Confidence region of least squares solution for single-arc observations

Gennaro Principe
Astronautics Research Group, University of Southampton, UK
Roberto Armellin
Departamento de Matematica y Computacion, Universidad de La Rioja, Spain
Hugh Lewis
Astronautics Research Group, University of Southampton, UK

ABSTRACT

The total number of active satellites, rocket bodies, and debris larger than 10 cm is currently about $20,000$. Considering all resident space objects larger than 1 cm this rises to an estimated minimum of $500,000$ objects. Latest generation sensor networks will be able to detect small-size objects, producing millions of observations per day. Due to observability constraints it is likely that long gaps between observations will occur for small objects. This requires to determine the space object (SO) orbit and to accurately describe the associated uncertainty when observations are acquired on a single arc. The aim of this work is to revisit the classical least squares method taking advantage of the high order Taylor expansions enabled by differential algebra. In particular, the high order expansion of the residuals with respect to the state is used to implement an arbitrary order least squares solver, avoiding the typical approximations of differential correction methods. In addition, the same expansions are used to accurately characterize the confidence region of the solution, going beyond the classical Gaussian distributions. The properties and performances of the proposed method are discussed using optical observations of objects in LEO, HEO, and GEO.

1. INTRODUCTION

Since the era of space exploration started, the number of space objects (SO) orbiting the Earth has on average grown. This population includes active spacecraft and space debris [1]. The dominating contributors to the evolution and stability of the space debris environment are fragmentation processes, e.g. explosions and collisions. Because of their frequency and the impact these phenomena have on debris creation, the space debris population is expected to grow with a cascade effect (already predicted some 40 years ago, [2]). Due to very high speed in orbit, even relatively small pieces can damage satellites in a collision resulting in the need of monitoring even small fragments.

Identifying observations belonging to the same object will be one of the main challenges of orbit determination. One requirement to perform reliable data association is to have realistic uncertainties for initial orbit solutions, which could also be used to initialize Bayesian estimators for orbit refinement [3]. Meeting this requirement demands two main steps: solve the orbit determination (OD) problem, and accurately describe the associated confidence region.

In the OD process it is customary to distinguish between the initial orbit determination (IOD) and the accurate orbit estimation (AOE). The former is used for the computation of six orbital elements from six observations with no a priori knowledge of the spacecraft orbit. The AOE, on the other hand, is used for the improvement of a priori orbital elements from a large set of tracking data [4]. When an object is detected for the first time it is highly probable (depending on the observation strategy) that more than six observations on a very short orbital arc are taken. Thus, more observations than those required for IOD are available, but their distribution along the orbit is not the typical one of AOE (i.e. observations spread over several orbital revolutions).

One approach to deal with very short arcs (in the case of optical observations) is based on attributable set and admissible regions [5]. An optical attributable is made of two angles (e.g. right ascension and declination) and their angular rates $\mathscr{A} = \left(\alpha, \delta, \dot{\alpha}, \dot{\delta}\right)$ and it is computed by fitting a polynomial through observations of a very short arc (defining a so called tracklet). Independently from the number of observations acquired for a newly detected object,

only four quantities are retained in the attributable. As a result, the orbit is undetermined in the range ρ, range-rate $\dot{\rho}$ space. The admissible region lies the $2D$ plane generated by the two degrees of freedom of the attributable and its boundaries are defined by some physical constraints such as the energy and minimum/maximum distance from the Earth. This concept has been proposed to correlate optical observations of both SO orbiting the Earth as well as asteroids [6, 7]. Recently, extensions of the admissible region concept have been proposed, in particular by endowing it with statistical properties [8, 9, 10]. This resulted in the advantage of allowing for the inclusion of uncertainties in observations, measurements and timing.

In this work we take a different path. Instead of using attributable and admissible regions we attempt at solving a least squares (LS) problem, in which all the observations belonging to a same tracklet are used. The process is started with an initial guess provided by an IOD solver that uses the observations at the ends and in the middle of the observation arc. An arbitrary order LS solver is implemented to take fully advantages of differential algebra (DA) techniques [11, 12, 13]. DA supplies the tools to compute the derivatives of functions within a computer environment. More specifically, by substituting the classical implementation of real algebra with the implementation of a new algebra of Taylor polynomials, any function f of v variables is expanded into its Taylor polynomial up to an arbitrary order k with limited computational effort. In addition to basic algebraic operations, operations for differentiation and integration can be easily introduced in the algebra, thus finalizing the definition of the differential algebra structure of DA [14]. Similarly to algorithms for floating point arithmetic, methods to perform composition of functions, to invert them, to solve nonlinear systems explicitly, and to treat common elementary functions are available in DA [15]. By using DA we are able to implement an arbitrary order LS solver (referred to in the reminder of the paper as differential algebra least squares (DALS)) and more importantly we are able to approximate the target function as an arbitrary order polynomial. This allows us to define the confidence region of the solution by retaining high order terms typically neglected by linearised theories. These linearised theories imply ellipsoidal confidence regions and Gaussian statistics. Evaluating high order terms represents a key point of our study as an accurate representation of the confidence region in the initial phase of OD is crucial for observation correlation and initialization of Bayesian estimators.

In summary in this work we focus our attention on the OD of SO observed on short arcs with optical instruments with following goals: 1) studying the convergence properties of DALS, 2) characterizing the confidence region of the LS solution by using high order Taylor polynomials. The paper is organized as follows. First, an overview on the classical method of least squares and the representation of the confidence region of its solution is given. The description of the DA implementation of the least squares solver is presented next, followed by the description of a set of algorithms useful to describe the confidence region of the solution in a full nonlinear way. The properties and performances of the proposed algorithms are shown using the orbit determination of four objects in different orbital regimes as test cases. Some final remarks conclude the paper.

2. CLASSICAL LEAST SQUARES

The basic formulation of orbit determination as a nonlinear least squares problem is summarized. First, we introduce the LS problem set-up and classical methods employed for its solution and then we describe the uncertainty of the result as confidence ellipsoids.

2.1 Problem Formulation and Solution

The goal is to determine the orbit of an object given some noisy observations. The orbit is described in terms of n x 1 state vector $\boldsymbol{x}$ at a reference epoch t_0. The state vector can be expressed in different ways, for example by the position vector $\boldsymbol{r}$ and the velocity vector $\boldsymbol{v}$ in Earth-centered inertial (ECI) reference frame.

The standard approach to the OD is based on the LS method, devised by Gauss [16]. Starting from a tentative value of $\boldsymbol{x} = \boldsymbol{x}(t_0)$, the observations at each observation epoch are computed. $\boldsymbol{z_c}$ is a m x 1 vector that contains the computed observations, $\boldsymbol{z_c} = \boldsymbol{h}(\boldsymbol{x})$, with m being the number of measurements. Note that, as this paper focuses on short-arc OD, the number of measurements is (relatively) small. The nonlinear function $\boldsymbol{h}$ includes:

- propagation from t_0 to observation epochs;
- conversion of the state vector into measurements.

Due to sensor noise, the observations $\boldsymbol{z_c}$ differ from the actual ones, $\boldsymbol{z}$ (a vector with the same dimension as $\boldsymbol{z_c}$): the differences are called *residuals*.

Let $\boldsymbol{\xi} = \boldsymbol{z} - \boldsymbol{z}_c$ be the m x 1 vector containing the residuals. The LS solution is that state that minimize the target function

$$J(\boldsymbol{x}) = \boldsymbol{\xi}^T(\boldsymbol{x})\boldsymbol{\xi}(\boldsymbol{x}) \tag{1}$$

To find the minimum, stationary points of $J(\boldsymbol{x})$ need to be found, i.e.

$$\frac{\partial J}{\partial \boldsymbol{x}} = \boldsymbol{0} \tag{2}$$

It can be proved that the LS estimate is equivalent to a maximum likelihood estimate [17] when the noise of the observations is Gaussian[18, 19].

Two main difficulties arise in solving Eq. (2) [20]:

- It represents a system of nonlinear equations, generally without an explicit solution. As a result an iterative method is needed;
- A stationary point of J can be a maximum or a saddle as well. Thus, it is required that the Hessian of the target function, $H = \frac{\partial^2 J}{\partial \boldsymbol{x}^2}$, is definite positive.

A system of nonlinear equations can be solved by using Newton's method. The method will converge given an appropriate choice of the initial estimate used to start the process. This is generally provided by the solution of the IOD problem, i.e. when the minimum number of observations is available ($m = n$). At a generic step i, the solution $\boldsymbol{x}_i$ is available (or the IOD solution, when $i = 1$). The gradient of $J(\boldsymbol{x})$ can be expanded at first order around $\boldsymbol{x}_i$.

$$\frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}) \approx \frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}_i) + \frac{\partial^2 J}{\partial \boldsymbol{x}^2}(\boldsymbol{x}_i)(\boldsymbol{x} - \boldsymbol{x}_i) \tag{3}$$

Then, the solution at the subsequent step is found by equating the previous expansion to 0:

$$\boldsymbol{0} = \frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}_i) + \frac{\partial^2 J}{\partial \boldsymbol{x}^2}(\boldsymbol{x}_i)(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i),$$

where the terms on the right side can be computed as follows:

- From the value of the solution $\boldsymbol{x}_i$, it is possible to compute the residuals $\boldsymbol{\xi}(\boldsymbol{x}_i)$ and then

$$\frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}_i) = 2\boldsymbol{\xi}^T(\boldsymbol{x}_i)F \tag{4}$$

  where F is a m x n matrix with the partial derivatives of the residuals with respect to the state vector components,

$$F = \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{x}}(\boldsymbol{x}_i) \tag{5}$$

  For methods used to correctly compute matrix F (called the design matrix) see [21];

- For the second term,

$$\frac{\partial^2 J}{\partial \boldsymbol{x}^2}(\boldsymbol{x}_i) = H = 2(F^T F + \boldsymbol{\xi}^T(\boldsymbol{x}_i)S) = 2C \tag{6}$$

  where

$$S = \frac{\partial^2 \boldsymbol{\xi}}{\partial \boldsymbol{x}^2}(\boldsymbol{x}_i) \tag{7}$$

  is a three-index array of shape m x n x n, while

$$C = F^T F + \boldsymbol{\xi}^T(\boldsymbol{x}_i)S \tag{8}$$

  is a n x n matrix, called normal matrix.

