optimal (local) solution. However, if the cost function can be changed to ensure the existence of only one minimum, the iterative algorithm will converge to the optimal solution. This technique is called *convex relaxation*.

One of the most important aspects for this thesis is the set of rules that, when verified, enable one to conclude without further analysis if a function if convex or not. The scheme presented in Fig.(2) summarizes the constructions that guarantee the creation of a convex function.

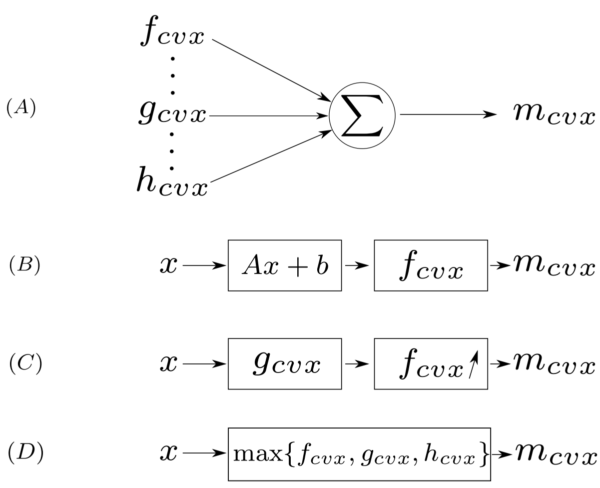

Figure 2: Disciplined Convex Programming ruleset.

5. Measurement model

5.0.1 Pseudo-Range

The *pseudo-range* (PR) is the distance between an emitter and a receiver computed from the travel time of a message exchanged between them. In this case, it corresponds to the difference between the time-stamp, t_s of a packet sent by a satellite and the time of reception by the ground station, t_r.

The expression that relates the position of the satellite $\boldsymbol{r}$ and the station $\boldsymbol{r}_{gs}$ with the PR measurement may be written as

$$\rho^k = c\,(t_r^k - t_s^k + \delta_{RF}^k). \tag{9}$$

where δ_{RF}^k is the time measurement error introduced by the RF system $\delta_{RF}^k \sim \mathcal{N}(0,\sigma_t^2)$ under the simplifying assumption that both agent clocks are synchronized.

The resulting log-likelihood function for this model is then

$$\ell(\rho^k;\boldsymbol{r}^k) = \log(\sqrt{2\pi}\sigma_p) - \frac{\left(\rho^k - \left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|\right)^2}{2\sigma_p^2} \tag{10}$$

5.0.2 Angle-of-Transmission

The angle-of-transmission corresponds to the unit vector pointing from the ground station antenna to the satellite. Under the assumption of an ideal emission/receiving system, the angle-of-transmission is related with the position of the ground station and a position of a satellite by

$$AoT^k = \frac{\boldsymbol{r} - \boldsymbol{r}_{gs}}{\|\boldsymbol{r} - \boldsymbol{r}_{gs}\|}. \tag{11}$$

This random variable may be described by the the Von-Mises Fisher distribution [12] with mean given by (11) and $\kappa > 0$. The log-likelihood function may then be written as

$$\ell(AoT^k;\boldsymbol{r}^k) = \log(2\pi(e^{\kappa} - e^{-\kappa})) + \tag{12}$$

$$\kappa \frac{(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k)^T AoT^k}{\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|} \tag{13}$$

5.0.3 Doppler Shift

An emitter that sends a signal while it is approaching/moving away from a receiver, causes a frequency change in the emitted signal that increases/decreases its frequency, respectively.

An equation that relates the Doppler Shift $\Delta f^k = f_r^k - f_t^k$ measured at ground station $\boldsymbol{r}_{gs}^k$ with the position of the satellite $\boldsymbol{r}^k$, is

$$\Delta f^k = -\frac{f_t^k}{c}\frac{d}{dt}\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|. \tag{14}$$

Taking the derivative in (14) and expanding, results in

$$\Delta f^k = -\frac{f_t^k}{c}\frac{(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k)^T(\boldsymbol{v}^k - \boldsymbol{v}_{gs}^k)}{\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|}. \tag{15}$$

The measured Doppler Shift error is assumed to follow a normal distribution, $\Delta f^k_{error} \sim \mathcal{N}(0, \sigma_f^2)$.

