ADVANCED MONTE CARLO SAMPLING TECHNIQUES FOR ORBITAL CONJUNCTIONS ANALYSIS AND NEAR EARTH OBJECTS IMPACT PROBABILITY COMPUTATION

M. Losacco[(1)], M. Romano[(1)], P. Di Lizia[(1)], C. Colombo[(1)], R. Armellin[(2)], A. Morselli[(3)], and J.M. Sánchez Pérez[(3)]

[(1)]*Department of Aerospace Science and Technology, Politecnico di Milano, 20156, Milano, Italy, E-mail: {matteo.losacco, matteo1.romano, pierluigi.dilizia, camilla.colombo}@polimi.it*
[(2)]*Surrey Space Centre, University of Surrey, GU2 7XH, Guildford, United Kingdom, E-mail: r.armellin@surrey.ac.uk*
[(3)]*European Space Operations Centre (ESOC/ESA), 64293, Darmstadt, Germany, E-mail: {alessandro.morselli, jose.manuel.sanchez.perez}esa.int*

ABSTRACT

This paper presents a survey of past and new results of the application of advanced sampling techniques for orbital conjunction analysis and Near Earth Objects impact probability computation. The theoretical background of the methods is presented, along with the results of their applications and a critical discussion of the benefits introduced.

Keywords: Space Debris; Conjunction Analysis; Near Earth Objects; Impact Probability; Monte Carlo.

1. INTRODUCTION

When dealing with space debris and Near Earth Objects (NEO), the major threat for human activities in space and global safety is represented by possible collisions. The risk of in-orbit collisions between operative satellites and space debris, indeed, is a crucial issue in satellite operation. When a close approach is identified, it is necessary to define an indicator that can reasonably tell how risky the predicted conjunction may be, and a common practice for space agencies and satellite operators is to consider the collision probability for this purpose, together with conjunction geometry and miss-distance [15]. On the other hand, whenever a new asteroid is discovered, it is of crucial importance to have accurate tools for detection and prediction of possible impacts. The task introduces relevant challenges due to the imperative of early detection and propagation of its state and associated uncertainty, and the problem is made even more complicated by the fact that the dynamics describing the motion of these objects is highly nonlinear, especially during close encounters with major bodies.

Present day approaches for orbital conjunctions analysis and robust detection and prediction of planetary encounters and potential impacts by NEO include simplified models or full nonlinear approaches. Simplified models significantly reduce the required computational burden, but their range of application is very narrow. On the other hand, full nonlinear models, i.e. canonical Monte Carlo (MC) simulations, represent a general and flexible way of approaching collision and impact probability estimation, but they involve unavoidable intensive computations, which may be not suitable for satellite-debris daily collision probability computation or fast NEO impact probability estimation. An elegant and effective compromise was introduced by Milani, based on the concept of the Line Of Variations (LOV) [18].

In recent times, new advanced Monte Carlo techniques, such as the Importance Sampling (IS), Line Sampling (LS) and Subsets Simulation (SS) methods were developed, aiming at reducing the computational burden by either restricting the sampling phase space or identifying optimal sampling paths within it. In parallel, an intense study on alternative uncertainty propagation methods was carried out, and innovative approaches based on the use of differential algebra (DA) were proposed. This paper presents a series of combinations and applications of the mentioned techniques to orbital conjunctions analysis and NEO impact probability computation. The problem of orbital conjunctions is faced combining the mentioned SS and LS techniques with DA, obtaining a high-order, efficient tool that can be applied to any set of initial states and using any arbitrary reference frame. The challenging task of NEO impact probability computation is faced both directly applying the SS and LS techniques and combining the IS method with DA. The presented approaches offer competitive results in both fields, providing a significant improvement with respect to standard MC performance while maintaining a sufficiently high level of accuracy. For all the approaches, test cases are presented, and a detailed analysis of the performance of the proposed methods is offered.

The paper is organized as follows. First, a detailed theoretical description of differential algebra, Importance Sampling, Line Sampling and Subset Simulation is of-

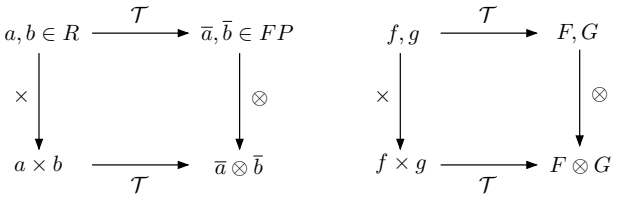

Figure 1: Analogy between the FP representation of real numbers in computer environment (left) and the algebra of Taylor polynomials in DA framework (right)[12].