  In conclusion, from Eq. (2.1), the solution of the iterative method is

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - C^{-1}F^T\boldsymbol{\xi} \tag{9}$$

Due to practical problems in computing the second derivatives in the matrix S, the full Newton's method is typically avoided for the OD problem [22]. The quantity $\boldsymbol{\xi}^T S$ in the expression of C is often neglected (as it contains the residuals this terms is small when the residuals are small), leading to the so called differential correction technique [20], a variant of Newton's method. As the differential correction does not compute the full Hessian of the cost function it is impossible to establish whether the convergence point is a local minimum (that is, the Hessian matrix H is positive semidefinite).

In LS formulation often the residuals are weighted to take into account sensors with different accuracies. Let ξ_i' be the true residual and ξ_i the normalized one:

$$\xi_i = \frac{\xi_i'}{\sigma_i},$$

with σ_i being the standard deviation of the sensors. It is essential to highlight that a uniform weight does not matter in the solution, although it matters when describing the confidence region of the solution [20, 21].

2.2 Classical Confidence Region

The solution of the LS $\boldsymbol{x}^*$ is the value of the state vector that minimizes the cost function, $J(\boldsymbol{x}^*) = J^*$. However, $\boldsymbol{x}^*$ does not represent the true orbit, which lies within a confidence region. In order to outline this region, let us consider the value of the target function J in a neighbourhood of $\boldsymbol{x}^*$,

$$J(\boldsymbol{x}) = J^* + \delta J(\boldsymbol{x}) \tag{10}$$

in which $\delta J(\boldsymbol{x})$ is called the penalty. (Throughout the paper δJ will be used to indicate the functional expression of the penalty, whereas ΔJ will be used when the penalty assumes or it is assigned a numerical value.)

The confidence region of the solution is defined as the region in which δJ is less than a fixed value (also called the control value) K^2. In terms of the penalty function the confidence region $Z(K)$ is defined as

$$Z(K) = \{\boldsymbol{x} \in \mathbb{R}^n : \delta J(\boldsymbol{x}) \leq K^2\}. \tag{11}$$

An expression for $\delta J(\boldsymbol{x})$ can be obtained by expanding $J(\boldsymbol{x})$ around $\boldsymbol{x}^*$ and neglecting terms of order ≥ 3 in $\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^*$:

$$J(\boldsymbol{x}) \approx J^* + \frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}^*)^T \delta\boldsymbol{x} + \frac{1}{2}\delta\boldsymbol{x}^T \frac{\partial^2 J}{\partial \boldsymbol{x}^2}(\boldsymbol{x}^*)\delta\boldsymbol{x} = J^* + \frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}^*)^T \delta\boldsymbol{x} + \frac{1}{2}\delta\boldsymbol{x}^T H \delta\boldsymbol{x}, \tag{12}$$

where $H = \frac{\partial^2 J}{\partial \boldsymbol{x}^2}(\boldsymbol{x}^*)$. As in the LS solution $\boldsymbol{x}^*$ the gradient of J is zero, $\frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x}^*) = \boldsymbol{0}$, it follows

$$J(\boldsymbol{x}) = J^* + \frac{1}{2}\delta\boldsymbol{x}^T H \delta\boldsymbol{x}. \tag{13}$$

By equating Eq. (10) and (13), the confidence region definition becomes

$$\delta J(\boldsymbol{x}) \approx \frac{1}{2}\delta\boldsymbol{x}^T H \delta\boldsymbol{x} = \delta\boldsymbol{x}^T C \delta\boldsymbol{x} \leq K^2. \tag{14}$$

Equation (14) represents the classical expression of the confidence region, with the Taylor expansion of the function $J(\boldsymbol{x})$ limited to 2-nd order. As the penalty has a quadratic form, the confidence region is described by an ellipsoid, whose axes can be determined by the eigenvector decomposition of the normal matrix (or its inverse, the covariance matrix).

The LS method can be given a probabilistic interpretation. If the error of each observation is an independent random variable with normal distribution and zero mean, the solution of a LS problem is a random variable with a multivariate gaussian probability density function (pdf) $p_{\boldsymbol{x}}(\boldsymbol{x})$. In particular, $\boldsymbol{x}^*$ is the mean of the Gaussian distribution and the covariance matrix is the inverse of the normal matrix $P = C^{-1}$ [16]. The LS solution is statistically described by

$$p_{\boldsymbol{x}}(\boldsymbol{x}) = \frac{\sqrt{|C|}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^T C(\boldsymbol{x}-\boldsymbol{x}^*)} = \frac{\sqrt{|C|}}{(2\pi)^{n/2}} e^{-\frac{1}{2}\delta J(\boldsymbol{x}^*)}. \tag{15}$$

The contour levels of the penalty function are ellipsoids of equal probability and a certain percent confidence region of the LS solution is obtained by properly selecting the value of the control value K^2.

3. DIFFERENTIAL ALGEBRA LEAST SQUARES

In this section we present a high order iterative procedure to solve the LS problem together with algorithms to describe the confidence region of the solution when the orders above the second are retained in the representation of the penalty function.

3.1 Solution of LS Problem

We start by describing a general algorithm to find the solution of a system of nonlinear equations in DA formalism. The aim is to solve $\boldsymbol{g}(\boldsymbol{x}) = 0$. At a general step i:

1. Given the solution $\boldsymbol{x}_i$ (from the previous step, or from the initial guess when $i = 1$), initialize the components of the state vector $\boldsymbol{x}_i$ as k-th order DA variables:

$$\boldsymbol{x}_i = \boldsymbol{x}_i + \delta\boldsymbol{x}_i;$$

2. The evaluation of $\boldsymbol{g}$ in DA framework delivers its k-th order Taylor expansion around $\boldsymbol{x}_i$, $\mathcal{T}_{\boldsymbol{g}}^k(\delta\boldsymbol{x}_i)$. Thus, $\boldsymbol{g}$ will be the sum of a constant part $\boldsymbol{g}_i$ (given by the solution of the last step, $\boldsymbol{g}(\boldsymbol{x}_i) = \boldsymbol{g}_i$) and an origin-preserving Taylor polynomial $\delta\boldsymbol{g}$, function of the DA variables $\delta\boldsymbol{x}_i$

$$\boldsymbol{g}(\boldsymbol{x}) \approx \mathcal{T}_{\boldsymbol{g}}^k(\delta\boldsymbol{x}_i) = \boldsymbol{g}(\boldsymbol{x}_i) + \delta\boldsymbol{g}(\delta\boldsymbol{x}_i) = \boldsymbol{g}_i + \mathcal{T}_{\delta\boldsymbol{g}}^k(\delta\boldsymbol{x}_i) \tag{16}$$

   The following direct map is available

$$\delta\boldsymbol{g} \approx \mathcal{T}_{\delta\boldsymbol{g}}^k(\delta\boldsymbol{x}_i); \tag{17}$$

3. Invert the map of Eq. (17), obtaining

$$\delta\boldsymbol{x}_i \approx \mathcal{T}_{\delta\boldsymbol{x}_i}^k(\delta\boldsymbol{g}); \tag{18}$$

4. Evaluate the inverse map in $-\boldsymbol{g}_i$ to compute the correction $\Delta\boldsymbol{x}_i$

$$\Delta\boldsymbol{x}_i = \mathcal{T}_{\delta\boldsymbol{x}_i}^k(-\boldsymbol{g}_i)$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta\boldsymbol{x}_i$$

5. Iterate until a convergence criterion is met or the maximum number of iterations is reached.

After convergence, the algorithm provides the solution of the set of nonlinear equations as well as the Taylor expansion of the function $\boldsymbol{g}(\boldsymbol{x})$ at high order around the solution $\boldsymbol{x}^*$.

The solution of the LS problem requires finding the stationary point of the cost function $J(\boldsymbol{x})$. If in the previous algorithm we set $\boldsymbol{g}(\boldsymbol{x}) = \frac{\partial J}{\partial \boldsymbol{x}}(\boldsymbol{x})$, then an arbitrary order solver of the LS problem is obtained, i.e. the DALS solver. The DALS solver has two main advantages with respect to the classical differential correction:

- As the objective function $J(\boldsymbol{x})$ is expanded to an arbitrary order, we have the correct (full) expression of the Hessian matrix H. This can be used to check whether the stationary point is actually a minimum;
- The polynomial representation of the objective function $J(\boldsymbol{x})$ allows us to analytically represent it in a neighbourhood of the minimum. This feature enables the nonlinear representation of the solution confidence region.

Like for all the iterative procedures, a convergence criterion needs to be defined. In our implementation we use two convergence criteria: the first one is based on the size of the correction $\Delta\boldsymbol{x}$, while the second is based on the change of the target function J. Thus, the iterative process is terminated when one of the two following conditions is met:

$$\begin{aligned} &\|\Delta\boldsymbol{x}\|_\infty \leq \varepsilon_x \\ &\Delta J \leq \varepsilon_J \end{aligned} \tag{19}$$

where ε_x and ε_J are established tolerances.

Although the algorithm presented in this section works at arbitrary order, it has been noticed that the inclusion of terms above the second does not improve the convergence while significantly increasing the execution time. Thus, a second order DALS solver is used in this work.

3.2 Confidence Region Representation

When considering short observational arcs, terms above second order in the expression of $\delta J(\boldsymbol{x})$ are in general non negligible. The contour levels of $\delta J(\boldsymbol{x})$ still characterise solutions with equal probability but, due to the presence of high order terms, the solution statistics are no longer Gaussian and surfaces of equal probability are no longer ellipsoids. In this section we present two algorithms to describe the confidence region of the LS solution taking full advantage of the high order representation of $\delta J(\boldsymbol{x})$ enabled by DA.

