M. Sciarra, M. Losacco, D. Santeramo, P. Di Lizia
Impact Probability Computation for Neo Resonant Returns Through a Polynomial Representation of the Line of Variations
Acta Astronautica, Vol. 168, 2020, p. 80-91
doi:10.1016/j.actaastro.2019.07.009

Impact probability computation for NEO resonant returns through a polynomial representation of the Line of Variations

Marcello Sciarra[1,*], Matteo Losacco[1], Daniele Santeramo[1], Pierluigi Di Lizia[1]

Abstract

A differential algebra based representation and propagation of the Line of Variations for Near Earth Objects impact monitoring is presented in this paper. The Line of Variations is described at the initial epoch by a high-order polynomial that is propagated forward in time. An Automatic Domain Splitting algorithm is embedded in the numerical integrator, in such a way that when the polynomials truncation error becomes too large, the line is split as many times as necessary to meet accuracy requirements. The Line of Variations is propagated forward in time until an intersection with a properly defined target plane occurs for all the generated subdomains. The subdomains are then projected onto the target plane to compute the impact probability by numerically integrating an associated one-dimensional probability density function. The proposed approach is applied to different test-cases to assess the performance of the method for the different possible shapes of the initial confidence region. Starting from a case of direct encounter, the technique is tested up to the case of a resonant return, which features critical nonlinearities.

Keywords: Line of Variations, Impact probability, Near Earth Asteroids, Resonant Return, Differential Algebra, Automatic Domain Splitting

1. Introduction

The increasing attention towards Space Situational Awareness (SSA) requires the development of efficient and reliable numerical techniques for different applications. In particular, as an increasing amount of Near Earth Objects (NEOs) are being detected, the efficient estimation of their impact probability (IP) with the Earth is gaining relevance. This is especially true in case of availability of short observational arcs for newly discovered NEOs. The uncertainty of the initial orbital estimate tends to be large and the computational complexity of the IP assessment increases accordingly. In addition, when the NEO experiences a close planetary encounter along the path to the potential impact, the IP

*Corresponding author
Email address: marcello.sciarra@mail.polimi.it (Marcello Sciarra)
[1]Politecnico di Milano, Department of Aerospace Science and Technology, via la Masa 34, 20156, Milan, Italy

computation becomes extremely challenging, as the highly nonlinear dynamics critically expand the uncertainty set about the NEO's trajectory. In case of Earth close encounters, the broadening of the range of possible post-encounter orbital periods may even trigger resonant returns: if the asteroid's period is $P = h/k$ years, then after k revolutions of the asteroid and h revolutions of the Earth, the two bodies may come again to a risky conjunction.

Monte Carlo simulations offer a reliable solution to the problem. Nevertheless, their naive implementation tends to be extremely time consuming. On the other hand, approaches based on linearizations provide low accuracy standards. The current state-of-the-art technique in the field is based on the idea of the Line of Variations (LOV), introduced for the first time by A. Milani in [1]. This approach exploits the particularly elongated shape of the six-dimensional uncertainty region to describe it along a one-dimensional curve, that is a line of weakness of the orbit determination solution.

When a NEO is observed, the initial orbit determination is based on the least-square minimization of a target function

$$Q = \frac{1}{m}\xi \cdot \xi \tag{1}$$

where $\xi \in \mathbb{R}^m$ are the observation residuals and $m = 2N_{obs}$, assuming two coordinates per observation and considering N_{obs} observations. Calling X^* the nominal state, the target function can be expanded in Taylor series as

$$Q(X) = Q(X^*) + \underbrace{\frac{1}{m}(X - X^*) \cdot (X - X^*) + O(X - X^*)}_{\Delta Q(X)} \tag{2}$$

We can define the *region of confidence* for a given value of σ as the area described by

$$m\Delta Q(X) \leq \sigma^2 \tag{3}$$

When the high order terms are neglected, the confidence region becomes a *confidence ellipsoid*, described by a covariance matrix $\Gamma = C^{-1}$ where the *normal matrix* C is defined as $C = B^T B$, with $B = \dfrac{\mathrm{d}\xi}{\mathrm{d}X}(X)$. When the asteroid is newly discovered and it has been observed for a very short timespan, the solution is weak and Γ is ill-conditioned. In this case, there is a very weak direction along the eigenspace of the largest eigenvalue of Γ which is called *Line of Variations*.

As presented by Chesley and Chodas in [2], the LOV offers the possibility to work with a reduced-dimension yet representative sampling of the confidence region. In the nonlinear case, calling $\sigma(X)$ the largest eigenvalue of $\Gamma(X)$ and $V_1(X)$ the corresponding eigenvector, then $\sigma(X) V_1(X)$ is a well-defined and smooth vector field and therefore the LOV is the unique solution of the differential equation

$$\frac{\mathrm{d}X}{\mathrm{d}\sigma} = \sigma(X) V_1(X) \tag{4}$$

On the contrary, when a linear approximation is feasible, the LOV is simply the major axis of the confidence ellipsoid, thus corresponding to the direction of the eigenvector V_1.

A further definition of the LOV is offered by Milani et al. in [3] and it is directly inspired by the operational procedure typically used to compute the multiple solutions along the LOV. For every point X, the direction V_1 is defined and therefore also the orthogonal hyperplane $H(X)$ can be defined as

$$H(X) = \left\{ Y \,|\, (Y - X) \cdot V_1(X) = 0 \right\} \tag{5}$$

Starting from a guess X, a differential correction procedure constrained to the hyperplane $H(X)$ is performed. For such correction, the constrained normal equation $C_H \Delta X = D_H$ is considered, where $\Delta X = \Delta H$ along the direction H and $\Delta X = 0$ along V_1. The weak direction and the hyperplane are computed and iterated until convergence: if $\overline{X}$ is the convergence value, then $D_H\left(\overline{X}\right) = 0$ must hold and therefore the right-hand side of the *unconstrained* normal equation is parallel to the weak direction:

$$D\left(\overline{X}\right) \parallel V_1\left(\overline{X}\right) \tag{6}$$

Following this idea, the LOV can be defined as the set of points X for which the target function gradient is along the weak direction. Such definition opens the door to the application of the LOV technique to cases in which a nominal solution is not given. Current NEO surveillance systems at JPL and at the University of Pisa adopt this last definition. In our application, the initial orbit determination is not included in the algorithm and the data are retrieved in the form of nominal state and related covariance matrix. In this framework, the LOV is only defined as the direction corresponding to the largest eigenvalue of the covariance ellipsoid.

In the standard approach by Chesley and Chodas [2], the LOV is sampled according to the chosen probability distribution, either uniform or Gaussian. Each generated sample is called virtual asteroid (VA) and has a certain probability to represent the true orbit. The VAs that directly impact the Earth take the name of virtual impactors (VI) and in the simplest approach the IP is computed as the ratio of VIs over VAs. This approach does not consider all impacts that may occur for nearby off-LOV solutions and therefore some local analysis in the neighborhood of the close approaches is performed. To this aim, a target plane is defined and used to eliminate the role of time. The b-plane (orthogonal to the incoming asymptote of the geocentric hyperbola) can be adopted if the nominal orbit enters the sphere of influence of the Earth. Alternatively, the modified target plane (MTP) is defined as orthogonal to the closest approach point along the trajectory. The MTP was introduced in 1999 by Milani and Valsecchi in [4] although the concept can be found earlier, see Chodas and Yeomans in [5].

After projecting the initial covariance onto the target plane at the impact epoch, the IP can be computed by integrating a one-dimensional or two-dimensional probability density function (pdf), depending on the elongation of

the confidence ellipse on the target plane, that is by what amount the major axis Λ is larger than the minor axis w:

- if $\Lambda \gg w$ then $P_I = \dfrac{p(\sigma_\Lambda)}{\Lambda}$
- if $\Lambda > w$ then $P_I = \dfrac{p(\sigma_\Lambda)\, p(\sigma_w)}{\Lambda w} \pi r_\oplus^2$

An alternative approach is proposed in this paper. The idea is to move from a sampled description of the LOV to a polynomial representation of the curve, obtained with differential algebraic techniques. The continuous description of the curve allows us to find the portion of the LOV that leads to a collision with the Earth, rather than searching for a set of impacting samples as in the case of the VI approach. According to this approach, a highly accurate description of the LOV evolution up to the collision epoch can be obtained and the interpolation between different samples can be avoided.

The paper is organized as follows. Section 2 briefly introduces differential algebra (DA) and the automatic domain splitting (ADS) algorithm. Section 3 reports the numerical technique adopted to define the target plane and to propagate and projecct the LOV onto the MTP. In Section 4, the IP computation approach is presented based on the integration of the one-dimensional pdf. Section 5 shows the application of the technique to a direct impact test case. Section 6 points out the criticalities introduced by a resonant return and the technique adopted to solve them. Section 7 presents the issue of an enlarged initial confidence region and introduces a potential solution. Finally, Section 8 performs a comparison with other techniques to assess advantages and disadvantages of this approach. Section 9 summarizes the results and concludes the paper.

2. Differential algebra and Automatic Domain Splitting

Differential algebraic techniques were developed after the need to solve analytic problems with algebraic means. The theory of DA was developed for the first time by Joseph Liouville during the XIX century but the complete algebraic theory of the technique was presented by Ritt in [6]. The following overview is based on the description given by Martin Berz in [7].