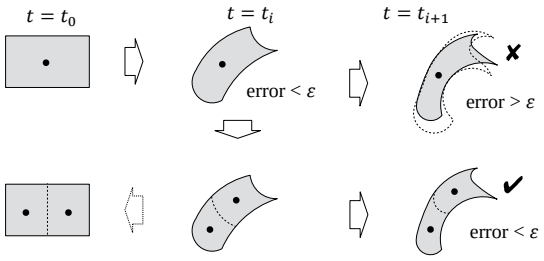

Figure 2: ADS algorithm schematic illustration [28].

fered. Then, the results of their application to orbital conjunction analysis and NEO impact probability computation are presented, along with a critical discussion of benefits and limitations of the presented approaches.

2. THEORETICAL BACKGROUND

A detailed description of the methods presented in the article is offered in this section. The differential algebra approach and its recent enhancements are presented in the first part. Then, three advanced orbital sampling techniques are described: Importance Sampling, Line Sampling and Subset Simulation.

2.1. Differential Algebra and Automatic Domain Splitting

Differential algebra provides the tools to compute the derivatives of functions within a computer environment [25, 26, 23, 24, 16, 7]. The basic idea of DA is to bring the treatment of functions and the operations on them to the computer environment in a similar way as the treatment of real numbers [7]. Real numbers, indeed, are approximated by floating point (FP) numbers with a finite number of digits. In a similar way, suppose two sufficiently regular functions f and g are given. In the framework of DA, these functions are converted into their Taylor series expansions, F and G respectively (see Fig. 1). For each operation in the function space, an adjoint operation in the space of Taylor polynomials is defined. As a results, the implementation of DA in a computer environment provides the Taylor coefficients of any function of v variables up to a specific order n: any function f of v variables can be expanded into its Taylor expansion up to an arbitrary order n, along with the function evaluation, with a limited amount of effort. The Taylor coefficients of order n for sum and product of functions, as well as scalar products with real numbers, can be directly computed from those of summands and factors. In addition to basic algebraic operations, differentiation and integration can be easily introduced in the algebra, thus finalizing the definition of the differential algebra structure of DA [5, 6]. The DA used in this work is implemented in the DACE software [22].

A relevant application of DA is the automatic high order expansion of the solution of an Ordinary Differential Equation (ODE) with respect to the initial conditions [7, 13, 22]. This expansion can be achieved by considering that any integration scheme, explicit or implicit, is characterized by a finite number of algebraic operations, involving the evaluation of the ODE right hand side (RHS) at several integration points. By replacing the operations between real numbers with those on DA numbers, the nth order Taylor expansion of the flow of the ODE, $\phi(t; \delta x_0, t_0) = \mathcal{M}_\phi(\delta x_0)$, at each integration time, for any perturbed initial condition $x_0 + \delta x_0$ can be obtained.

While a single polynomial map can accurately describe the evolution of an uncertainty set for short term propagation or dynamics characterized by a low level of nonlinearities, which is a common case of orbital conjunctions, the accuracy of a single polynomial maps drastically decreases as the propagation window increases or the effect of nonlinearities becomes more important. This is the typical case of NEO long term propagations. The approximation error of the DA representation is strictly related to the size of the domain the polynomial is defined in [28]. This means that, if one divides the initial domain into smaller domains and then computes the Taylor expansion around the center points of the new domains, the error can be reduced, still covering the entire initial set with all the generated polynomial expansions. Starting from these considerations, Automatic Domain Splitting (ADS) employs an automatic algorithm to determine at which time t_i the flow expansion over the set of initial conditions is no longer able to describe the dynamics with enough accuracy [28]. Once this event has been detected, the domain of the original polynomial expansion is divided along one of the expansion variables into two domains of half their original size. By re-expanding the polynomials around the new centre points, two separate polynomial expansions are obtained. In this way, the splitting procedure guarantees a more accurate description of the whole uncertainty set at the current time epoch t_i. After such a split occurs, the integration process is resumed on both generated subsets, until new splits are required. A representation of the ADS procedure is shown in Fig. 2. The main degrees of freedom of the algorithm are the tolerance for the splitting procedure and the maximum number of splits per subset. A detailed description of the role of these parameters and ADS is offered in [28] and [17].

2.2. Advanced orbital sampling techniques

Three advanced Monte Carlo sampling techniques are presented in this section: Importance Sampling, Line Sampling and Subset Simulation. All presented methods aim at identifying the failure region of an uncertainty set with the minimum amount of samples. For the problem under study, failure is represented either by a collision in orbital conjunction analysis or an impact for NEO impact monitoring. The occurrence of the events is expressed by the value of the performance function

$$g_{\boldsymbol{x}}(\boldsymbol{x}_0) = d(\boldsymbol{x}_0) - D \tag{1}$$

where $\boldsymbol{x}_0$ represents a single initial condition for NEO impact monitoring, a couple $\boldsymbol{x}_0 = (\boldsymbol{x}_0^1, \boldsymbol{x}_0^2)$ for conjunction analysis, d is a function that maps an initial condition $\boldsymbol{x}_0$ to a performance index, and D is the selected critical distance. For conjunction analysis, the performance index is the Distance of Closest Approach (DCA). For impact monitoring, it is the minimum planetocentric distance. The critical distance is a threshold beyond which a "failure" is identified. For conjunction analysis, this threshold can be identified as half of the diameter of the sphere including both objects [20]. For impact monitoring, D represents the planet radius. According to this definition, it follows that

$$g_{\boldsymbol{x}}(\boldsymbol{x}_0) \quad \begin{cases} < 0 \rightarrow \text{failure} \\ = 0 \rightarrow \text{limit state} \\ > 0 \rightarrow \text{no failure} \end{cases} \tag{2}$$