After convergence, DALS provides the k-th order Taylor expansion of $J(\boldsymbol{x})$ around the optimal solution

$$J(\delta\boldsymbol{x}^*) \approx J(\boldsymbol{x}^*) + \mathscr{T}^k_{\delta J}(\delta\boldsymbol{x}^*) \tag{20}$$

$$\delta J(\delta\boldsymbol{x}^*) = \mathscr{T}^k_{\delta J}(\delta\boldsymbol{x}^*), \tag{21}$$

where $\delta\boldsymbol{x}^* = \boldsymbol{x} - \boldsymbol{x}^*$. In Eq. (12) terms up to order k are retained. Given an arbitrary direction and a confidence level, the boundaries of the confidence region along this direction can be found by exploiting the polynomial expression of $J(\boldsymbol{x})$. The availability of algorithms to compute such boundaries is of fundamental importance, for example when samples of equal probability need to be drawn to correlate observations or to initialize a particle filter [23]. In such cases a set of random directions may be generated and points with equal probability along these directions may be computed. Two algorithms are presented. The first method, referred to as state-dependent quadratic approximation along fixed direction (SDQA-FIXED), is based on a state-dependent quadratic approximation of J in the proximity of $\boldsymbol{x}^*$. In contrast, the arbitrary order fixed direction (AO-FIXED) uses the polynomial representation of J and a nonlinear solver to compute the boundaries of the confidence region along the desired direction.

SDQA-FIXED. From the expression of $\mathscr{T}^k_{\delta J}(\delta\boldsymbol{x}^*)$ provided by DALS it is possible to compute the gradient $\frac{\partial \mathscr{T}^k_{\delta J}}{\partial \boldsymbol{x}}$ and the Hessian matrix $\frac{\partial^2 \mathscr{T}^k_{\delta J}}{\partial \boldsymbol{x}^2}$ at a generic point $\boldsymbol{x}_i$ in the neighbourhood of $\boldsymbol{x}^*$. A quadratic approximation of δJ around $\boldsymbol{x}_i$ is (let $\delta\boldsymbol{x}_i = \boldsymbol{x} - \boldsymbol{x}_i$)

$$\delta J(\delta\boldsymbol{x}_i) \approx \frac{\partial \mathscr{T}^k_{\delta J}}{\partial \boldsymbol{x}}\bigg|^T_{x_i - x^*} \delta\boldsymbol{x}_i + \frac{1}{2}\delta\boldsymbol{x}_i^T \frac{\partial^2 \mathscr{T}^k_{\delta J}}{\partial \boldsymbol{x}^2}\bigg|_{x_i - x^*} \delta\boldsymbol{x}_i = \boldsymbol{g}(\boldsymbol{x}_i)^T \delta\boldsymbol{x}_i + \delta\boldsymbol{x}_i^T C(\boldsymbol{x}_i)\delta\boldsymbol{x}_i \tag{22}$$

As both $\boldsymbol{g}$ and C are functions of the point at which they are computed this formulation is referred to as state-dependent quadratic form. The confidence region is given by

$$\boldsymbol{g}(\boldsymbol{x}_i)^T \delta\boldsymbol{x}_i + \delta\boldsymbol{x}_i^T C(\boldsymbol{x}_i)\delta\boldsymbol{x}_i \leq K^2 - (J(\boldsymbol{x}_i) - J(\boldsymbol{x}^*)) \tag{23}$$

where the term $\Delta J(\boldsymbol{x}_i) = J(\boldsymbol{x}_i) - J(\boldsymbol{x}^*)$ takes into account the presence of a penalty in $\boldsymbol{x}_i$ with respect to $\boldsymbol{x}^*$. According to Eq. (23), it is possible to use the quadratic approximation to evaluate the confidence region at a generic point $\boldsymbol{x}_i$ and not only in the solution $\boldsymbol{x}^*$. When higher order terms are considered the confidence region provided by the state-dependent quadratic approximation changes with $\boldsymbol{x}_i$. The following algorithm has this purpose: find the points along a given direction $\boldsymbol{v}$ (i.e. a n dimensional unit vector) such that $\Delta J = K^2$.

1. Initialize the starting point with the solution of LS, i.e. $\boldsymbol{x}_i = \boldsymbol{x}^*$;
2. Compute, at $\boldsymbol{x}_i$, the gradient $\boldsymbol{g}(\boldsymbol{x}_i)$ and the normal matrix $C(\boldsymbol{x}_i)$. At the first step $\boldsymbol{g}(\boldsymbol{x}_i) = \boldsymbol{g}(\boldsymbol{x}^*) = 0$, but it will be non-zero in successive steps;
3. Let $\delta\boldsymbol{x}_i = \alpha\boldsymbol{v}$. The Taylor expansion of $J(\boldsymbol{x})$ at 2-nd order around $\boldsymbol{x}_i$ gives
$$\delta J(\delta\boldsymbol{x}_i) \approx \boldsymbol{g}(\boldsymbol{x}_i)^T \alpha\boldsymbol{v} + \alpha\boldsymbol{v}^T C(\boldsymbol{x}_i)\alpha\boldsymbol{v} = K^2 - (J(\boldsymbol{x}_i) - J(\boldsymbol{x}^*)); \tag{24}$$
4. Transform $\boldsymbol{v} = V\tilde{\boldsymbol{v}}$, where V is the matrix whose columns are the eigenvectors of the normal matrix at $\boldsymbol{x}_i$, to get
$$\boldsymbol{g}(\boldsymbol{x}_i)^T V\alpha\tilde{\boldsymbol{v}} + \alpha\tilde{\boldsymbol{v}}^T V^T C(\boldsymbol{x}_i)V\alpha\tilde{\boldsymbol{v}} = K^2 - (J(\boldsymbol{x}_i) - J(\boldsymbol{x}^*)); \tag{25}$$
5. The matrix $V^T C(\boldsymbol{x}_i)V$ is diagonal, with the eigenvalues $\frac{1}{\gamma_i^2}$, $i = 1 \ldots n$ ([24]). The second term in Eq. (25) can be written as
$$\alpha^2 \sum_{i=1}^{n} \frac{\tilde{v}_i^2}{\gamma_i^2};$$

Eq. (25) becomes

$$\boldsymbol{g}(\boldsymbol{x}_i)^T V\tilde{\boldsymbol{v}}\alpha + \sum_{i=1}^{n} \frac{\tilde{v}_i^2}{\gamma_i^2}\,\alpha^2 = K^2 - (J(\boldsymbol{x}_i) - J(\boldsymbol{x}^*)), \tag{26}$$

which is a quadratic form in α. The equation provides two solutions for α, α_+ and α_-. When $\boldsymbol{x}_i = \boldsymbol{x}^*$, $\boldsymbol{g}(\boldsymbol{x}_i) = 0$ and $J(\boldsymbol{x}_i) = J(\boldsymbol{x}^*)$, then

$$\alpha^2 \sum_{i=1}^{n} \frac{\tilde{v}_i^2}{\gamma_i^2} = K^2 \tag{27}$$

$$\alpha_{+,-} = \pm K \sqrt{\frac{1}{\sum_{i=1}^{n} \frac{\tilde{v}_i^2}{\gamma_i^2}}} \tag{28}$$

The two solutions are symmetric with respect to $\boldsymbol{x}^*$ (due to the quadratic approximation). In contrast, when $\boldsymbol{x}_i \neq \boldsymbol{x}^*$, $\alpha_+ \neq \alpha_-$ due to the gradient term $\boldsymbol{g}(\boldsymbol{x}_i)^T V\tilde{\boldsymbol{v}}\alpha$. After solving the Eq. (26), the current solution is updated as $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_{+,-}\boldsymbol{v}$, in which only one of two solutions is chosen.

6. Evaluate $\Delta J(\boldsymbol{x}_{i+1}) = J(\boldsymbol{x}_{i+1}) - J(\boldsymbol{x}^*)$ with the k-order polynomial provided by DALS (Eq. (21)). If $\Delta J \neq K^2$, steps 2-5 are repeated. It is worth noting that, if δJ is expanded at second order in the classical way, the process stops after a single iteration, because $\Delta J \neq K^2$ directly at first iteration.

AO-FIXED. Given an arbitrary direction and confidence level the boundaries of the confidence region can be computed directly with the polynomial representation of the penalty function, i.e. without resorting to the state-dependent formulation. The algorithm can be summarized as follows.

1. Start by approximating δJ in $\boldsymbol{x}^*$ as a 2-nd order polynomial. Because $\boldsymbol{g}(\boldsymbol{x}^*) = 0$,

$$\delta J \approx \delta\boldsymbol{x}^T C\delta\boldsymbol{x}, \tag{29}$$

where $\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^*$;

2. Evaluate $\delta\boldsymbol{x}$ along $\boldsymbol{v}$:

$$\delta\boldsymbol{x} = \alpha^{(2)}\boldsymbol{v}. \tag{30}$$

$\alpha^{(2)}$ is the magnitude of the displacement along $\boldsymbol{v}$. At second order the boundary of the confidence region along $\boldsymbol{v}$ is given by

$$\delta J \approx \alpha^{(2)}\boldsymbol{v}^T C\boldsymbol{v}\alpha^{(2)} = K^2; \tag{31}$$

3. Solve Eq. (31) for $\alpha^{(2)}$

$$\alpha_{+,-}^{(2)} = \pm K\sqrt{\frac{1}{\boldsymbol{v}^T C\boldsymbol{v}}} \tag{32}$$

$\alpha^{(2)}\boldsymbol{v}$ is the point along $\boldsymbol{v}$ such that $\Delta J(\delta\boldsymbol{x}) = \Delta J(\alpha^{(2)}\boldsymbol{v}) = K^2$. There are two solutions, symmetric with respect to $\boldsymbol{x}^*$ (due to the quadratic approximation).