The log-likelihood function may then be written as

$$\ell\left(\Delta f^k; \boldsymbol{r}^k, \boldsymbol{v}^k\right) = \log(\sqrt{2\pi}\sigma_f) - \frac{\left(\Delta f^k + \frac{f_t^k}{c}\frac{(\boldsymbol{r}^k - \boldsymbol{r}^k_{gs})^T(\boldsymbol{v}^k - \boldsymbol{v}^k_{gs})}{\left\|\boldsymbol{r}^k - \boldsymbol{r}^k_{gs}\right\|}\right)^2}{2\sigma_f^2} \tag{16}$$

5.1. Complete ML formulation

The total log-likelihood function may then be written as

$$\begin{aligned}\ell_(\rho^k, AoT^k, \Delta f^k; \boldsymbol{r}^k, \boldsymbol{v}^k) = \sum_{N}^{k=1} & \ell(\rho^k; \boldsymbol{r}^k) + \\ & \ell(AoT^k; \boldsymbol{r}^k) + \\ & \ell\left(\Delta f^k; \boldsymbol{r}^k, \boldsymbol{v}^k\right)\end{aligned} \tag{17}$$

and the problem of finding the ML estimator follows from (7),

$$\hat{\theta}_{ML} = -\arg\min_{\boldsymbol{r}^k, \boldsymbol{v}^k} \ell(\rho^k, AoT^k, \Delta f^k; \boldsymbol{r}^k, \boldsymbol{v}^k) \tag{18}$$

Now, it is necessary to find relaxations that may be applied to (17) in order to solve (18) using convex optimization tools.

5.2. Known relaxations

5.2.1 Pseudo-range

The function in (10) is non-convex due to the argument of the exponential term in the numerator.

To overcome this issue, a relaxation technique developed in [15] may me adapted to this case to formulate an alternative convex function. Instead of trying to minimize de difference between the real distance and the measured distance between the agents, the requirement is relaxed to minimize the difference only for positions that are outside the circle with a radius equal to the measurement value, d, as depicted in Fig. 3.

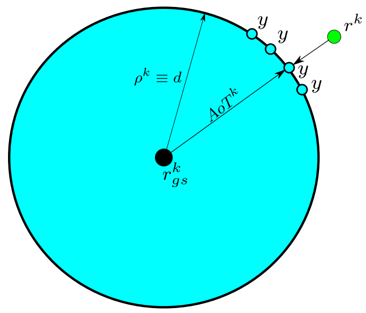

Figure 3: Visual representation of the convex relaxation for distance and angle measurement. Based on [16]. A disk was used for simplicity of representation but the conclusions are identical for higher dimensions. The range measurement obtained from the sensor ρ^k now define the value of d and $y \in \mathcal{D}(b_k, d)$.

Although in [16] this modification enables a parallel implementation of the localization algorithm, in this centralized case there is no need to introduce this new estimation variable since, equivalently, the difference $\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}\right\| - \rho^k$ may simply be clipped to zero when it takes on negative values.

The resulting log-likelihood function of applying this relaxation to (10) is

$$\ell(\rho^k; \boldsymbol{r}^k) = \log(\sqrt{2\pi}\sigma_\rho) - \frac{\left[\left(\left\|\boldsymbol{r}^k - \boldsymbol{r}^k_{gs}\right\| - \rho^k\right)_+\right]^2}{2\sigma_\rho^2} \tag{19}$$

where $(x)_+ = \max(0, x)$.