Typically, in a computational environment, numbers take the place of functions and the algorithms are based on the pointwise functions evaluation. DA techniques try to increase the information that can be obtained from a function with respect to its mere evaluation. This additional information is provided by the implementation of an algebra of Taylor polynomials to which differentiation and integration can be easily added, thus obtaining a DA.

A relevant application of DA techniques arises for the propagation of an ODE set (see [8] for details): in this framework, a high order polynomial expansion of the propagated state can be obtained with respect to the initial condition. This result is achieved by substituting all the operations involved in the adopted

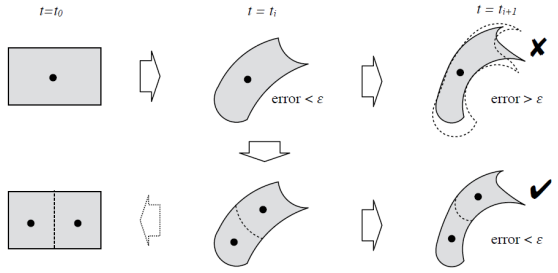


Figure 1: Basic principle of Automatic Domain Splitting

numerical integration scheme with the corresponding DA operations. In particular, if the initial polynomial represents the n-σ uncertainty region about the nominal state, its numerical propagation yields a polynomial description of the confidence region at the final integration time.

When the dynamics is highly nonlinear and for long-term propagations, a single polynomial expansion is not capable of providing an accurate description of the real confidence region evolution. To maintain a certain level of accuracy, an Automatic Domain Splitting (ADS) algorithm can be implemented to support the simple DA propagation, see [9]. The ADS algorithm, whose working principle is sketched in Fig. 1, estimates the truncation error of the polynomials over their own domain and decides wheather the truncation error is above or below a selected tolerance. When the estimated error is larger than the tolerance, the previous integration state (i.e. the last accurate state in the propagation history) is retrieved and split in two equal parts along a certain direction. This direction corresponds to the state variable along which the truncation error is estimated to be the largest. By dividing the domain by a factor 2, the truncation error is reduced by a factor 2^{n+1}. The polynomials are then evaluated on the two generated subdomains and propagated independently until the truncation error overcomes again the selected tolerance. This procedure is repeated as many times as necessary, until a certain limit of splits $\mathcal{N}_{max}$ is reached. Once the splitting tolerance tol_{split} and N_{max} are defined, the procedure is fully automatic and guarantees an accurate nonlinear propagation of uncertainties. In this work, tol_{split} is set to 10^{-10} and $\mathcal{N}_{max}$ to 16 whereas DA is used to obtain 12th-order polynomial expansions.

The DA routines exploited in this work are implemented in the DACE software [2] and more details on the numerical implementation can be found in [10]. The adopted numerical integration scheme is a DA version of a 7/8 Dormand-Prince (8th order solution for propagation, 7th order solution for step size control) RK integrator.

[2] https://github.com/dacelib/dace

3. DA-based Line of Variations

3.1. LOV generation

All the analyses presented in this paper were carried out considering data obtained from the NEODyS database (`https://newton.spacedys.com/neodys/`), which provides nominal states in mean equinoctial elements and the associated covariance matrices Γ. The LOV is a coordinate-dependent subspace of the confidence region and a proper choice of the coordinate system is crucial. According to [11], the LOV produces better results if computed in Cartesian coordinates and this choice is adopted in this work accordingly. The initialization of the LOV requires that the uncertainty ellipsoid represented by Γ is transformed from Equinoctial into Cartesian coordinates.

In this work, this nonlinear covariance transformation is performed through a DA-based approach. Full details on the method and its generalization can be found in [12], while the main steps are here presented. By indicating with $\vec{x}_E$ the equinoctial representation of the orbital state and with $\vec{x}_C$ its Cartesian representation, the two representations are related by a nonlinear coordinates transformation

$$\vec{x}_C = \mathcal{F}(\vec{x}_E) \tag{7}$$

We consider covariance matrices only, therefore all statistical moments above second order are here neglected. In the DA framework, the state in the original coordinates is written as

$$[\vec{x}_E] = \vec{\bar{x}}_E + \delta\vec{x}_E \tag{8}$$

where $\vec{\bar{x}}_E$ is the initial mean and the $\delta\vec{x}_E$ is the 1σ variation. The n-th order Taylor expansion of $\vec{x}_C$ with respect to $\delta\vec{x}_E$ is obtained by applying the transformation $\mathcal{F}$ to $[\vec{\bar{x}}_E]$ in the DA framework, i.e.

$$[\vec{x}_C] = \mathcal{F}([\vec{x}_E]) = \vec{\bar{x}}_C + \mathcal{T}_{\delta\vec{x}_C}(\delta\vec{x}_E) = \sum_{p_1+\cdots+p_n\leq k} c_{p_1\ldots p_n} \cdot \delta x_{E,1}^{p_1} \ldots \delta x_{E,n}^{p_n} \tag{9}$$

where $\vec{\bar{x}}_C$ is the zeroth-order term of the expansion, $\mathcal{T}_{\delta\vec{x}_C}$ is the Taylor map of the final state with respect to the initial state, and $c_{p_1\ldots p_n}$ are the coefficients of the resulting Taylor polynomial. The Taylor polynomial in the form of Eq. 9 can be used to efficiently compute the propagated statistics by computing the j-th moment of the transformed pdf using the polynomial map. The analytical expressions of the first two statistical moments for a generic scalar random variable z are

$$\begin{cases} \mu = E\{z\} \\ P = E\{(z-\mu)^2\} \end{cases} \tag{10}$$

where μ is the mean value, P is the covariance and the expectation value of z is

$$E\{z\} = \int_{-\infty}^{+\infty} z g(z) \mathrm{d}z \tag{11}$$

The moments of the transformed pdf can be computed by applying the multivariate form of Eq. 10 to the Taylor expansion in Eq. 9. The result for the first two moments are

$$\begin{cases} \vec{\mu}_i & = E\{[\vec{x}_{C_i}]\} = \sum_{p_1+\cdots+p_n\leq k} \vec{c}_{i,p_1\ldots p_n} \vec{E}\{\delta x_1^{p_1}\ldots\delta x_n^{p_n}\} \\ \vec{P}_{ij} & = E\{([\vec{x}_{C_i}]-\mu_i)([\vec{x}_{C_j}]-\mu_j)\} \\ & = \sum_{\substack{p_1+\cdots+p_n\leq k,\\ q_1+\cdots+q_n\leq k}} \vec{c}_{i,p_1\ldots p_n}\vec{c}_{j,p_1\ldots p_n} E\{\delta x_1^{p_1+q_1}\ldots\delta x_n^{p_n+q_n}\} \end{cases} \tag{12}$$

where $\vec{c}_{i,p_1\ldots p_n}$ are the Taylor coefficients of the Taylor polynomial describing the i-th component of $[\vec{x}_C]$. The expectations in Eq. 12 are computed through the Isserlis' formula.

Eq. 12 allows the covariance to be retrieved in Cartesian coordinates. The LOV can then be initialized, according to the linear approximation definition presented in Section 1, by computing the eigenvalues and eigenvectors of the Cartesian covariance matrix and considering the eigenvector associated to the eigenvalue of largest magnitude. Considering the eigenvalues and eigenvectors as sorted in ascending order of magnitude (i.e. the last component being the major axis), the LOV is first initialized in the DA framework in the space of the eigenvectors as

$$[\vec{x}_{\mathrm{EIG}}] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \Lambda\cdot\delta x_6 \end{bmatrix}^T \tag{13}$$

where Λ is a scaling factor for the DA polynomial expressed as function of δx_6, which is a coordinate spanning the direction of the eigenvector corresponding to the largest eigenvalue λ_6 of the covariance matrix Γ. If the desired analysis covers the $\pm 3\sigma$ interval, then $\Lambda = 3\sqrt{\lambda_6}$. The LOV is finally obtained by rotating $\vec{x}_{\mathrm{EIG}}$ back from the space of the eigenvectors to the real Cartesian space through the eigenvectors matrix U and adding the nominal state:

$$[\vec{x}_{\mathrm{LOV}}] = [\vec{x}_0[+U\cdot[\vec{x}_{\mathrm{EIG}}] \tag{14}$$

As a result, $[\vec{x}_{\mathrm{LOV}}]$ is a vector of polynomials in δx_6 that spans the LOV, i.e. it is a polynomial representation of the LOV at the initial epoch.