4. By using $\alpha_{+,-}^{(2)}$ as the initial condition, numerically compute the points, along the direction $\boldsymbol{v}$, in which the k-th order approximation of δJ is K^2. This requires the solution of a k-th order polynomial equation in α, $\mathscr{T}_{\delta J}^k(\alpha\boldsymbol{v}) = K^2$, which can be easily obtained by a nonlinear solver using both $\alpha_{+,-}^{(2)}$ as initial guesses. We refer to the final solution as $\alpha_{+,-}^{(k)}$.

Both SDQA-FIXED and AO-FIXED deliver the same values of the boundaries of the confidence region along a fixed direction when a same value of K^2 is used. The value K^2 defines the confidence region and it must be chosen based on the percent confidence level we want to draw. According to the F-test method [25] the value of K^2 for confidence level of $100(1-\alpha)\%$ can be estimated by

$$\delta J(\boldsymbol{x}) \leq \frac{n}{m-n} J^* F_{n,m-n}^{\alpha} = K^2, \tag{33}$$

in which $F_{n,m-n}^{\alpha}$ is the upper α percentage point of the F-distribution.

3.3 Evolution of the Weak Direction

The confidence region associated to the LS solution is in general a n dimensional region. However, in many cases of practical interest this region is stretched along one direction, called the weak direction. The weak direction is defined as the predominant direction of uncertainty in an orbit determination problem [26], i.e. it is the direction along which the penalty δJ is less sensitive to variations of the state vector. The determination of the weak direction can be a key element when studying the time evolution of the confidence region. In fact, when this region is highly stretched, it can be approximated as a mono-dimensional set and sampling can be performed along the weak direction, thus drastically reducing the number of samples needed. Note that this concept can be extended to 2-D when the confidence region is significantly elongated along two dimensions.

When a pure quadratic form is used to represent the penalty function, the weak direction is determined by the eigenvector corresponding to minimum eigenvalue of the normal matrix or the maximum eigenvalue of the covariance matrix. However, when high order terms are retained, the weak direction is point dependent. Similarly to the SDQA-FIXED algorithm, the computation of the evolution of the weak direction can be performed by taking advantage of a state-dependent quadratic approximation of the penalty δJ. The algorithm for computing the evolution of the weak direction is referred to as the state-dependent quadratic approximation along the weak direction (SDQA-WEAK), and it can be summarised as follows:

SDQA-WEAK

1. Start from the solution point of LS, $\boldsymbol{x}_i = \boldsymbol{x}^*$. As done in Eq. (22), approximate
$$\delta J(\delta \boldsymbol{x}_i) \approx \boldsymbol{g}(\boldsymbol{x}_i)^T \delta \boldsymbol{x}_i + \delta \boldsymbol{x}_i^T C(\boldsymbol{x}_i) \delta \boldsymbol{x}_i \tag{34}$$
2. Compute $\boldsymbol{g}(\boldsymbol{x}_i)$ and the normal matrix $C(\boldsymbol{x}_i)$. At the first step $\boldsymbol{g}(\boldsymbol{x}_i) = \boldsymbol{g}(\boldsymbol{x}^*) = 0$, but it is non-zero in successive steps;
3. Evaluate eigenvectors and eigenvalues of C. Indicate with V the eigenvector matrix and C_d the diagonal form of C with elements $\frac{1}{\gamma_1^2} \dots \frac{1}{\gamma_n^2}$. Order the eigenvectors such that $\boldsymbol{v}_1$ corresponds to the minimum eigenvalue $\frac{1}{\gamma_1^2}$;
4. Use the transformation $\delta \boldsymbol{x}_i = V \delta \tilde{\boldsymbol{x}}_i$ such that Eq. (34) becomes
$$\delta J(\delta \boldsymbol{x}_i) \approx \boldsymbol{g}(\boldsymbol{x}_i)^T V \delta \tilde{\boldsymbol{x}}_i + \delta \tilde{\boldsymbol{x}}_i^T C_d(\boldsymbol{x}_i) \delta \tilde{\boldsymbol{x}}_i; \tag{35}$$
5. As $V = \left[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\right]$, evaluate $\delta \tilde{\boldsymbol{x}}_i$ along the eigenvector corresponding to the minimum eigenvalue (this eigenvector defines the weak direction),
$$\delta \tilde{\boldsymbol{x}}_i = \alpha \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
where α is the magnitude of the variation. Eq. (35) becomes
$$\delta J(\delta \boldsymbol{x}_i) = \boldsymbol{g}(\boldsymbol{x}_i)^T \boldsymbol{v}_1 \alpha + \frac{\alpha^2}{\gamma_1^2} \tag{36}$$
By solving for α
$$\boldsymbol{g}(\boldsymbol{x}_i)^T \boldsymbol{v}_1 \alpha + \frac{\alpha^2}{\gamma_1^2} = K^2 - (J(\boldsymbol{x}_i) - J^*), \tag{37}$$
the boundary of the uncertainty region along the weak direction is computed;
6. Proceed along $\boldsymbol{v}_1$ taking only a fraction of α to update the point
$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta \boldsymbol{x}_i = \boldsymbol{x}_i + V \delta \tilde{\boldsymbol{x}}_i = \boldsymbol{x}_i + h \alpha \boldsymbol{v}_1 \tag{38}$$
where h is a chosen step size;
7. Compute the value of $\Delta J(\boldsymbol{x}_{i+1}) = J(\boldsymbol{x}_{i+1}) - J(\boldsymbol{x}^*)$ with the k-order polynomial delivered by the DALS solver. Repeat the steps $2-6$, until the value of $\Delta J(\boldsymbol{x}_i)$ exceeds the fixed threshold K^2 that determines the boundary of the confidence region.

4. Test Cases

Four different orbits are used as test cases: a low Earth orbit (LEO), a geostationary Earth orbit (GEO), a geostationary transfer orbit (GTO), and a Molnyia. The orbital parameters and the object's catalogue number (SSC) are reported in Table 1 ($a, e, i, \Omega, \omega, M$ are the classical orbital elements).

Table 1: Test cases, orbital parameters

Orbit type		LEO	GEO	GTO	Molnyia
SSC		04784	26824	23238	40296
Epoch	JED	2457158.4277	2457163.2824	2457167.1008	2457165.0708
a	km	7353.499	42143.781	24628.972	26569.833
e	-	0.002640	0.000226	0.699849	0.723221
i	deg	74.0295	0.0356	3.962	62.794
Ω	deg	179.6401	26.278	315.676	344.538
ω	deg	359.079	42.052	240.885	271.348
M	deg	22.252	72.455	13.735	347.726

For all the objects optical observations (i.e. right ascension and declination) are simulated from Teide observatory, Tenerife, Canary Islands, Spain (observation code 954). Objects in GEO, GTO and Molnyia are observed with a telescope similar to the optical ground station (OGS) with an observation strategy that re-observes the same portion of the sky every 2 minutes. The measurement noise is Gaussian with zero mean and standard deviation $\sigma = 0.5$ arcsec. The object in LEO is observed with a wide field of view camera, with observations taken every 7 seconds and exposure times of 3 seconds. In this case $\sigma = 5$ arcsec. Two scenarios are simulated with 8 or 15 observations. When 8 observations are used, the observational arcs range from 1.93 deg for the Molnyia to 3.51 deg for the GEO; when 15 are considered, the arc length ranges from 3.96 deg for the Molnyia to 7.02 deg for the GEO. Tab. 2 summarizes the observation conditions and Fig. 1 depicts the observation geometries for each test case.

The results discussed in the following sections assume the availability of an initial orbit provided by the solution of an IOD problem. This preliminary solution is computed with a high order algorithm based on the solution of two Lambert's problems between the central epoch and the two ends of the observation arc (for more details the reader can refer to [27]). Finally, it is worth mentioning that Kepler's dynamics is considered throughout this section.

4.1 DALS Convergence Properties

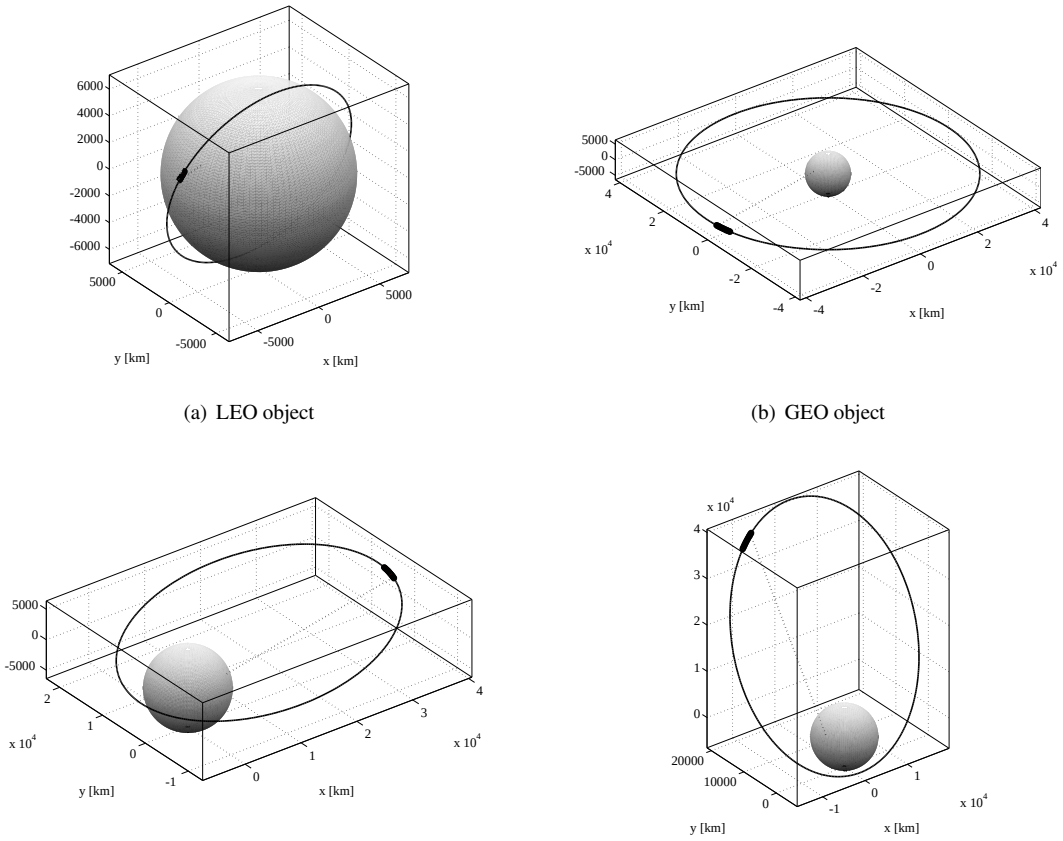
For each test case in Table 2 100 simulations were run in which synthetic observations were generated by adding Gaussian noise to the ideal observations. The DALS solver was used to estimate the orbit at the center of the observation window (at observation #5 for the 8-observation scenario and #8 for the 15-observation one) as this was found to maximize the algorithm performances and robustness. The tolerances ε_x and ε_J were chosen such that convergence was reached when one the following conditions was met

$$||\Delta \boldsymbol{x}||_\infty \leq \begin{cases} 1 \text{ m} & \text{for position} \\ 1 \text{ mm/s} & \text{for velocity} \end{cases} \qquad \Delta J \leq m \left(\frac{\sigma}{100}\right)^2 .$$