5.2.2 Angle of Transmission

The cost function in (13) is also non-convex due to the denominator term $\|\boldsymbol{r} - \boldsymbol{r}_{gs}\|$ in the argument of the exponential. A relaxation for the angle term was developed in [16] based on the reasoning made for the distance term relaxation.
According to Fig. 3, since $\boldsymbol{y}$ corresponds to the projection of $\boldsymbol{r}^k$ onto the disk when $\boldsymbol{r}^k$ is outside of it, $\frac{\boldsymbol{y}}{d}$ will be a unit norm vector that encodes the angle between the two agents, so, the non-convex term $\frac{(\boldsymbol{r}^k - \boldsymbol{r}^k_{gs})^T AoT^k}{\left\|\boldsymbol{r}^k - \boldsymbol{r}^k_{gs}\right\|}$ can be approximated by $\frac{\boldsymbol{y}^T AoT^k}{d}$.
Adapting this formulation to the problem addressed in this thesis where the range error is a tiny fraction of the true range between the agents, the normalization may be simply approximated by $\frac{\boldsymbol{r}^k - \boldsymbol{r}^k_{gs}}{\rho^k}$.
Given the differences in the magnitude of the measured range and the true range, the impact of this approximation on the minimization is small since the denominator is acting almost like a normalization constant with the great advantage having just half of the variables to optimize. The resulting log-likelihood function of applying this relaxation to

(13) is

$$\ell(AoT^k; \boldsymbol{r}^k) = \log(2\pi(e^\kappa - e^{-\kappa})) + \kappa \frac{\left(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right)^T AoT^k}{\rho^k} \tag{20}$$

.

6. Implementation

Equation (16) does not fit the framework of *Disciplined Convex Programming* problems since its construction does not follow the rules summarized in Sec. 4.3.

In fact it's easy to spot two terms that can't be constructed according to those rules, namely, a bi-linear term $(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k)^T(\boldsymbol{v}^k - \boldsymbol{v}_{gs}^k)$ and a vector norm in the denominator $\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|$.

6.0.1 Doppler function relaxation options

For simplicity of writing, all optimization problems formulated in this section only focus on the Doppler-shift cost function term.

6.1. Discrete derivative relaxation

Since the measurements are taken in discrete time, the derivative term in (14) can't be directly used. The first and most intuitive approach is to approximate the continuous time derivative by its discrete counterpart.

For a generic function f, an approximate second-order finite-divided-difference first derivative, $f'(x_k)$, may be obtained from 4 consecutive values of $f(x_k)$ from instant $(k-2)$ to $(k+2)$ using

$$f'(x_k) = \frac{-f(x_{k+2}) + 8f(x_{k+1}) - 8f(x_{k-1}) + f(x_{k-2})}{12\Delta t} \tag{21}$$

Other approximations and further discussion may be found in [2].

The result of the substitution of (21) in (14) is given by (22) at the bottom of this page.

Inspecting (22), it becomes clear that it can't be built by the rules presented in Sec. 4.3 since there is a term containing the difference of two convex functions which is not guaranteed to be convex.

6.1.1 Epigraph relaxation

The first and most direct approach to the convex relaxation problem is the well known epigraph relaxation.

The method consists in substituting the convex functions (which form the full cost function) with a new set of variables and enforcing the solution of the problem for the new variables to be contained on the epigraph of the convex functions replaced by them. The approximated problem written in the canonical form of a convex optimization problem is given in (23).

It is now a convex optimization problem since it has a convex cost function (composition of an affine function with a convex function) and second-order cone constraints which amounts to requiring that the affine function $(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k, t^k)$ lie in a convex cone in $\mathbb{R}^4$.

Note also that a penalization term is required for the vector $\boldsymbol{t}$ [18], otherwise the constraints become loose. Note also that a weighting factor must be applied to the penalizing term, otherwise the norm terms in inequality constraints would be strangled to zero since the penalization term has a much bigger value than the Doppler cost term.

6.1.2 Direct substitution by measurements

Another possible way to ensure convexity in (22) is to remove the term that ruins the convexity of the expression. In the case where range measurements are available, they may be used to replace the norm being subtracted, leading to a modified convex optimization problem written in (24).

However, this creates a strong influence of the range measurement error on the estimated position using the Doppler-shift equation.