3.2. Numerical propagation

The numerical propagation of the LOV is performed in an N-body dynamics including relativistic corrections. This is required for complex encounters, as an inaccurate description of a close planetary approach may lead to unacceptably wrong post-encounter trajectories. The dynamics, taken from [9], is:

$$\begin{aligned} \ddot{\vec{r}} = G\sum_i \frac{m_i(\vec{r}_i-\vec{r})}{r_i^3}\Bigg\{1 - \frac{2(\beta+\gamma)}{c^2}G\sum_j\frac{m_j}{r_j} - \frac{2\beta-1}{c^2}G\sum_{j\neq i}\frac{m_j}{r_{ij}} + \frac{\gamma|\dot{\vec{r}}|^2}{c^2} + \\ + \frac{(1+\gamma)|\dot{\vec{r}}_i|^2}{c^2} - \frac{2(1+\gamma)}{c^2}\dot{\vec{r}}\cdot\dot{\vec{r}}_i - \frac{3}{2c^2}\left[\frac{(\vec{r}-\vec{r}_i)\cdot\dot{\vec{r}}_i}{r_i}\right]^2 + \frac{1}{2c^2}(\vec{r}_i-\vec{r})\cdot\ddot{\vec{r}}_i\Bigg\} \\ + G\sum_i\frac{m_i}{c^2 r_i}\cdot\left\{\frac{3+4\gamma}{2}\ddot{\vec{r}} + \frac{\{[\vec{r}-\vec{r}_i]\cdot[(2+2\gamma)\dot{\vec{r}}}{r_i^2} - \frac{(1+2\gamma)\dot{\vec{r}}_i]\}(\dot{\vec{r}}-\dot{\vec{r}}_i)}{r_i^2}\right\} \end{aligned} \tag{15}$$

where $\vec{r}$ is the position of the object in Solar System barycentric coordinates, G is the gravitational constant, m_i and $\vec{r}_i$ are the mass and the barycentric position of Solar System body i, $r_i = |\vec{r}_i - \vec{r}|$ and c is the speed of light in vacuum. β and γ are the parametrized post-Newtonian parameters measuring the nonlinearity in superposition of gravity and space curvature produced by unit rest mass and are set to one, since the analysis is performed in the framework of general relativity.

3.3. Target plane

Before propagating the LOV, the Target Plane has to be identified and fixed. In this work we choose to adopt the MTP instead of the b-plane. The main reason for this selection is that the b-plane requires the nominal trajectory to enter the sphere of influence of the Earth and this is not guaranteed to be true, particularly when the impacts occur towards the tips of the LOV. The MTP is identified by propagating the object's nominal trajectory until the condition of minimum geocentric distance is met. Since the propagation spans many revolutions of the asteroid, the condition $\rho = \vec{r}_{\mathrm{GEO}} \cdot \vec{v}_{\mathrm{GEO}} = 0$ (where the two elements of the scalar product are the geocentric position and velocity of the asteroid, respectively) is verified several times. In the current version of the tool, we decided to work with a given estimate of the impact epoch. Therefore, the condition above is considered to be met only if it occurs inside a *collision timespan* of ± 100 days around the expected impact epoch. When ρ changes sign during the propagation, a bisection routine is started to identify the exact $\rho = 0$ condition and subsequently fix the MTP in the heliocentric reference frame. Working with an estimate of the impact epoch, taken from NEODyS, is a simplifying assumption that is not suitable for operational monitoring. Yet, the goal of this paper is to assess the performance of the proposed technique in terms of IP computation which is not affected by this assumption.

3.4. ADS-based LOV propagation

The LOV can now be propagated forward in time to the MTP. The propagation is performed by means of an algorithm that embeds ADS inside an RK numerical integrator as illustrated in Fig. 2. The initial conditions are initialized as a vector of DA variables that spans the LOV at the initial epoch, as shown in Eq. 14. As each RK step is performed, the truncation error of the polynomials is estimated and, when it exceeds a selected tolerance, the domain is split in two subdomains. This procedure is repeated as many times as necessary in order to guarantee that the propagation is carried on with an accurate description of the LOV evolution. One by one, all the generated subdomains are propagated until a final integration condition is met. There are three possible conditions to stop the integration:

1. The reference point (i.e. the center) of the subdomain impacts the Earth surface before or during the *collision timespan*. If this happens, the integration is stopped to avoid that large nonlinearities compromise the

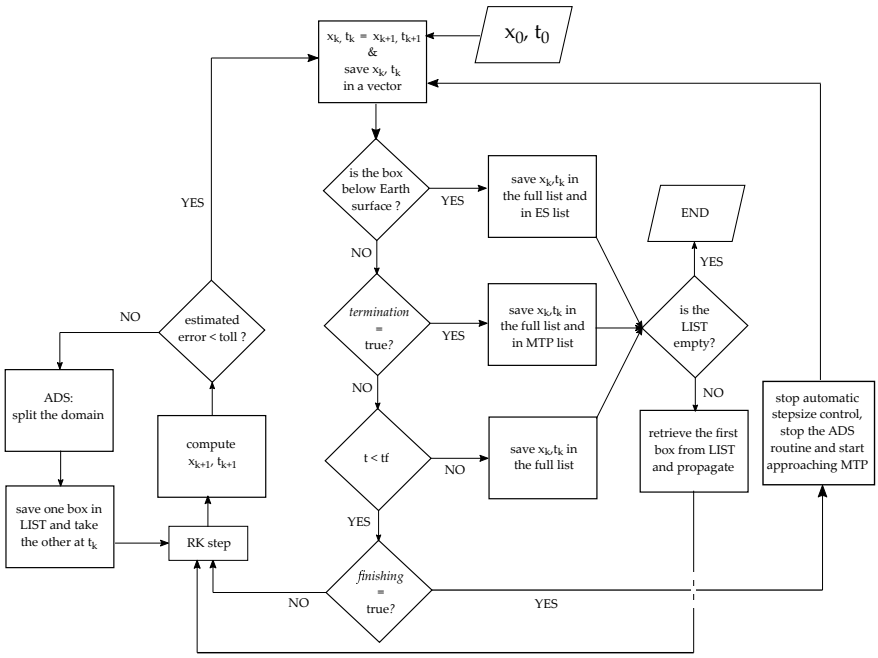


Figure 2: ADS-based propagation algorithm. Several lists are mentioned, in particular: LIST is the vector of subdomains generated by the splitting processes and that must still be propagated. ES contains the subdomains that cross the Earth surface during the propagation; MTP contains all the subdomains that hit the target plane; full list contains all the propagated subdomains.

efficiency of the tool and the subdomain is saved in a separate list for IP computation (ES in the scheme).

2. The subdomain reaches the MTP inside the *collision timespan* before its center hits the Earth surface. In this case the subdomain is saved in a second list (MTP in the scheme).

3. The subdomain does not cross the MTP inside the *collision timespan*. If this happens, the integration is stopped right after the subdomain exits the timespan and the subdomain is excluded from IP computation.

The intersection with the target plane is a rather critical aspect of the algorithm. In fact, the scalar product ρ changes its value under several conditions and this might lead to including some fictitious MTP crossings in the IP computation. More specifically, the following conditions may infer a wrong MTP intersection:

- the scalar product changes sign at an epoch far away from a potential impact condition.

- the scalar product changes sign when the two bodies (i.e. the asteroid and the Earth) lie on opposite sides of the Sun.

- when a subdomain reaches the MTP or the final integration time and the following subdomain is initialized for propagation, the sign of the scalar product at the two states may be opposite. Therefore the algorithm may detect a fictitious crossing of the MTP that must be avoided.

- when a subdomain experiences a split while lying across the MTP, particular care has to be adopted to avoid that one of the two generated subdomains misses the intersection with the plane.

A simple yet very effective solution to fictitious or missed MTP crossing is offered by the following idea: when a subdomain is initialized for propagation, the initial stepsize is forced to be equal to 10^{-8}, which guarantees a better control on the presented numerical issues. Indeed, by exploiting the knowledge of the stepsize we can either include or exclude the first steps of a subdomain's propagation from being checked for potential intersection:

1. If the propagation of a new subdomain is started with a very small and fixed stepsize, the first steps can be easily excluded from the check of the MTP crossing. The issue of a fictitious inversion of the scalar product is thus avoided and, at the same time, the path of the object in those first steps is so small that real MTP intersections cannot be missed.

2. On the other hand, by limiting the control to the first steps of the propagation (by requiring the stepsize to be small), we avoid failures of MTP intersection detections due to a large initial stepsize for subdomains that lie on the MTP since the very beginning of their trajectory. By making reference to Fig. 3, this is crucial to make sure that, after a split occurred across the plane, the intersection is detected when the propagation of the red subdomain is started

As far as coding is concerned, the correct intersections of the MTP are identified by checking at every integration step the following conditions:

$$\begin{aligned}
&\text{(a)} \quad t_k > (t_{\mathrm{IMP}} - 100 \text{ d}) \\
&\text{(b)} \quad (\vec{r}_{\mathrm{GEO}_k} \cdot \vec{V}) \cdot (\vec{r}_{\mathrm{GEO}_{k-1}} \cdot \vec{V}) < 0 \\
&\text{(c)} \quad \vec{r}_{\mathrm{GEO}_k} \cdot \vec{V} > 0 \\
&\text{(d)} \quad |h| > 2 \cdot 10^{-8} \\
&\text{(e)} \quad t_{k-1} < (t_{\mathrm{f}} - \Delta t_{max})
\end{aligned} \tag{16}$$

where t_{IMP} represents the estimated collision epoch, t_f is the final integration time, h is the current stepsize, k subscript is the integration step counter and Δt_{max} is a fixed time interval slightly larger than maximum integration stepsize. Moreover, $\vec{V}$ represents the velocity of the center point of the subdomain and $\vec{r}_{\mathrm{GEO}}$ is its geocentric positon vector. This set of conditions represents the standard MTP intersection: time belonging to the *collision timespan* (a), ρ inversion (b) with negative-to-positive value (c), at least one step has occurred