The DALS solver converged for 99.72 % of the tests, taking on average 2 only iterations. (Note that algorithm convergence does not provide any information on solution quality, e.g. convergence to a local minimum). In Table 3

Table 2: Test cases, observation window

Test Case	First observation				Δt	$\sigma_{\alpha,\delta}$
	yr	mo	day_0	hr_0	s	arcsec
LEO	2015	MAY	15	22.15	7	5
GEO	2015	MAY	22	21.34	120	0.5
GTO	2015	JUN	02	05.07	120	0.5
Molnyia	2015	MAY	22	22.00	120	0.5



(a) LEO object

(b) GEO object

(c) GTO object

(d) Molnyia object

Figure 1: Observation geometry for the different test cases with 15 observations.

the median of the absolute error (with respect to the reference orbit) in position (km) and velocity (km/s) is reported, for the four test cases and different numbers of observations. On average the estimation errors of DALS solution are lower than those associated to the IOD solution, showing that the inclusion of all the individual observations can improve the orbit estimation even for very short arcs. In addition, the improvement in accuracy of the LS is greater when longer observational arcs is considered. Finally, as one may expect, the median of the errors decreases with the number of observations, i.e. the orbit estimation becomes more accurate for longer observational arcs.

To analyse the solution accuracy without resorting to the (unknown in a real scenario) true solution, the absolute values of the residuals (scaled by the measurements σ) were analyzed. For each test case the median of these absolute values was then computed for the 100 simulations, and the maximum value reported in Table 4. The values reported in the Table 4 are compatible with measurement statistics. Fig. 2 reports the results of the simulations for Molnyia orbit when 8 observations are considered. The statistics of the absolute value of the dimensionless residuals are plotted and compared with the IOD solutions. First, it is worth noting that the residuals of the IOD solutions are almost zero at the 1-st, 5-th, and 8-th observations, i.e. those used for the IOD. This result is obvious as the IOD solutions are deterministic and exactly produce the available observations. On the other hand, the residuals significantly grow at other observation epochs. On average the residual are much smaller and more uniformly distributed when LS solutions are analyzed. From these analysis it can be concluded that the LS solution improves the orbit provided by the IOD even when only few measurements distributed on a very short arc are available

4.2 Confidence Region

The confidence region of the LS solution is represented in modified equinoctial elements (MEE). The state vector $\boldsymbol{x}_{\text{MEE}} = (p, f, g, h, k, L)$, in which $p = a(1-e^2)$ is the semilatus rectum, $f = e\cos(\omega+\Omega)$ and $g = e\sin(\omega+\Omega)$,

Table 3: Median of the absolute value of the error (with respect to the true solution) in position (km) and velocity (km/s), for IOD and DALS solutions

		Number of observations			
		8		15	
		IOD	DALS	IOD	DALS
LEO	Pos	5.832e+01	4.449e+01	1.540e+01	6.050e+00
	Vel	2.569e-01	1.861e-01	7.414e-02	3.035e-02
GEO	Pos	5.756e+02	4.454e+02	1.115e+02	6.856e+01
	Vel	4.443e-02	3.262e-02	8.358e-03	5.407e-03
GTO	Pos	4.758e+02	3.227e+02	9.718e+01	6.576e+01
	Vel	2.680e-02	1.884e-02	5.893e-03	3.393e-03
Molnyia	Pos	4.093e+02	2.564e+02	1.025e+02	5.826e+01
	Vel	2.235e-02	1.442e-02	5.628e-03	3.043e-03

Table 4: Maximum of the median of the absolute values of the normalized residuals.

		Number of observations	
		8	15
LEO	$\xi_\alpha\ [1/\sigma]$	6.661e-01	8.085e-01
	$\xi_\delta\ [1/\sigma]$	6.980e-01	6.922e-01
GEO	$\xi_\alpha\ [1/\sigma]$	6.853e-01	7.066e-01
	$\xi_\delta\ [1/\sigma]$	7.117e-01	7.402e-01
GTO	$\xi_\alpha\ [1/\sigma]$	6.955e-01	7.250e-01
	$\xi_\delta\ [1/\sigma]$	6.790e-01	7.762e-01
Molnyia	$\xi_\alpha\ [1/\sigma]$	6.255e-01	7.177e-01
	$\xi_\delta\ [1/\sigma]$	6.615e-01	7.423e-01

$h = \tan(i/2)\cos(\Omega)$ and $k = \tan(i/2)\sin(\Omega)$, and the true longitude $L = \Omega + \omega + \nu$. MEE were chosen as they absorb part of the nonlinearity of orbital dynamics and thus bring benefits when propagating the confidence region [28, 29].

When the cost function J is approximated with a 2-nd order polynomial, the solution vector is normally distributed and its statistics are fully describe by its mean (the LS solution) and covariance matrix as shown in Eq. (15). The contour surfaces of the penalty function are ellipsoids whose axes are proportional to the square root of the eigenvalues of the covariance matrix.

In Tables 5 and 6 the covariance matrices for the GEO object are reported in the case of 8 and 15 observations respectively. The diagonal terms of the covariance matrix are very large when short arcs are considered. By comparing the two tables it can be appreciated how these terms lower when the observations are spread on a longer arc. This finding is further confirmed when the square root of the maximum eigenvalue of the covariance matrix with 15 observations γ^{15}_{max} is compared to that of 8 observations γ^{8}_{max}. On average for the LEO case $\frac{\gamma^{15}_{max}}{\gamma^{8}_{max}} = 0.1849$, for the GEO $\frac{\gamma^{15}_{max}}{\gamma^{8}_{max}} = 0.1914$, for the GTO $\frac{\gamma^{15}_{max}}{\gamma^{8}_{max}} = 0.2112$ and for the Molnyia $\frac{\gamma^{15}_{max}}{\gamma^{8}_{max}} = 0.1824$.

In Table 7 the covariance matrix for scenario with 8 observations separated by 4 minutes (same observational arc as 15 measures separated by 2 minutes) is reported. This covariance matrix is similar to the one with 15 observations, confirming that the length of the observational arc plays a major role in the definition of the confidence region. In addition, it can be noticed that for arcs of same lengths the confidence region is slightly smaller when more observations are included.

The results shown thus far are relative to a classical second order representation of the penalty δJ. By inspection of the terms above the second order in the expression of $\mathscr{T}^k_{\delta J}(\delta \boldsymbol{x}^*)$ it can be readily understood that for very short arcs higher order terms play a significant role in the definition of the confidence region. This is clearly shown in Figures 3-6 in which the 85, 90 and 95 percent confidence regions computed at different orders are displayed for all the test cases when 8 observations are used. Given the value of n, m, and $J(\boldsymbol{x}^*)$ and desired confidence percent Eq. (33) is used to find the proper values of K^2. The boundaries of the confidence regions are then computed by either SDQA-FIXED

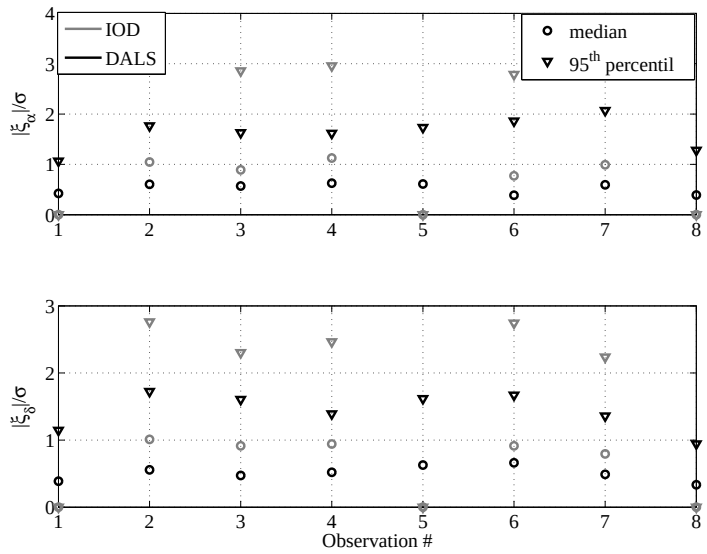


Figure 2: Statistics of the absolute value of the dimensionless residuals on α and δ, for IOD and DALS solutions. The observed object is in Molnyia orbit

Table 5: Covariance matrix in MEE for GEO and 8 observations

p [km]	f	g	h	k	L [rad]
5.0920123e+06	-8.9547143e+01	-1.0650759e+01	1.5536300e-01	-1.2065074e+00	-1.3639190e-01
-8.9547143e+01	1.5748250e-03	1.8675471e-04	-2.7543301e-06	2.1214725e-05	2.3983609e-06
-1.0650759e+01	1.8675471e-04	2.6824098e-05	-1.4105794e-07	2.5457500e-06	2.8691229e-07
1.5536300e-01	-2.7543301e-06	-1.4105794e-07	1.2522745e-08	-3.5873937e-08	-4.0955955e-09
-1.2065074e+00	2.1214725e-05	2.5457500e-06	-3.5873937e-08	2.8598508e-07	3.2324982e-08
-1.3639190e-01	2.3983609e-06	2.8691229e-07	-4.0955955e-09	3.2324982e-08	3.6545165e-09

or the AO-FIXED algorithm along directions in the plane defined by the two eigenvectors associated with the two largest eigenvalues of the covariance matrix. The figures show that, when δJ is expanded at order higher than 2, the confidence region is no longer represented by ellipsoids. It is also worth observing how the shape of confidence region remains essentially unchanged above order 4, thus suggesting order 4 as default expansion order.