6.1.3 Spherical-like Coordinates

Another possible method is to separate the problem into two independent sub-problems for distance estimation and direction estimation, something that could be accomplished by using spherical coordinates.

Ideally, the polar coordinate system would have the same origin as the ECI frame. However, since all measurements are taken from a ground station, the origin's placement of the spherical coordinate frame is forced to lie at the station's location.

This eases the use of the Doppler-shift cost function but severely limits any constraint relative to satellite dynamics since these constraints are expressed in the ECI frame as presented in Sec. 3.1. In this manner $\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k$ is written in terms of its magnitude and direction vector given by r^k and $\hat{\boldsymbol{u}}^k$, respec-

$$F(r^k) = \Delta f^k + \frac{f_{emm}}{c} \frac{-\left\|\boldsymbol{r}^{k+2} - \boldsymbol{r}_{gs}^{k+2}\right\| + 8\left\|\boldsymbol{r}^{k+1} - \boldsymbol{r}_{gs}^{k+1}\right\| - 8\left\|\boldsymbol{r}^{k-1} - \boldsymbol{r}_{gs}^{k-1}\right\| + \left\|\boldsymbol{r}^{k-2} - \boldsymbol{r}_{gs}^{k-2}\right\|}{12\Delta t} \tag{22}$$

tively. Now it is straightforward to use the Doppler cost function with these new variables since it only involves the information about the range which is represented by r.
The resulting optimization problem is given in (25).

6.2. Analytical derivative relaxation

The previously introduced equation (15) uncovers a new unknown quantity: the satellite's velocity.

Again, this expression cannot directly be used to build a convex cost function since it depends on the bi-linear term $(\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k)^T(\boldsymbol{v}^k - \boldsymbol{v}_{gs}^k)$ and contains the already mentioned problematic norm term in the denominator.

6.2.1 Substitution by estimated positions

The first approach to convexify (15) using a convex relaxation may be to replace the positions, $\boldsymbol{r}^k$, by the respective measurements, $\hat{\boldsymbol{r}}^k$, and approximate $\hat{\boldsymbol{v}}^k$ according to (21). If $\boldsymbol{r}^k$ is chosen to be replaced, the problem of the denominator is solved and the expression ends up as an affine function.

$$F = \sum_{k=2}^{N}\left(\Delta f^k + \frac{f_{emm}}{c}\frac{(\hat{\boldsymbol{r}}^k - \boldsymbol{r}_{gs}^k)^T(\boldsymbol{v}^k - \boldsymbol{v}_{gs}^k)}{\left\|(\hat{\boldsymbol{r}}^k - \boldsymbol{r}_{gs}^k)\right\|}\right)^2. \tag{26}$$

Nevertheless, this can be transformed to a problem whose estimation variables are only positions if the velocity term is approximated by its discrete derivative, as in (22), so it also needs the additional constraint

$$\boldsymbol{v}^k = \frac{-\boldsymbol{r}^{k+2} + 8\boldsymbol{r}^{k+1} - 8\boldsymbol{r}^{k-1} + \boldsymbol{r}^{k-2}}{12\Delta t} \tag{27}$$

6.2.2 Ping-Pong Position-Velocity

An attempt to reduce the influence of the distance measurement errors on the cost function in (26) is to resort to alternating projections, commonly known as the ping-pong method. Beginning with raw position estimates from sensors $\hat{\boldsymbol{r}}$, the optimization problem is solved first with $\boldsymbol{v}$ as the optimization variable but introducing a regularization term, g^k, to induce agreement between position and velocity

$$g^k = \delta\left(\boldsymbol{v}^k - \frac{-\hat{\boldsymbol{r}}^{k+2} + 8\hat{\boldsymbol{r}}^{k+1} - 8\hat{\boldsymbol{r}}^{k-1} + \hat{\boldsymbol{r}}^{k-2}}{12\Delta t}\right)^2, \tag{28}$$

with a penalizing term δ which should be chosen according to the quality of the estimates.