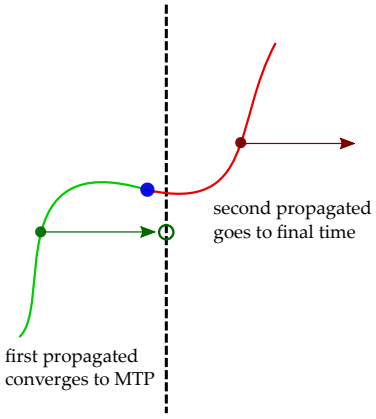


Figure 3: Example of subdomain splitting occuring across the MTP. The full polynomial is split into two: the green subdomain would reach the MTP while the red one would miss the intersection since at the beginning of its propagation its center lies beyond the plane.

after starting propagating the subdomain (d) and previous step not being the final integration time (e). The key role of conditions (d) and (e) is to prevent comparison of states belonging to separate subdomains (e.g. the last state of one subdomain's propagation and the first state of the following one);

Pertaining to the detection of MTP intersections when splits occurred across the plane, this is the set of required conditions:

$$\begin{aligned}
&\text{(a)} \quad t_k > (t_{\text{IMP}} - 100 \text{ dd}) \\
&\text{(b)} \quad \vec{r}_{\text{GEO}_k} \cdot \vec{V} > 0 \\
&\text{(c)} \quad |h| < 2 \cdot 10^{-8} \\
&\text{(d)} \quad |\vec{r}_{\text{GEO}_k} \cdot \vec{V}| > 10^{-10} \\
&\text{(e)} \quad \vec{r}_{\text{GEO}_k}|_{-1.0} \cdot \vec{r}_{\text{GEO}_k}|_{1.0} < 0 \\
&\text{(f)} \quad t_{k-1} < (t_{\text{f}} - \Delta t_{max})
\end{aligned} \tag{17}$$

which refer to: time belonging to the *collision timespan* (a), ρ being positive (b) with previous value larger than impact tolerance (d), at most one step occurred after starting propagating the subdomain (c), previous step not being the final integration time (f), and subdomain's tips lying on opposite sides of the MTP (e). In this case, conditions (d), (c) and (f) assure that we are taking into account only the first steps of a subdomain's propagation while condition (e) detects the crossing of the MTP.

Note that condition (d) of Eq. 16 and the condition (c) of Eq. 17, that refer to the scenarios number 1 and 2 respectively in the list above, are complementary and therefore the two sets cannot be satisfied simultaneously. By choosing a stepsize threshold of $2 \cdot 10^{-8}$, we can look separately for the two conditions: when the stepsize is below this threshold, that is the propagation just started, we only search for possible intersections due to splits that occur on the plane.

When the stepsize is larger than the theshold, that is a few steps have already been performed, we only search for standard intersections. If one of the two sets is fully satisfied, a Newton's method can be triggered to approach the plane.

3.5. Projection of the subdomains onto the MTP

After the propagation, we obtain a set of subdomains whose center point belongs to the MTP but that expand across the plane. In order to eliminate the role of time from the analysis, these subdomains are projected onto the MTP. During the propagation, indeed, the confidence region spreads very quickly along track. For this reason, the subdomains reach the target plane at time epochs that differ from each other but that are very close to the actual closest approach to the Earth. Therefore, by analyzing the projection of the LOV onto the MTP, the impact probability can be properly computed.

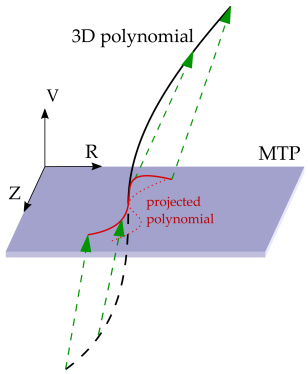


Figure 4: Projection of a subdomain onto the MTP: each point is projected along the direction of its own velocity vector, which is approximated by the available Taylor expansion of the velocity on the subdomain.

Instead of projecting along the direction normal to the plane, the projection is performed along the direction of the velocity vector in order to better approximate the dynamical path that leads each point of the subdomain to the MTP. Figure 4 illustrates the approach that is adopted to perform such a projection using DA. The idea is to perform a DA-based projection by relying on the available polynomial expansion of the asteroid velocity on each subdomain. The polynomial is used to project each point of the subdomain along its own velocity vector. After LOV propagation, all subdomains cross the MTP and are endowed with the Taylor expansion of their position vector $\vec{r}$ and velocity vector $\vec{v}$. In order to project the subdomain onto the MTP, the position vector is expressed as

$$\vec{r} = c_1\vec{R} + c_2\vec{v} + c_3\vec{Z} \tag{18}$$

where $\vec{v}$ is the polynomial expression of the subdomain's velocity vector while $\vec{R}$ and $\vec{Z}$ are the in-plane unit vectors of the MTP frame. The projection is achieved by imposing the out-of-plane component of the position vector to be

zero and finding the resulting coefficients c_1, c_2, and c_3. To this aim the following scalar products are computed:

$$\begin{aligned} \vec{r}\cdot\vec{R} &= c_1\vec{R}\cdot\vec{R} + c_2\vec{v}\cdot\vec{R} + c_3\vec{Z}\cdot\vec{R} = \\ &= c_1 + c_2\vec{v}\cdot\vec{R} \end{aligned} \tag{19}$$

$$\begin{aligned} \vec{r}\cdot\vec{v} &= c_1\vec{R}\cdot\vec{v} + c_2\vec{v}\cdot\vec{v} + c_3\vec{Z}\cdot\vec{v} = \\ &= c_1\vec{R}\cdot\vec{v} + c_2 v^2 + c_3\vec{Z}\cdot\vec{v} \end{aligned} \tag{20}$$

$$\begin{aligned} \vec{r}\cdot\vec{Z} &= c_1\vec{R}\cdot\vec{Z} + c_2\vec{v}\cdot\vec{Z} + c_3\vec{Z}\cdot\vec{Z} = \\ &= c_2\vec{v}\cdot\vec{Z} + c_3 \end{aligned} \tag{21}$$

The terms in the left-hand sides are known and are denoted with $A = \vec{r}\cdot\vec{R}$, $B = \vec{r}\cdot\vec{v}$ and $C = \vec{r}\cdot\vec{Z}$ in the following for the sake of notation. Furthermore we call $\cos\alpha = \vec{v}\cdot\vec{R}$ and $\cos\beta = \vec{v}\cdot\vec{Z}$. If the whole procedure is carried out with normalized velocity vector, then $v^2 = 1$. The simplified system to be solved is

$$\begin{cases} A = c_1 + c_2\cos\alpha \\ B = c_1\cos\alpha + c_2 + c_3\cos\beta \\ C = c_2\cos\beta + c_3 \end{cases} \tag{22}$$

The projection condition B=0 is then imposed and the system is easily solved to find the resulting expression of the three coefficients c_1, c_2 and c_3. Working within DA, all the elements in the expressions above are polynomials and therefore the obtained coefficients are also polynomials. The resulting coefficients can then be introduced in Eq. 18 to obtain the Taylor expansion of the projection of the subdomain onto the MTP.

$$\mathcal{T}_{\vec{r}_{MTP}} = \begin{bmatrix} \mathcal{T}_{c_1}|\vec{r}|\,\vec{R} \\ \mathcal{T}_{c_3}|\vec{r}|\,\vec{Z} \end{bmatrix} \tag{23}$$

where the notation $\mathcal{T}_{c_i}$ identifies the polynomial representation of the i-th coefficient. Besides granting an accurate projection, the above procedure provides a polynomial expansion of the position vector on the MTP, which means that the LOV is eventually fully described as a single-variate polynomial map in MTP coordinates.

4. 1D impact probability computation

Once the propagation and the projection of the LOV have been performed, the LOV is represented onto the MTP by a series of subdomains of different length, some of them having a geocentric distance lower than the Earth radius. In Fig. 5 a sketch of the LOV on the MTP is illustrated: the different subdomains of the LOV are highlighted with different colours and the impacting ones are dashed. Considering a one-dimensional pdf associated to the LOV, the IP between the asteroid and the Earth can be computed by integrating the pdf over the impacting subdomains. Different approaches can be exploited to perform this computation:

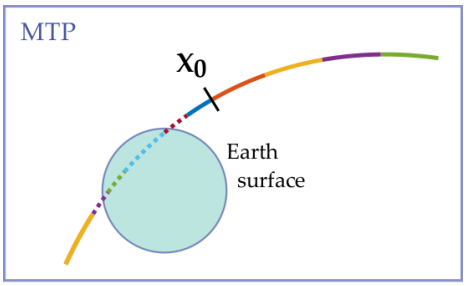


Figure 5: Sketch of the LOV projected on the MTP: dashed subdomains are the impactors which are considered in the IP computation.