As a final result, Figure 7 shows the projection of the samples generated along the weak direction in the $p-L$, $f-g$ and $h-k$ planes together with the true solution. The gray scale is proportional to the value of δJ, i.e. to the solution probability. It is apparent that the weak direction is bent and the confidence region is longer when very shorter arcs are considered. When the observations span a longer arc the weak direction resembles more a straight line that passes closer to the true solution.

5. Conclusions

In this work we focus our investigation on the OD problem when optical observations are taken on very short arcs, i.e. less than 10 deg. A classical least squares problem was formulated and DA techniques were used to implement an arbitrary order solver (referred to as DALS solver) and to nonlinearly describe the solution confidence region. For this purpose, three different algorithm were implemented: two to determine the confidence region along an assigned direction (SDQA-FIXED and AO-FIXED) and one to study the evolution of the weak direction (SDQA-WEAK). The main findings of this work can be summarized as

- The formulation of a least squares problem and its solution via the DALS on average improve the solution made available by IOD;

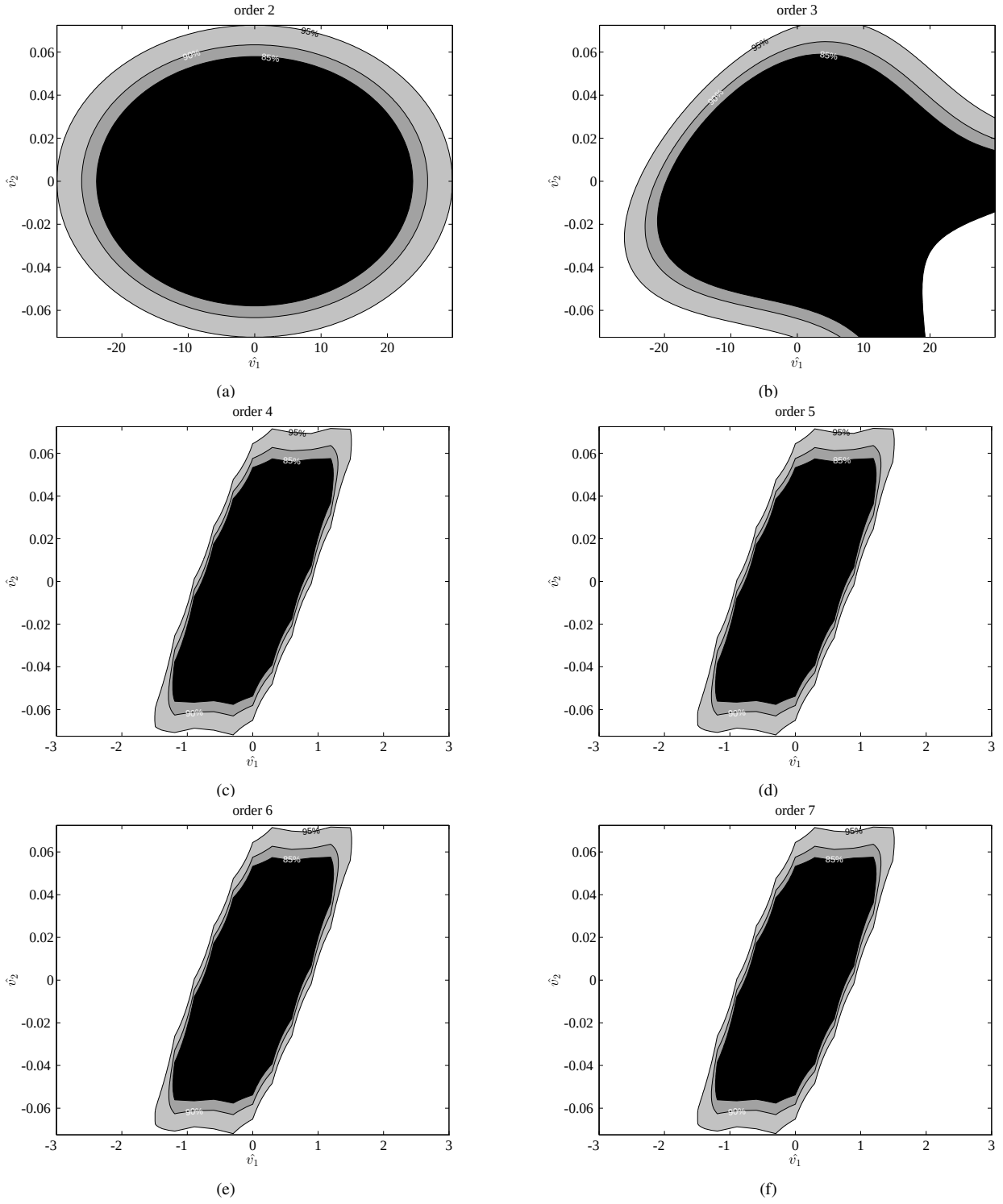


Figure 3: 2-D confidence region for the LEO test case with 8 observations at different expansion orders. The 2-D plane is defined by the two eigenvectors associated with the two largest eigenvalues of the covariance matrix. The axes are in scaled units.

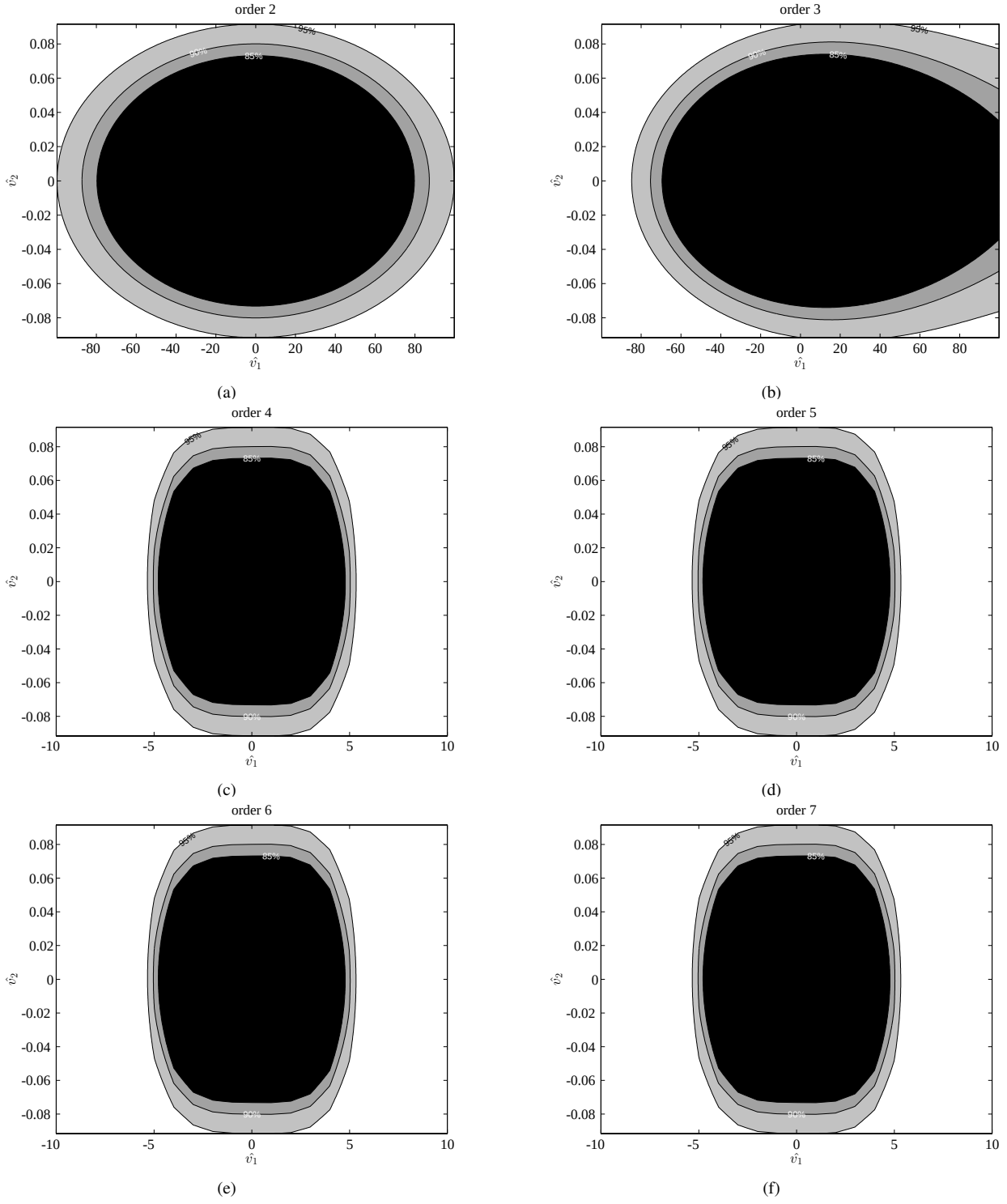


Figure 4: 2-D confidence region for the GEO test case with 8 observations at different expansion orders. The 2-D plane is defined by the two eigenvectors associated with the two largest eigenvalues of the covariance matrix. The axes are in scaled units.