After that, the problem is solved again with the same cost function and the same regularization term, but alternating the optimization variable to $\boldsymbol{r}$ and using $\hat{\boldsymbol{v}}$ from the previous iteration.
The case when $\boldsymbol{r}$ is to be estimated is more problematic since the norm in the denominator is no longer a constant. To circumvent this problem, that norm is approximated with the estimated position vector $\hat{\boldsymbol{r}}$ of the previous iterations.
As in typical alternating minimization methods, each sub-problem converges to its minimum but there are no guarantees that a global optimum is ever reached.

The same procedure can be applied to the case where the angles are substituted, but it is omitted here for brevity.

6.2.3 SDP based relaxation

One of the most common choices to tackle a non-convex problem with quadratic constraints is to transform it into a Semidefinite Program (SDP). Although the relaxation developed does not lead to an SDP since it produces a quadratic cost function (instead of a linear one), it borrows the techniques used in that class of problems. In particular, the idea to leverage the complicated handling of non-convex bi-linear terms which appear when one expands (15) by encoding them as entries of a suitably constructed matrix, $\boldsymbol{G}$.

Expanding (15) one obtains

$$\begin{aligned}&\frac{1}{\rho}(\Delta f_k \left\|\boldsymbol{r}_k - \boldsymbol{r}_{gs}^k\right\| + \frac{f_{emm}}{c}(\frac{\boldsymbol{r}_k^T\boldsymbol{r}_{k+1}}{\Delta t}\\ &- \frac{\boldsymbol{r}_k^T\boldsymbol{r}_k}{\Delta t} - \boldsymbol{r}_k^T v_k - \frac{\boldsymbol{r}_{gs}^{k\,T}\boldsymbol{r}_{k+1}}{\Delta t} + \\ &\frac{\boldsymbol{r}_{gs}^{k\,T}\boldsymbol{r}_k}{\Delta t} + \boldsymbol{r}_{gs}^{k\,T} v_k)) = \epsilon\end{aligned} \tag{29}$$

$$\begin{aligned}\underset{\boldsymbol{x},\boldsymbol{t}}{\text{minimize}} \quad & \sum_{k=2}^{N-2}\left(\Delta f^k + \frac{f_{emm}}{c}\frac{-t^{k+2} + 8t^{k+1} - 8t^{k-1} + t^{k-2}}{12\Delta t}\right)^2 + \delta\mathbf{t}^T\mathbf{t}\\ \text{subject to} \quad & \left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\| < t^k \qquad k = 1,\dots,N\end{aligned} \tag{23}$$

$$F(\boldsymbol{r}^k) = \left(\Delta f^k + \frac{f_{emm}}{c}\frac{-\rho^{k+2} + 8\left\|\boldsymbol{r}^{k+1} - \boldsymbol{r}_{gs}^{k+1}\right\| - 8\rho^{k-1} + \left\|\boldsymbol{r}^{k-2} - \boldsymbol{r}_{gs}^{k-2}\right\|}{12\Delta t}\right)_+^2. \tag{24}$$

To overcome the non-convexity introduced by the norm at denominator in (15), the norm is factored out of the expression and approximated by the range measurement, as proposed in [14].
It is important to note that a norm term appears in (29) multiplying the measured noisy frequency difference, Δf^k, which in the end may amplify the error in measured Doppler-shift.

A suitable choice for $\boldsymbol{G}$ could be

$$\mathbf{G} = \boldsymbol{g}\boldsymbol{g}^T \tag{30}$$

where

$$\boldsymbol{g} = \begin{bmatrix} a_x & a_y & a_z & b_x & b_y & b_z & n_1 & n_2 & 1 \end{bmatrix}. \tag{31}$$

substituting, for shorter representation,

$$\boldsymbol{a} = \boldsymbol{r}^{k-1} \tag{32}$$

$$\boldsymbol{b} = \boldsymbol{r}^{k} \tag{33}$$

$$n_1 = \left\|\boldsymbol{r}^{k-1} - \boldsymbol{r}_{gs}^{k-1}\right\| \tag{34}$$

$$n_2 = \left\|\boldsymbol{r}^{k} - \boldsymbol{r}_{gs}^{k}\right\| \tag{35}$$

$$\tag{36}$$

Now it is possible to find the linear and bi-linear terms to construct (29) since these appear directly in specific entries of $\boldsymbol{G}$.