- *Integration of the propagated pdf* – The nonlinear dynamics stretches the asteroid's confidence region, and its LOV. If a single subdomain is considered, the ratio of its length at the final and initial epoch provides a stretching factor Λ[3]. This stretching factor can be used to propagate the pdf associated to the subdomain. By multiplying it by the length l of the subdomain at the final epoch, the impact probability can be computed:

$$IP = \frac{p(\sigma)}{\Lambda} l \tag{24}$$

- *DA-based Monte Carlo* – Thanks to DA, the polynomial expansion of the position vector on the MTP is available, as explained in Section 2. Therefore, the state on the MTP corresponding to each point of the initial LOV can be retrieved by evaluating the polynomial map. Therefore the initial LOV can be sampled according to the chosen pdf and the impact probability can be computed as the ratio of the number of samples whose corresponding final position impacts with the Earth and the total number of samples:

$$IP = \frac{N_{\oplus}}{N} \tag{25}$$

- *Integration of the pdf over the initial domain* – Once the impacting subdomains are detected, the IP can be computed by simply integrating the pdf on the initial LOV over the portion of the line that collides with the Earth.

The first method enables a target plane analysis, if needed. Presented in large detail in [2], a target plane analysis allows to project the initial uncertainty

[3]The symbol used here is the same as the one adopted for the major axis of the confidence ellipse on the MTP seen in Section 1. Indeed there is a strong correlation between the two parameters. If we consider the initial 6D confidence region and project it onto the MTP to become a 2D ellipse, in a first-order approximation, its semi-major axis is exactly the propagated and projected LOV.

region onto the plane and to search for impactors out of the LOV. On the other hand the DA-based Monte Carlo is very simple and accurate, provided that the final polynomial map is accurate enough. As the accuracy may not be sufficient in case of highly nonlinear dynamics, the third method is adopted. A Gaussian pdf is associated to the initial LOV in the form:

$$\mathrm{pdf}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{26}$$

The mean μ is equal to zero in this case, as the pdf is centered at the expansion point of the LOV. In the DA framework, the DA independent variable δ spanning the LOV goes from -1 to $+1$ in the corresponding real interval $[-3\sigma, +3\sigma]$. To compute the probability density at a generic coordinate $x = q\sigma$ (i.e. $\delta = \frac{q\sigma}{3\sigma} = \frac{q}{3}$) the exponent in Eq. 26 becomes:

$$-\frac{(x-\mu)^2}{2\sigma^2} = -\frac{(q\sigma)^2}{2\sigma^2} = -\frac{1}{2}q^2 = -\frac{1}{2}(3\delta)^2 \tag{27}$$

and the pdf in the DA framework is:

$$\mathrm{pdf}(\delta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(3\delta)^2} \tag{28}$$

On first approximation, the subdomains whose center or one of the extrema lie within the Earth disk can be considered as fully impacting. A refinement can be obtained by finding the exact portion of the LOV that intersects the Earth surface, increasing the precision of IP computation. With reference to Fig. 6, in which a sketch of the split LOV is illustrated, the subdomains are classified as:

- *Internal* – Subdomains whose center lies within the Earth surface and are not adjacent to the blue-to-red transition. They are considered as full impactors.
- *External* – Subdomains whose center lies outside the Earth disk and are not adjacent to the transitiom. They are excluded from the IP computation.
- *Crossing* – Moving from left to right along the LOV, the last red subdomains before one or more blue subdomain and the first blue subdomain after one or more red subdomains. For these subdomains the exact intersection shall be found.

For the crossing subdomains, the point corresponding to the intersection with the Earth surface can be computed by numerically finding the root of the function:

$$f = ||\vec{r}_{\mathrm{MTP}}|| - R_{\oplus} \tag{29}$$

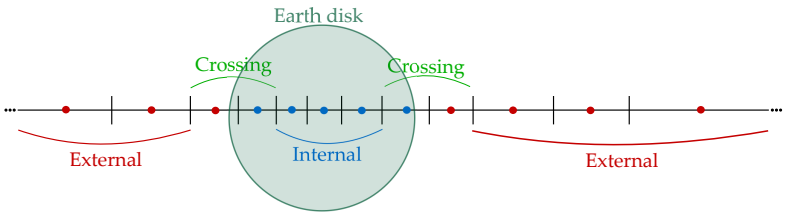


Figure 6: IP computation refinement: classification of the LOV subdomains.

where $R_\oplus$ is the Earth radius and $||\vec{r}_{\mathrm{MTP}}||$ is the polynomial expansion of the norm of the asteroid's position on the MTP. For each subdomain i, finding the exact intersection between the LOV and the Earth surface allows two coefficients $\mathcal{B}_i$ and $\mathcal{D}_i$ to be identified, which are then used to refine the IP computation as illustrated hereafter. Before applying the bisection method, the presence of an actual zero is checked. To this aim, the function is evaluated at the extrema and at the center point and the three resulting values are denoted as

$$a_i = f(+1) \qquad b_i = f(0) \qquad c_i = f(-1) \tag{30}$$

Calling $\mathcal{Z}_i$ the root of the function for the i-th subdomain, different possibilities arise for the assignment of the coefficients $\mathcal{B}_i$ and $\mathcal{D}_i$

case 1) $a_i \cdot b_i < 0$: Search for $\mathcal{Z}_i$ between 0 and +1.

(a) $a_i < 0$

$$\mathcal{B}_i = 0.5 \cdot (1 - |\mathcal{Z}_i|)$$
$$\mathcal{D}_i = 0.5 * (1.0 + \mathcal{Z}_i)$$

(b) $a_i > 0$

$$\mathcal{B}_i = 0.5 + 0.5|\mathcal{Z}_i|$$
$$\mathcal{D}_i = 0.5 * (-1.0 + \mathcal{Z}_i)$$

case 2) $c_i \cdot b_i < 0$: Search for $\mathcal{Z}_i$ between -1 and 0.

(a) $c_i < 0$

$$\mathcal{B}_i = 0.5 \cdot (1 - |\mathcal{Z}_i|)$$
$$\mathcal{D}_i = 0.5 * (-1.0 + \mathcal{Z}_i)$$

(b) $c_i > 0$

$$\mathcal{B}_i = 0.5 + 0.5|\mathcal{Z}_i|$$
$$\mathcal{D}_i = 0.5 * (1.0 + \mathcal{Z}_i)$$

case 3) $a_i < 0$ & $b_i < 0$ & $c_i < 0$:

$$\mathcal{B}_i = 1$$
$$\mathcal{D}_i = 0$$

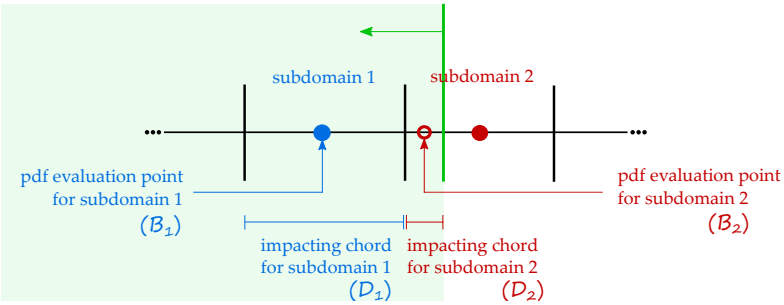


Figure 7: IP computation refinement: illustration of the role of the coefficients $\mathcal{B}_i$ and $\mathcal{D}_i$.

case 4) $a_i > 0$ & $b_i > 0$ & $c_i > 0$:

$$\mathcal{B}_i = 0$$
$$\mathcal{D}_i = 0$$

At the end of the above procedure, the IP is computed as:

$$IP = \sum_i \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\Big[\big(c_i + \mathcal{D}_i \cdot 0.5 w_i\big)\sigma\Big]^2\right\} \cdot 3\sigma \mathcal{B}_i w_i \tag{31}$$

in which c_i is the center of the i-th subdomain expressed in terms of its displacement with respect to the LOV center, and w_i is the width of the subdomain. As illustrated in Fig. 7, $\mathcal{D}_i$ defines by what amount the point of evaluation of the pdf has to be shifted with respect to the reference point and $\mathcal{B}_i$ represents the portion of the subdomain that impacts the Earth and therefore multiplies the pdf.

5. Asteroid 2010 RF12 – Direct encounter

The proposed technique is now applied to a test case with a potential direct impact, i.e. during the propagation, no planetary deep encounters occur between the initial epoch and the impact. According to the NEODyS database[4], asteroid 2010 RF12 has a very elongated initial confidence region with a ratio of $3.8 \cdot 10^4$ between the largest and the second eigenvalue of the initial Cartesian covariance matrix, from now on called *eigenvalue ratio* χ. Due to the high eigenvalue ratio, the LOV is expected to provide a very good approximation of the behaviour of the full confidence region. This asteroid, whose initial conditions are taken at 00:00 of March, 23$^{\text{rd}}$ 2018, has a potential impact with Earth in 2095.