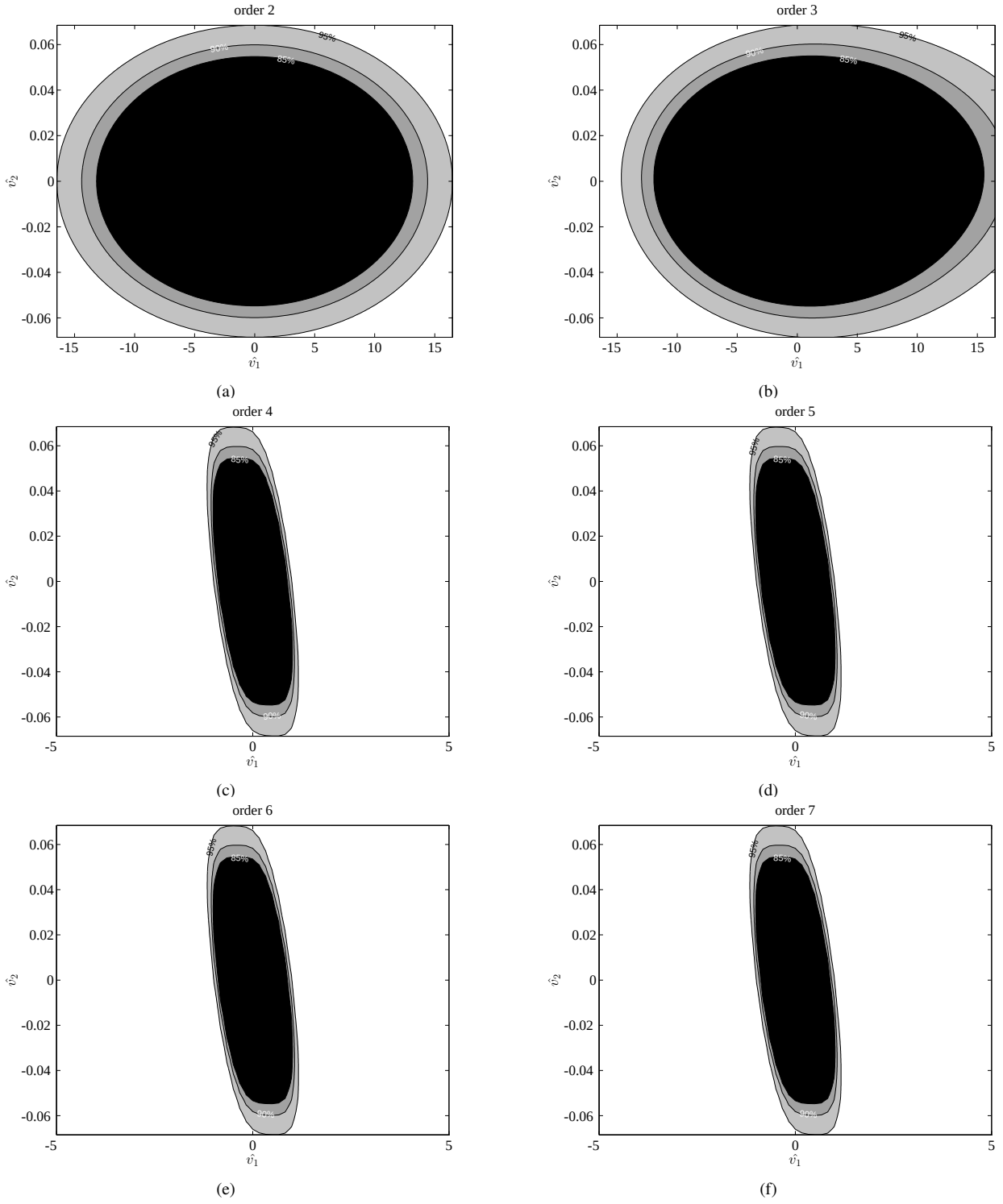


Figure 5: 2-D confidence region for the GTO test case with 8 observations at different expansion orders. The 2-D plane is defined by the two eigenvectors associated with the two largest eigenvalues of the covariance matrix. The axes are in scaled units.

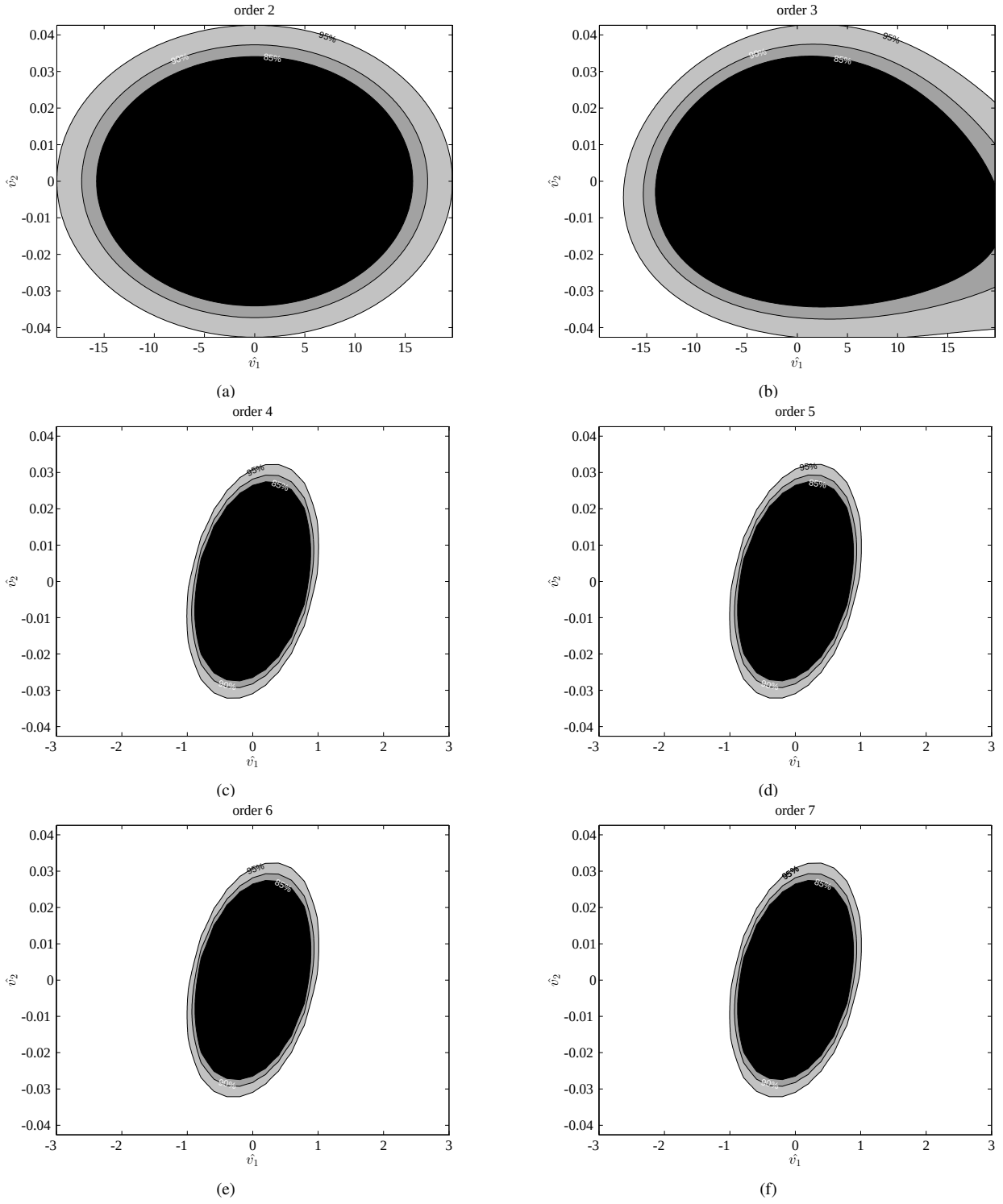
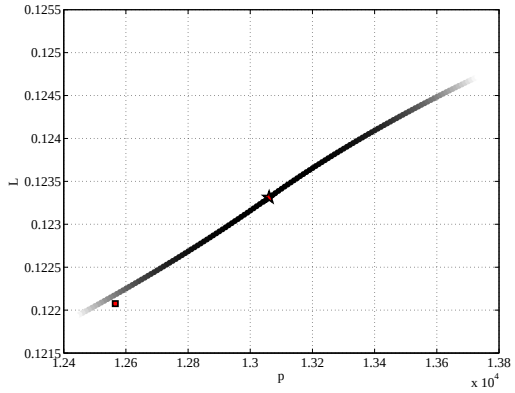
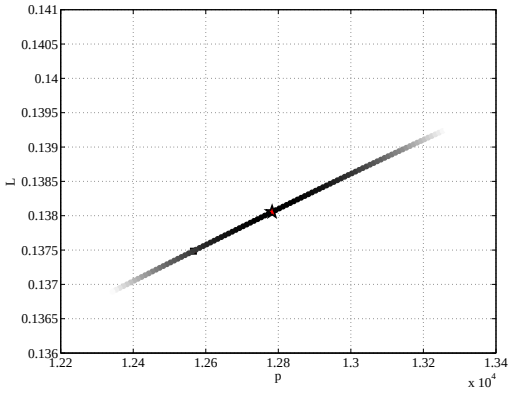


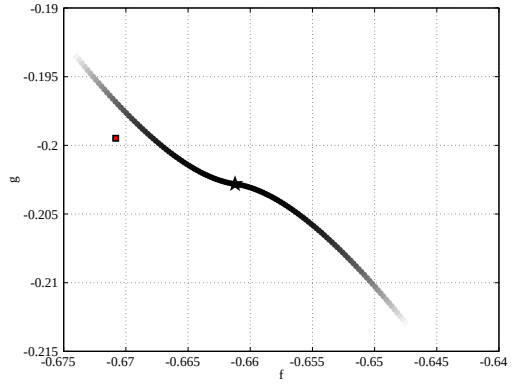
Figure 6: 2-D confidence region for the Molnyia test case with 8 observations at different expansion orders. The 2-D plane is defined by the two eigenvectors associated with the two largest eigenvalues of the covariance matrix. The axes are in scaled units.