Due to the structure of $\boldsymbol{G}$, additional constraints are needed to ensure a more complete description of the relationships between matrix entries. The construction of $\boldsymbol{G}$ according to (30) automatically implies that this matrix is symmetric positive semidefinite and has unitary rank.

$$\boldsymbol{G}(4N+1, 4N+1) = 1 \tag{37}$$

$$\mathbf{G} \succeq 0 \tag{38}$$

$$\operatorname{rank}(\boldsymbol{G}) = 1 \tag{39}$$

Although (37) and (38) are convex constraints, (39) is not. The easiest solution is to drop this constraint, although there are approaches based on statistical properties of the solution as discussed in [9].

Also, one may observe that the squared norm terms in the diagonal and the position terms in the last column of $\boldsymbol{G}$ may be used to create the following additional constraints

$$\left\|\boldsymbol{r}^k - \boldsymbol{r}_{gs}^k\right\|^2 = \left\|\boldsymbol{r}^k\right\|^2 - 2{\boldsymbol{r}_{gs}^k}^T \boldsymbol{r}^k + \left\|\boldsymbol{r}_{gs}^k\right\|^2 \qquad k = 1, \ldots, N. \tag{40}$$

Finally, an additional constraint arrives from satellite dynamics by forcing the cross-product between position and velocity to be constant over time. It is then possible to add q as a new variable to the optimization problem and create an additional set of constraints

$$\begin{aligned} \underset{\boldsymbol{r},\boldsymbol{k}}{\text{minimize}} \quad & F_f^k(\boldsymbol{r}) \\ \text{subject to} \quad & \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \frac{1}{\Delta t} \begin{bmatrix} \boldsymbol{r}_y^k \boldsymbol{r}_z^{k+1} - \boldsymbol{r}_z^k \boldsymbol{r}_y^{k+1} \\ \boldsymbol{r}_z^k \boldsymbol{r}_x^{k+1} - \boldsymbol{r}_x^k \boldsymbol{r}_z^{k+1} \\ \boldsymbol{r}_x^k \boldsymbol{r}_y^{k+1} - \boldsymbol{r}_y^k \boldsymbol{r}_x^{k+1} \end{bmatrix} \\ & k = 1, \ldots, N \end{aligned} \tag{41}$$

7. Dynamics based complementary functions

Until now, the cost functions developed were based only on the available measurements (except in the SDP based formulation, with the introduction of the angular momentum constraint). However, some complementary cost functions may be combined with each Doppler-shift cost function obtained with the convex relaxations discussed in previous sections. These complementary functions contain information about the satellite's dynamics and thus may be used to improve the position estimation as they introduce correlation between positions. Once again, given the rigorous set of rules regarding the construction of convex cost functions, it is hard to find admissible regularization terms based on the dynamic's properties.

Due to space limitations, this topic was only briefly outlined here.

8. Results

To produce simulated measurements, the (simulated) true measurements were used as mean value for a Gaussian distribution with the standard deviations indicated in Tab. 1.

Standard Deviation	Small Noise
Range (σ_p)	0.2 km
Angle (σ_a)	0.1°
Frequency (σ_f)	0.01 Hz

Table 1: Summarized standard deviations

From Fig. 4 one can confirm a predictable conclusion about Doppler-shift measurements: with the most favorable conditions to use Doppler information (pass with high elevation and frequency mea-

$$\underset{\boldsymbol{r}}{\text{minimize}} \quad \sum_{k=2}^{N} \left(\Delta f^k + \frac{f_{emm}}{c} \frac{-r^{k+2} + 8r^{k+1} - 8r^{k-1} + r^{k-2}}{12\Delta t} \right)^2. \tag{25}$$

surements with small added noise) but just one station and a satellite at a time and no complementary functions to regularize the position estimates, none of the relaxations found could improve the position estimates significantly.