Figure 8 shows the resulting projection of the LOV onto the MTP associated to the potential impact. The straight shape of the propagated LOV confirms

[4]accessed in May, 2018

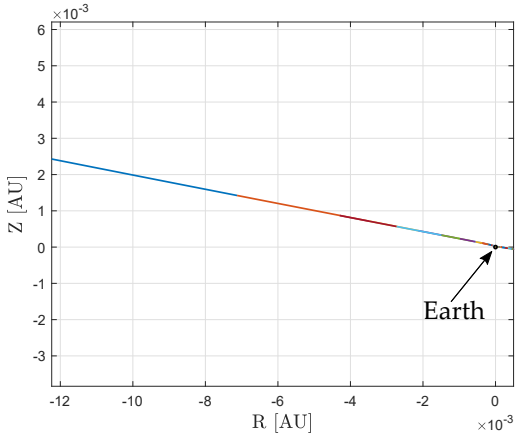


Figure 8: 2010 RF12: LOV projection onto the MTP at the epoch of impact in 2095

the relatively weak nonlinearity of the dynamics due to the absence of deep encounters. The subdomains (represented by the different colours) become smaller and smaller as the asteroid moves close to Earth (the small black dot close to the right end of the LOV). Inside the Earth disk (see Fig. 9 for a detailed view of Fig. 8) they become extremely small. Table 1 reports the IP computation for this test case, together with the reference value taken from the NEODyS database. The accuracy of the result and the computational time are satisfactory, especially considering the timespan covered by the numerical propagation (about 80 years) and the fact that the code has not been parallelized.

order	subdomains	comp. time [s]	IP	$\text{IP}_{\text{NEODyS}}$
12	91	1017	0.05943	0.06

Table 1: 2010 RF12 IP computation. The computational time is the running-time on an AMD Opteron 6376 processor with a total of 64 cores @ 2.3 GHz and 256 Gbytes of RAM.

6. Asteroid Apophis – Resonant return

A resonant return introduces relevant complexity in the computation. To illustrate the associated criticalities, the case of asteroid Apophis is here investigated. When first observed, the asteroid was estimated to have a very high impact probability with the Earth in 2029. When this collision was ruled out, the 2029 deep encounter still accounted for a possible resonant return collision in 2036. Further observations ruled out also this risk. The initial conditions used for the following analysis were provided by G. Tommei (private communication) and they refer to the results of a 2004-2005 observational campaign that still included the possibility of a collision in 2036. In this case $\chi = 5 \cdot 10^2$,

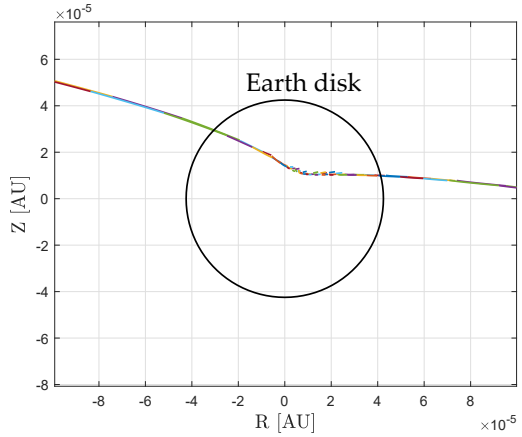


Figure 9: 2010 RF12: detail of the LOV projection onto the MTP at the epoch of impact in 2095 close to the Earth disk

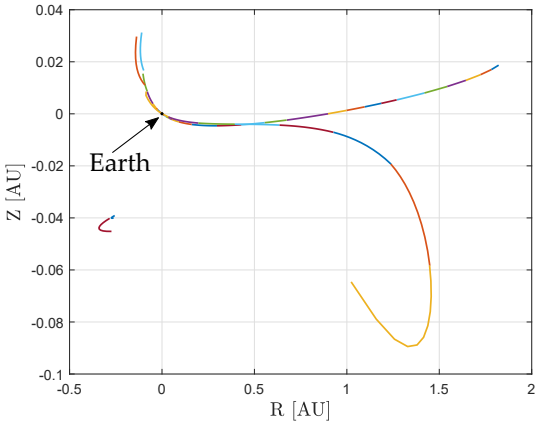


Figure 10: Apophis: LOV projection onto the MTP at the epoch of impact in 2036

which is lower than the previous case but still describes an elongated initial confidence region. The nonlinearities due to the deep encounter with the Earth in 2029 critically enlarge the confidence region and the LOV is strongly stretched and tangled, as evidently illustrated in Fig. 10. These nonlinearities induce a very high number of splits and many subdomains tend to reach the maximum number of splits $\mathcal{N}_{max}$. When this happens, the final polynomial map is not accurate , which affects in turn the reliability of IP computation. To better illustrate this issue, Fig. 11 shows the projection of the portion of the LOV close to the Earth disk. It is evident that the inner subdomains (in red and light blue) are completely inaccurate since they could not be split further after reaching N_{max}.

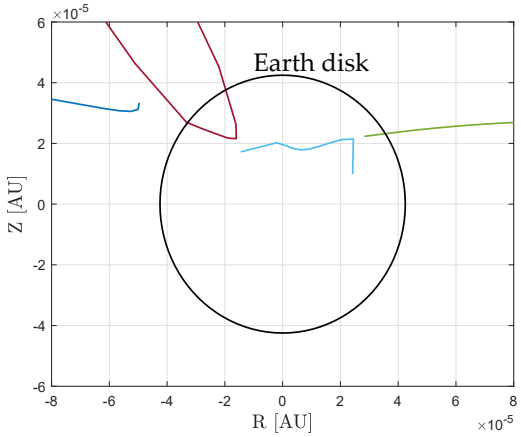


Figure 11: Apophis: detail of the LOV projection onto the MTP at the epoch of impact in 2036, close to the Earth disk

This problem must be necessarily circumvented to guarantee sufficiently accurate results for asteroids featuring a resonant return. To this aim, a refining routine is proposed:

1. After LOV propagation and projection, the subdomains (either accurate or inaccurate) whose center lies below the Earth surface are identified.

2. The subdomains are classified as *Internal*, *External* or *Transition* subdomains.

3. Internal subdomains are considered full impactors even if they are inaccurate in order to significantly reduce the computational burden.

4. Consider the Transition subdomains and retrieve the last accurate state within their propagation history.

5. Initialize a new ADS-based propagation of these subdomains that allows to further split the sudomains and to obtain an accurate polynomial map close to the Earth surface.

6. Identify the exact intersection between the LOV and the Earth disk to compute the IP.

The last item deserves to be discussed in more detail. With reference to Fig. 12, the center points of the blue subdomains lie below the Earth surface, whereas the center points of the red subdomains lie above the surface. Moving from left to right, the last blue subdomain and the first red one are the crossing subdomains, as the intersection between LOV and Earth surface may lie in either one or the other. These subdomains are repropagated and a schematic

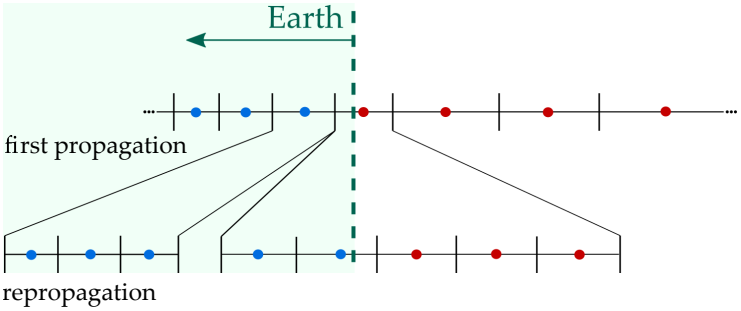


Figure 12: Subdomains repropagation approach: the crossing subdomains of the initial propagation (above) are repropagated and a finer grid is obtained (below)

representation of the possible outcome of the procedure is reported at the bottom of Fig. 12: several additional subdomains are generated from the two initial subdomains. By applying a procedure similar to the one presented in Section 4, the exact intersection can be identified inside the sub-subdomains. The IP computation is then performed as described in Section 4.

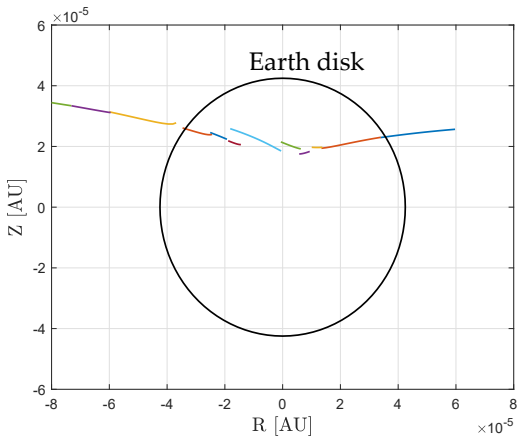


Figure 13: Apophis: detail of the LOV projection onto the MTP at the impact in 2036 after the application of the repropagation routine.

The beneficial effect of the repropagation routine for Apophis is evident from Fig. 13. The inaccurate subdomains are repropagated and this grants a better description of the LOV. Some of the subdomains inside the disk seem to introduce discontinuities along the LOV: this is not due to inaccuracy of the polynomials but rather to the projection routine. Indeed, the propagation of the impacting subdomains is stopped as soon as they reach Earth surface and they are then directly projected onto the MTP even if their center does not lie on the plane. This introduces an error that turns out to be negligible for IP

order	subdom.	comp. time [s]	IP	IP_{REF}
12	130	1474	$3.0134 \cdot 10^{-5}$	$3 \cdot 10^{-5}$

Table 2: Apophis IP computation

computation. The result of the IP computation is presented in Tab. 2: once again, the predicted value is close to the reference from NEODyS.