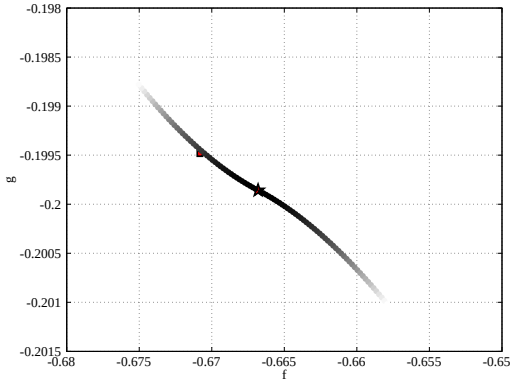
(a) 8 observations



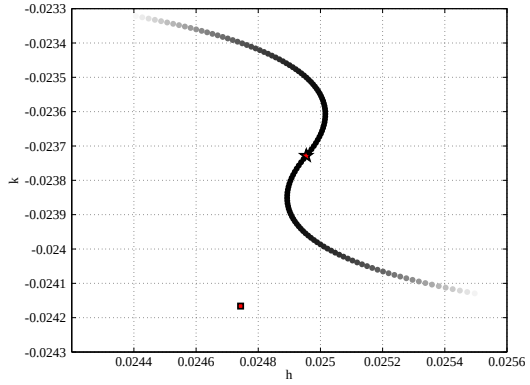
(b) 15 observations



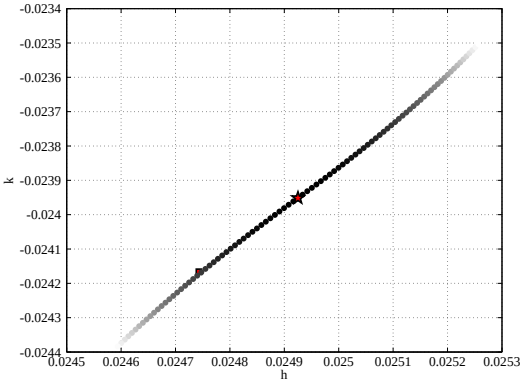
(c) 8 observations



(d) 15 observations



(e) 8 observations



(f) 15 observations

Figure 7: Projection on $p-L$, $f-g$, $h-k$ planes of the evolution of the weak directions for GTO object with 8 (left) and 15 (right) observations. The star indicates the optimal solution found by the DALS and the square the true solution. Note that the true solution is not the same in the $p-L$ planes as the LS solution refers to different epochs for the 8- and 15-observation cases.

Table 6: Covariance matrix in MEE for GEO and 15 observations

p [km]	f	g	h	k	L [rad]
1.8168618e+05	-3.2018546e+00	-4.8442433e-01	7.1359627e-03	-4.4223249e-02	-5.0162799e-03
-3.2018546e+00	5.6430383e-05	8.5102068e-06	-1.2686786e-07	7.7917549e-07	8.8363321e-08
-4.8442433e-01	8.5102068e-06	1.4683747e-06	-1.1699532e-08	1.1902194e-07	1.3627376e-08
7.1359627e-03	-1.2686786e-07	-1.1699532e-08	6.3490977e-10	-1.6830486e-09	-1.8656513e-10
-4.4223249e-02	7.7917549e-07	1.1902194e-07	-1.6830486e-09	1.0772493e-08	1.2225812e-09
-5.0162799e-03	8.8363321e-08	1.3627376e-08	-1.8656513e-10	1.2225812e-09	1.3915385e-10

Table 7: Covariance matrix in MEE for GEO and 8 observations separated by 4 minutes

p [km]	f	g	h	k	L [rad]
2.9925663e+05	-5.2508673e+00	-7.9459210e-01	1.1492387e-02	-7.1271923e-02	-8.0395247e-03
-5.2508673e+00	9.2140160e-05	1.3899915e-05	-2.0337197e-07	1.2502988e-06	1.4099964e-07
-7.9459210e-01	1.3899915e-05	2.3887892e-06	-1.9153710e-08	1.9096494e-07	2.1761432e-08
1.1492387e-02	-2.0337197e-07	-1.9153710e-08	9.8904909e-10	-2.6538518e-09	-2.9184732e-10
-7.1271923e-02	1.2502988e-06	1.9096494e-07	-2.6538518e-09	1.6987286e-08	1.9173605e-09
-8.0395247e-03	1.4099964e-07	2.1761432e-08	-2.9184732e-10	1.9173605e-09	2.1719357e-10

- The covariance matrix of the solution of the LS problem has large eigenvalues whose magnitude rapidly decreases with observation lenght;
- The confidence region of the solution of the LS problem is not accurately described by ellipsoids as terms above the second order in J are not negligible;
- The weak direction can be significantly bent

Our future work will be devoted to using the developed tools for observation correlation problem and particle filters initialization.

6. Acknowledgments

R. Armellin acknowledges the support received by the Marie Sklodowska-Curie grant 627111 (HOPT - Merging Lie perturbation theory and Taylor Differential algebra to address space debris challenges). This work has been partially funded by the Spanish Finance and Competitiveness Ministry under Project ESP2014-57071-R.
Gennaro Principe was funded by an EPSRC Doctoral Training Grant awarded by the Faculty of Engineering and the Environment of UoS.

References

[1] H. Klinkrad. *Space Debris - Models and Risk Analysis*. Springer-Verlag Berlin Heidelberg, 2006.

[2] Donald J Kessler and Burton G Cour-Palais. Collision frequency of artificial satellites: The creation of a debris belt. *Journal of Geophysical Research: Space Physics*, 83(A6):2637–2646, 1978.

[3] Paul W. Schumacher, Chris Sabol, Clayton C. Higginson, and Kyle T Alfriend. Uncertain Lambert Problem. *Journal of Guidance, Control, and Dynamics*, 38(9):1573–1584, 2015.

[4] O. Montenbruck and E. Gill. *Satellite orbits: models, methods, and applications*. Physics and Astronomy Online Library. Springer, 2000.

[5] Andrea Milani, Giovanni Gronchi, Mattia de' Michieli Vitturi, and Zoran Knezevic. Orbit determination with very short arcs. I admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 90(1-2):59–87, 2004.

[6] Andrea Milani, Giovanni Gronchi, Zoran Knezevic, Maria Eugenia Sansaturio, and Oscar Arratia. Orbit determination with very short arcs. II. Identifications. *Icarus*, 179(2):350–374, 2005.

[7] Jared M. Maruskin, Daniel J. Scheeres, and Kyle T. Alfriend. Correlation of Optical Observations of Objects in Earth Orbit. *Journal of Guidance, Control, and Dynamics*, 32(1):194–209, 2009.

[8] Johnny L. Worthy III and Marcus J. Holzinger. Incorporating Uncertainty in Admissible Regions for Uncorrelated Detections. *Journal of Guidance, Control, and Dynamics*, 38(9):1673–1689, 2015.

[9] Kohei Fujimoto. New Methods in Optical Track Association and Uncertainty Mapping of Earth-Orbiting Objects. 2013.

[10] Kyle J. DeMars and Moriba K. Jah. Probabilistic Initial Orbit Determination Using Gaussian Mixture Models. *Journal of Guidance, Control, and Dynamics*, 36(5):1324–1335, 2013.

[11] M. Berz. *Modern Map Methods in Particle Beam Physics*. Academic Press, 1999.

[12] M Berz. The new method of tpsa algebra for the description of beam dynamics to high orders. los alamos national laboratory. Technical report, Technical, Report AT-6: ATN-86-16, 1986.

[13] Martin Berz. The method of power series tracking for the mathematical description of beam dynamics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 258(3):431–436, 1987.

[14] Martin Berz. The new method of TPSA algebra for the description of beam dynamics to high orders. Technical report, Los Alamos National Laboratory, 1986.

[15] M. Berz. *Modern Map Methods in Particle Beam Physics*. Academic Press, 1999.

[16] C.F. Gauss. *Theoria motus corporum coelestium in sectionis conicis solem ambientum*. 1809.

[17] Ronald Aylmer Fisher et al. On an absolute criterion for fitting frequency curves. *Messenger of Mathmatics*, 41:155–160, 1912.

[18] Harold W Sorenson. Least-squares estimation: from Gauss to Kalman. *IEEE spectrum*, 7(7):63–68, 1970.

[19] A.H. Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[20] A. Milani and G. Gronchi. *Theory of Orbit Determination*. Cambridge University Press, 2010.

[21] D.A. Vallado. *Fundamentals of Astrodynamics and Applications*. Space Technology Library, 2013.

[22] Andrea Milani. The asteroid identification problem: I. recovery of lost asteroids. *Icarus*, 137(2):269–292, 1999.

[23] Dan Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[24] Zhe-Xian Wan. *Geometry of matrices: in memory of professor LK Hua (1910–1985)*. World Scientific, 1996.

[25] A. F. Seber and C. J. Wild. *Nonlinear Regression*. John Wiley & Sons, 2003.

[26] Andrea Milani, Maria Eugenia Sansaturio, Giacomo Tommei, Oscar Arratia, and Steven R Chesley. Multiple solutions for asteroid orbits: computational procedure and applications. *Astronomy & Astrophysics*, 431(2):729–746, 2005.

[27] Roberto Armellin, Pierluigi Di Lizia, and Renato Zanetti. Dealing with uncertainties in angles-only initial orbit determination. *Celestial Mechanics and Dynamical Astronomy*, 125(4):435–450, 2016.

[28] C. Sabol, T. Sukut, K. Hill, K.T. Alfriend, B. Wright, Y. Li, and P. Schumacher. Linearized orbit covariance generation and propagation analysis via simple monte carlo simulations.

[29] V. Vittaldev, R. Linares, and R.P. Russell. Spacecraft uncertainty propagation using gaussian mixture models and polynomial chaos expansions.