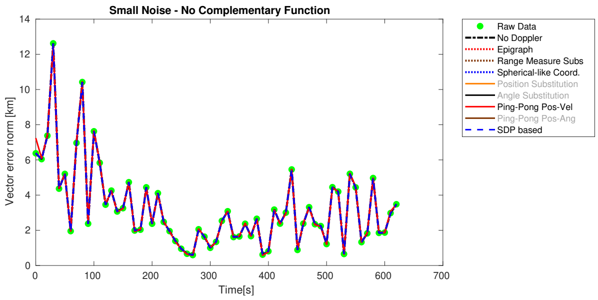

Figure 4: Simulation 1 - Small Noise - No Complementary Function - Comparing best performing cost functions without additional terms to regularize position estimates.

Each of the remaining possible relaxations lead to an estimation worse than the raw output of the sensors.

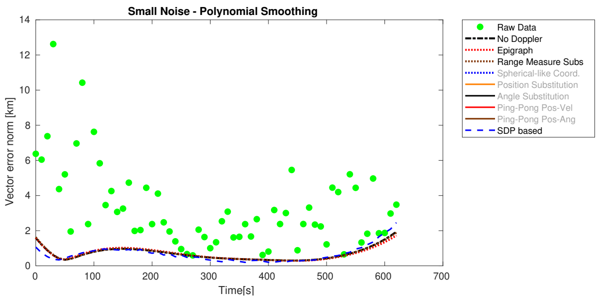

Figure 5: Simulation 2 - Small Noise - Polynomial Smoothing - Comparing best performing cost functions with additional polynomial terms to penalize non-smoothness of estimates.

Once more, the best performing relaxations, aided by the regularization term, do not significantly outperform the equivalent optimization problem without Doppler information. However, the introduction of the smoothing requirement presents a major improvement from the previous case since now there is a correlation between positions which better matches the reality (compared to the case without correlation).

9. Conclusions

A preliminary theoretical study strongly suggests that the use of approximate functions for the Doppler-Shift cost function introduces additional information to the estimation problem. However, this study is not sufficient to conclude about the tightness of the developed relaxations which may degrade the estimator performance. In fact, this problem was verified numerically, since by either approximating the Doppler-Shift cost function using discrete derivatives or analytical derivatives based methods, there was no significant gain of information, and so Doppler relaxations are not expected to provide major improvements in the accuracy of position estimates when compared to the case with no frequency information . This study assumes that both range and angular data are always available, and if that is not the case the above conclusion on the marginal contribution of Doppler may not hold. Also, the ultimate test to find how much the Doppler-shift measurements could improve the estimation would be to solve a non-linear optimization problem with the, non-convex, original measurement model. But in that case, at each instant, six variables would have to be estimated from only three measurements which is not feasible.
Nevertheless, this analysis was made with only two agents at a time, so the conclusion may not hold when several stations are observing a single satellite at the same time while sharing their information.
In the case with little information studied in this thesis, the alternative found to circumvent the lack of redundant information was to introduce complementary functions based on the known properties of satellite dynamics. Interestingly, the best results were obtained with the simplest ones, namely, the polynomial smoothing and dynamics relaxation.
For a case with no outliers and for an adequate weighting factor, the introduction of a polynomial to regularize the position estimates is beneficial. However, there are hyper-parameters such as the order of the polynomial which are dependent on the observation window length which introduces a processing overhead when one has to employ grid searching methods to obtain a guess about the best order and the best weight to apply. One possible solution to overcome these limitations is to use splines that enable the use of an assemblage of smaller order polynomials in smaller partitions of the observation window.
Finally, there is an issue transversal to every optimization problem formulation, not for this particular case, but in a general sense. The problem is that solving convex optimization problems does not generally produce the true best solution for the original (non-convex) problem. Instead, it produces the best solution for the mathematical formulation that was developed to encode that question. This appears in practice when one has to decide what are the best weighting factors to apply to each term of the global cost function, in the case where the global cost function is composed of the sum of convex functions. This poses an additional challenge of firstly determining these hyper-parameters. In this case, the parameters were found by grid search methods,

but these are cumbersome. Moreover, these hyperparameters are generally dependent on the intrinsic characteristics of the input data which compounds the problem even further since it requires a new grid search for different noise levels.