7. Low eigenvalue ratio and 2D method

7.1. Low eigenvalue ratio

The previous sections focused on two test cases in which the LOV is able to approximate the behaviour of the full confidence region and the computed IP is accurate. However, more generally, the impactors are gathered in a region that can be more or less elongated and variously distributed inside the initial 6D uncertainty ellipsoid. If this region lies across the LOV, then the one-dimensional LOV technique is capable of approximating the impact probability of the full portion of the curve leading to an impact by reducing the analysis along a single direction. If the confidence region is very elongated, this condition is likely to occur. On the other hand, this assumption may fail for non-elongated uncertainty ellipsoids, i.e. if the first two eigenvalues are about the same order of magnitude. Evidently, the eigenvalue ratio tips the balance between the different behaviours and assumes a key role in this analysis.

The reliability of the proposed method for low values of χ is now investigated. The case of asteroid 2016 LP10 is analysed. According to NEODys, this asteroid has an eigenvalue ratio of about 8. Thus, it features a less elongated initial uncertainty ellipsoid with respect to the previous test cases. The mere application of the technique as presented in the previous sections yields $IP = 0$ even if the impact probability, according to the NEODyS database, is $4.1 \cdot 10^{-5}$. This means that the portion of the initial uncertainty set leading to the impact lies in a region that is fully outside the LOV. Therefore, the one-dimensional technique fails (see Fig. 14). In such cases, the standard LOV technique performs a projection of the covariance associated to each VA onto the MTP in order to perform an off-LOV analysis. Yet, our analysis is carried out starting with a determined orbit and only one covariance matrix is known for the whole distribution. An alternative approach is proposed here, which is based on the idea of increasing the number of variables of the Taylor expansions by considering the first two eigenvalues of the covariance matrix.

7.2. Two-dimensional technique

When the one-dimensional approach is deemed to be unreliable, a two-dimensional technique can be adopted. The technique relies on the observation that DA can provide expansions with respect to an arbitrary number of

variables. An extended treatment of two-dimensional approaches is provided in [13] where the concept of the LOV is extended to introduce the *manifold of variations*. Similarly to Eq. 13, the manifold of variation is initialized in the space of the eigenvectors as

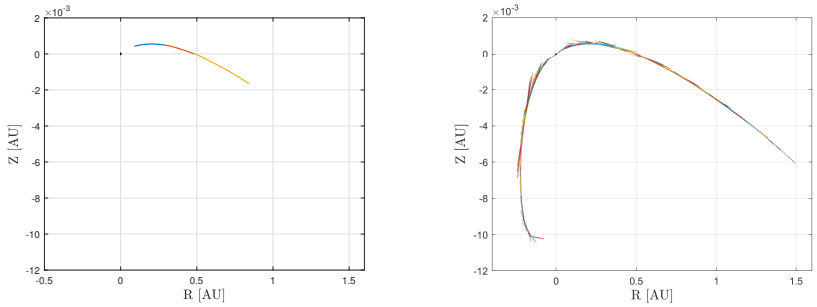
$$[\mathbf{x}_{\mathrm{EIG}}] = \begin{bmatrix} 0 & 0 & 0 & 0 & \Lambda_5 \cdot \delta x_5 & \Lambda_6 \cdot \delta x_6 \end{bmatrix}^T \tag{32}$$

Focusing on a $\pm 3\sigma$ analysis, the scaling factors are:

$$\Lambda_5 = 3\sqrt{\lambda_5} \qquad \text{and} \qquad \Lambda_6 = 3\sqrt{\lambda_6} \tag{33}$$

$[\mathbf{x}_{\mathrm{EIG}}]$ needs to be rotated in the real Cartesian space as in Eq. 14 and the result of this procedure is a 6D vector whose components are polynomials the two variables δx_5 and δx_6.

The two-dimensional confidence region stretches during the propagation and the projection onto the MTP, as shown in Fig. 14b for asteroid 2016 LP10. The subdomains are more subject to deformation after reaching $\mathcal{N}_{max}$. Consequently, for the sake of clarity, Fig. 14b is produced by evaluating a first order truncation of the 12th-order polynomials.



(a) *1D technique.* (b) *2D technique.*

Figure 14: 2016 LP10: 1D and 2D confidence region projected onto the MTP

The propagation of a multidimensional confidence region exacerbate the required computational effort. A larger number of subdomains are generated, typically in the order of thousands, and many of them reach the maximum number of splits. This has a significant effect on the reliability of the method. When a subdomain reaches $\mathcal{N}_{max}$, if the nonlinearities are large or if a long integration is performed, the accuracy of the final polynomial map is totally lost. The technique adopted to overcome this issue is presented in the followings.

During the propagation, the distance from the Earth is checked at the center of the subdomain and at the four boundary points identified by $(\delta x_5, \delta x_6) = (-1, 0), (1, 0), (0, -1), (0, 1)$. All the subdomains that have at least one point below the Earth surface are saved for further analysis. When all subdomains

have been propagated, the impacting ones are retrieved and, if they belong to the group of subdomains that reached $\mathcal{N}_{max}$, they are re-initialized and repropagated, starting from the integration step at which N_{max} was reached. This allows to obtain a further splitting only of the relevant boxes as depicted in Fig. 15.

Such repropagation is performed with a lower $\mathcal{N}_{max}$, that has to be tuned to compromise between accuracy and computational effort. In the multidimensional case, the subdomains are prone to many splits and may reach $\mathcal{N}_{max}$ even in the second propagation. Yet, their size would be way smaller than reasonably needed for an accurate IP computation and their propagation would require an extremely large computational effort. Instead, if the maximum number of splits is reduced up to 8 or 4, the computational time reduces to values in the order of minutes, making the approach feasible.

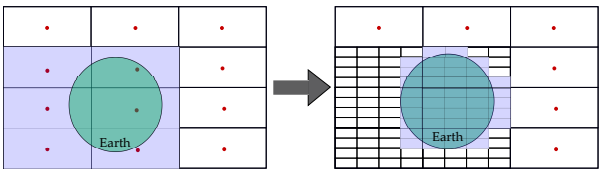


Figure 15: Repropagation of the impacting subdomains in the two-dimensional approach.

The IP computation for the 2D tool is performed by assigning a two-dimensional Gaussian pdf to the initial domain and integrating it over the impacting subdomains. The expression of a 2D pdf

$$\mathrm{pdf}(x_1, x_2) = \frac{1}{2\pi\sqrt{|\det\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{34}$$

is simplified thanks to the fact that the pdf is built on the space of the eigenvectors of the initial domain. Therefore, the covariance matrix $\boldsymbol{\Sigma}$ is diagonal and the pdf for the i-th subdomain becomes

$$\mathrm{pdf_i}(\delta_{1_i}, \delta_{2_i}) = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}} e^{-\frac{1}{2}\delta_{1_i}^2 3^2 - \frac{1}{2}\delta_{2_i}^2 3^2} \tag{35}$$

with the same nomenclature as in Eq. 28 and where 3 in the exponent accounts for the sigma-interval of the analysis: $[-3\sigma, +3\sigma]$.

On first approximation, all the impacting subdomains resulting from the first propagation are considered as full impactors. The 2D probability density is obtained by evaluating Eq. 35 at the center, i.e. $\delta_{1_i} = \delta_{2_i} = 0$ and their area is simply computed as:

$$\mathcal{A}_i = 3\sigma_1 3\sigma_2 w_{1_i} w_{2_i} \tag{36}$$

The IP is therefore computed as

$$IP = \sum_i \mathrm{pdf}_i \mathcal{A}_i \tag{37}$$

The result may be strongly affected by the fact that the impacting subdomains have reached $\mathcal{N}_{max}$ and therefore neither they can be reliably assumed as full impactors nor they can be used for a computation of the exact intersection. The repropagation of the impacting subdomains provides the necessary refinement. In this case, the impacting area is computed as:

$$\mathcal{A}_i = \mathcal{G}_i 3\sigma_1 3\sigma_2 w_{1_i} w_{2_i} \tag{38}$$

where $\mathcal{G}_i$ is a coefficient that accounts for the actual portion of the box that is identified to collide with the Earth after the repropagation. $\mathcal{G}_i$ is computed as

$$\mathcal{G}_i = \sum_j \frac{1}{2} w_{1_{i,j}} \frac{1}{2} w_{2_{i,j}} \tag{39}$$

where w_{1_j} and w_{2_j} are the widths in the two dimensions of the j-th sub-subdomain generated by the repropagation of the i-th subdomain.

The IP computation results are listed in Tab. 3. The repropagation of the impacting boxes, in this case with $\mathcal{N}_{max} = 4$, requires about 3 additional hours of computational time and it grants a much more accurate IP if compared to the reference value provided by NEODyS.

A critical role is played by the choice of $\mathcal{N}_{max}$ for the repropagation. In this case $\mathcal{N}_{max} = 4$ was a good choice but it may turn out to be low for other cases. In fact, when the IP is very small or the dynamics is highly nonlinear, the impactors may lie in a very tiny region of the impacting subdomains. Therefore, a higher value for $\mathcal{N}_{max}$ may be required.

	comp. time [h]	2D IP	$\text{IP}_{\text{NEODyS}}$
without repropagation	8.48	$5.8983 \cdot 10^{-4}$	$4.10 \cdot 10^{-5}$
with repropagation	11.51	$4.3003 \cdot 10^{-5}$	$4.10 \cdot 10^{-5}$

Table 3: 2016 LP10 IP computation

7.3. Intermediate eigenvalue ratio

As a final relevant scenario, let us consider the case of an intermediate value of χ, such as asteroid 2017 RH16 for which $\chi = 2.95 \cdot 10^3$. The initial conditions at 00:00 of March, 23$^{\text{rd}}$ 2018 include the possibility of an impact with Earth on 2026 with IP $= 1.45 \cdot 10^{-3}$, according to NEODyS. A one-dimensional and a two-dimensional propagation are performed and the impacting subdomains are identified. Fig. 16 shows the 2D impacting subdomanins (in blue) together with the LOV (the red line) and the 1D impactor (in black) on the plane of the first two eigenvectors. We can observe that the 2D impactors lie almost orthogonally to the LOV. This explains why the one-dimensional technique is capable of providing an accurate IP ($1.04 \cdot 10^{-3}$) even if most VIs lie outside the LOV. Indeed, if we integrate a one-dimensional pdf we are implicitely integrating the other dimensions from $-\infty$ to $+\infty$. More in detail, if we consider the 1D

impactor, this is equivalent to assuming as impactors all the points lying on the full line orthogonal to the LOV and crossing the line at the quote of the 1D impactor. We have seen in various cases that in this χ condition, the impactors lie almost perpendicularly to the LOV. Even when this is not strictly true, we can observe that the largest contribution to the IP value is given by the impactor on the LOV and, despite a slight deviation that might be accounted for, the two outputs turn out to be close. This justifies the adoption of a one-dimensional reduction of the problem.

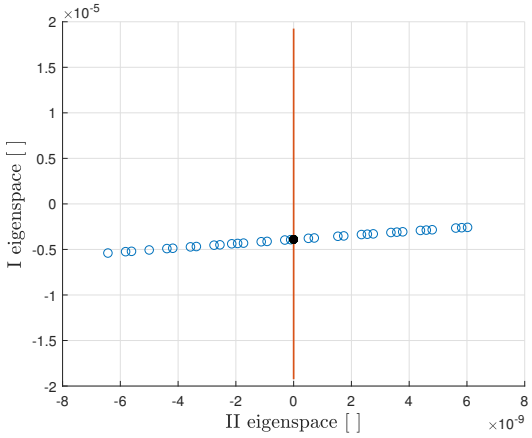


Figure 16: 1D and 2D impacting subdomains for 2017 RH16

8. Comparison with the existing IP computation methods

The technique investigated in this paper aims at obtaining an efficient and reliable IP computation. To assess whether the proposed technique does actually improve the currently adopted approaches, a comparison with the existing IP computation techniques is provided. All the following considerations apply to the one-dimensional technique.

Monte Carlo simulations provide a benchmark solution for any uncertainty propagation technique and, provided that a sufficient number of N-dimensionally distributed samples is generated, they offer a reliable result with a large computational burden. For the case of asteroid 2016 RD34, the longest time-per-sample required in a sample-wise propagation has been identified to be about 2.2 seconds on the computer and with the code adopted in this work. According to [14], the minimum number of samples required to detect any IP value above a certain IP_{lim} with a 50% relative error is $4/\mathrm{IP}_{lim}$. By considering that the impact probability is about $4 \cdot 10^{-4}$, the propagation of 10^4 samples guarantees a reliable IP computation. Yet, if the order of magnitude of the IP is not known a priori and if IP levels down to 10^{-6} are of interest, then the propagation of up to $4 \cdot 10^6$ samples is necessary. By multiplying this value by 2.2 seconds per

sample, the required simulation time turns out to be of the order of 100 days. For the sake of fairness, it is worth mentioning that the code adopted here for numerical propagation is not optimized for point-wise propagation and all the simulations are here performed on a single-core. Yet, it is worth observing that, under the same conditions the proposed approach provides a computational gain of three orders of magnitude with respect to standard Monte Carlo. The computational gain is even larger for asteroid 2010 RF12 due to the longer integration time span (which yields longer Monte Carlo simulation time) and the reduced DA-based computational time when the one-dimensional approach is sufficiently accurate.

An interesting comparison can be carried out with respect to the standard LOV approach. According to [11], a generic completion level of 10^{-7} requires the propagation of around 10^4 samples along the LOV (3-4 orders of magnitude lower than the samples required with standard Monte Carlo), which would require an IP computational time of about 1.5 to 2 hours per object on our platform and with our code. The analyses reported in this work have shown that the proposed approach yield accurate IP value in less than 20 minutes (see 2010 RF12). According to this considerations, the proposed approach seems to be competitive with respect to current techniques. The main issue is, at the moment being, the lower level of reliability of the algorithm. On a general level, the proposed approach could be adopted for preliminary analyses. All the detected objects with $\chi > 10^2$ can be tested with the DA-based one-dimensional approach and, depending on the results, different strategies may be adopted:

- IP $\neq 0$: the result can be considered reliable.
- IP $= 0$, $\chi < 10^4$: a different technique analysis shall be applied.
- IP $= 0$, $\chi > 10^4$ and very low minimum Earth-LOV distance: a different technique shall be applied to exclude the possibility of low-probability off-LOV impactors.
- IP $= 0$, $\chi > 10^4$ with safe minimum Earth-LOV distance: the approach can be considered reliable.

Since most NEOs fall in the first or fourth category, only few objects be analyzed with different techniques.

9. Conclusion

This paper has introduced DA techniques into the standard LOV approach for NEO impact probability computation. The one-dimensional technique, if supported by the repropagation algorithm, provides reliable results if the eigenvalue ratio is sufficiently large. Computational times are relatively low, ranging from few minutes to few hours in the most demanding cases. For low eigenvalue ratios, the 2D approach has to be adopted, with increased computational time and sensitivity to the tuning parameters

The 1D approach could be a valuable tool, for its capability of providing a fast and reliable IP computation. The majority of asteroids has a large χ value and therefore this represents a significant achievement. Further work is demanded, in particular to develop an autonomous version of the software, capable of working without an estimated collision epoch. This aspect is certainly non-trivial and is crucial in routine impact monitoring operations.

As a final remark, it would be interesting to assess the potential advantages offered by code parallelization. Indeed, while working with ADS, the different subdomains could be propagated independently thus reducing the computational time.

References

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

[2] S. R. Chesley, P. W. Chodas, Asteroid close approaches: analysis and potential impact detection, Asteroids III 55 (2002).

[3] A. Milani, M. E. Sansaturio, G. Tommei, O. Arratia, S. R. Chesley, Multiple solutions for asteroid orbits: computational procedure and applications, Astronomy & Astrophysics 431 (2) (2005) 729–746 (2005).

[4] A. Milani, G. B. Valsecchi, The asteroid identification problem: Ii. target plane confidence boundaries, Icarus 140 (2) (1999) 408–423 (1999).

[5] D. K. Yeomans, P. W. Chodas, Predicting close approaches of asteroids and comets to earth, Hazards due to Comets and Asteroids (1994) 241–258 (1994).

[6] J. F. Ritt, Differential equations from the algebraic standpoint, Vol. 14, American Mathematical Soc., 1932 (1932).

[7] M. Berz, Advances in imaging and electron physics, Academic Press, 1999 (1999).

[8] P. Di Lizia, R. Armellin, M. Lavagna, Application of high order expansions of two-point boundary value problems to astrodynamics, Celestial Mechanics and Dynamical Astronomy 102 (4) (2008) 355–375 (2008).

[9] A. Wittig, P. Di Lizia, R. Armellin, K. Makino, F. Bernelli-Zazzera, M. Berz, Propagation of large uncertainty sets in orbital dynamics by automatic domain splitting, Celestial Mechanics and Dynamical Astronomy 122 (3) (2015) 239–261 (2015).

[10] M. Rasotto, A. Morselli, A. Wittig, M. Massari, P. Di Lizia, R. Armellin, C. Valles, G. Ortega, Differential algebra space toolbox for nonlinear uncertainty propagation in space dynamics (2016).

[11] A. Milani, S. R. Chesley, M. E. Sansaturio, G. Tommei, G. B. Valsecchi, Nonlinear impact monitoring: line of variation searches for impactors, Icarus 173 (2) (2005) 362–384 (2005).

[12] M. Valli, R. Armellin, P. Di Lizia, M. Lavagna, Nonlinear mapping of uncertainties in celestial mechanics, Journal of Guidance, Control, and Dynamics 36 (1) (2012) 48–63 (2012).

[13] F. Spoto, A. Del Vigna, A. Milani, G. Tommei, P. Tanga, F. Mignard, B. Carry, W. Thuillot, P. David, Short arc orbit determination and imminent impactors in the gaia era, Astronomy & Astrophysics 614 (2018) A27 (2018).

[14] D. Farnocchia, S. Chesley, A. Milani, G. Gronchi, P. Chodas, Orbits, long-term predictions, impact monitoring, Asteroids IV (2015) 815–834 (2015).