10. Future Work

A more promising idea is to investigate if any of the Doppler-shift cost function relaxations presented in this thesis is useful in the case where several stations observe the satellite at the same time and share the measurements so that some information redundancy is obtained.

Beyond studying a scenario with more ground-stations, it would be valuable to study a scenario with a formation of CubeSats (since this is the way they are normally used) in which each one can perform the same kind of measurements used in this thesis and also communicate to determine their relative position. This problem would be much more close to a sensor network optimization problem for which there is extensive literature published.

References

[1] A. Almeida. A study on probabilistic satellite localization using messages metadata. Master's thesis, IST, Universidade de Lisboa, 2019.

[2] S. C. Chapra and R. P. Canale. *Numerical methods for engineers*, volume 33. McGraw-Hill Science/Engineering/Math, 1991.

[3] M. C. Dykstra. Single Station Doppler Tracking for Satellite Orbit Prediction and Propagation. Master's thesis, University of Texas, 2015.

[4] E. Golton. The use of the Doppler effect to deduce an accurate position for an artificial earth satellite. *Planetary and Space Science*, 9(10):607–623, 1962.

[5] R. Hart. Single-station Tracking for Orbit Determination of Small Satellites. In *Technical Session VI: Small Satellites - Support Systems (Small Satellite Conference)*, 1987.

[6] I. G. Izsak. Orbit determination from simultaneous doppler-shift measurements. *SAO Special Report*, 38, 1960.

[7] A. Jäggi, U. Hugentobler, and G. Beutler. Pseudo-stochastic orbit modeling techniques for low-Earth orbiters. *Journal of Geodesy*, 80(1):47–60, 2006.

[8] S. M. Kay. *Fundamentals of statistical signal processing.* Prentice Hall PTR, 1993.

[9] Z. Q. Luo, W. K. Ma, A. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[10] O. Montenbruck, E. Gill, and F. Lutze. *Satellite Orbits: Models, Methods, and Applications.* Springer-Verlag, 2000.

[11] M. Naeimi and J. Flury. *Global Gravity Field Modeling from Satellite-to-Satellite Tracking Data*. Springer, 2017.

[12] H. Nurminen, L. Suomalainen, S. Ali-Löytty, and R. Piché. 3D Angle-of-Arrival Positioning Using von Mises-Fisher Distribution. In *21st International Conference on Information Fusion, FUSION 2018*, 2018.

[13] R. B. Patton. Orbit determination from single pass doppler observations. *IRE Transactions on Military Electronics*, MIL-4(2/3):336–344, 1960.

[14] Y. Shen Du, P. Wei, W. Chun Li, and H. Shu Liao. Doppler-shift based target localization using semidefinite relaxation. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E97-A(1):397–400, 2014.

[15] C. Soares, J. Xavier, and J. Gomes. Simple and fast convex relaxation method for cooperative localization in sensor networks using range measurements. *IEEE Transactions on Signal Processing*, 63(17):4532–4543, 2015.

[16] F. Valdeira. Hybrid Localization in Underwater Environment. Master's thesis, IST, Universidade de Lisboa, 2018.

[17] D. A. Vallado and W. D. Mcclain. *Fundamentals of Astrodynamics and Applications, Fourth Edition.* Microcosm Press, 2013.

[18] K. Yang, G. Wang, and Z. Q. Luo. Efficient convex relaxation methods for robust target localization by a sensor network using time differences of arrivals. *IEEE Transactions on Signal Processing*, 57(7):2775–2784, 2009.