2.2.1. Importance Sampling

The Importance Sampling (IS) method amounts to replacing the original probability density function (pdf) $q_{\boldsymbol{x}}(\boldsymbol{x})$ with an Importance Sampling Distribution (ISD) $\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x})$ arbitrarily chosen by the analyst so as to generate a large number of samples in the importance region of the phase space F. This region represents the set of initial conditions leading to collisions in orbital conjunction analysis or impacts for NEO impact monitoring. The general IS algorithm is the following:

1. Identify a proper $\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x})$.
2. Express the failure probability $P(F)$ as a function of $\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x})$.

$$\begin{aligned} P(F) &= \int I_{\mathrm{F}}(\boldsymbol{x}) q_{\boldsymbol{x}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \\ &= \int \frac{I_{\mathrm{F}}(\boldsymbol{x}) q_{\boldsymbol{x}}(\boldsymbol{x})}{\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x})} \tilde{q}_{\boldsymbol{x}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \end{aligned} \tag{3}$$

   where $I_{\mathrm{F}}(\boldsymbol{x}) : \mathbb{R}^v \rightarrow \{0, 1\}$ is an indicator function such that $I_{\mathrm{F}}(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in F$, 0 otherwise.
3. Draw N_{T} samples $\boldsymbol{x}^k : k = 1, 2, \ldots, N_{\mathrm{T}}$ from the importance sampling distribution $\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x})$. If a good choice for the auxiliary pdf is made, the generated samples concentrate in the region F.
4. Compute the estimate $\hat{P}(F)$ for the failure probability $P(F)$ by resorting to equation (3):

$$\hat{P}(F) = \frac{1}{N_{\mathrm{T}}} \sum_{k=1}^{N_{\mathrm{T}}} \frac{I_{\mathrm{F}}(\boldsymbol{x}^k) q_{\boldsymbol{x}}(\boldsymbol{x}^k)}{\tilde{q}_{\boldsymbol{x}}(\boldsymbol{x}^k)} \tag{4}$$

5. Compute the variance of the estimator $\hat{P}(F)$ as:

$$\begin{aligned} \sigma^2(\hat{P}(F)) = \frac{1}{N_{\mathrm{T}}} \Bigg(\int \frac{I_{\mathrm{F}}^2(\boldsymbol{x}) q_{\boldsymbol{x}}^2(\boldsymbol{x})}{\tilde{q}_{\boldsymbol{x}}^2(\boldsymbol{x})} \tilde{q}_{\boldsymbol{x}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \\ - P^2(F) \Bigg) \approx \frac{1}{N_{\mathrm{T}}} \left(\widehat{P^2(F)} - \hat{P}^2(F) \right) \end{aligned} \tag{5}$$

The selection of the ISD represents the most critical point for the method. Several techniques have been developed in order to find the one giving small variance for the estimator [30]. As later described in Sec. 4, this sampling method can be combined with DA for NEO impact monitoring, and the ISD can be modelled according to the results of the DA propagation.

2.2.2. Line Sampling

The Line Sampling (LS) method is a Monte Carlo-based approach for the estimation of small probabilities. The main idea behind LS is to transform a high-dimensional problem into a number of conditional one-dimensional problems solved along an important direction $\boldsymbol{\alpha}$. The LS method follows four steps: 1) the mapping of random samples from the physical coordinate space into a normalized standard space, 2) the determination of the reference direction $\boldsymbol{\alpha}$, 3) the probing of the failure region along lines following the reference direction $\boldsymbol{\alpha}$, 4) the estimation of the failure probability.

In the LS approach, a vector of uncertain parameters $\boldsymbol{x}_0 \in \mathbb{R}^n$ is first transformed into the adjoint vector $\boldsymbol{\theta} \in \mathbb{R}^n$ belonging to the so-called "standard normal space", where each variable is represented by an independent central unit Gaussian distribution. This is done by relying on Rosenblatt's transformation [27]

$$\begin{aligned} \boldsymbol{\theta} &= T_{\boldsymbol{x},\boldsymbol{\theta}}(\boldsymbol{x}_0) \\ \boldsymbol{x}_0 &= T_{\boldsymbol{\theta},\boldsymbol{x}}(\boldsymbol{\theta}) \end{aligned} \tag{6}$$

A natural choice for the important direction is the normalized gradient of the performance function at the nominal point in the standard normal space

$$\boldsymbol{\alpha} = \frac{\nabla_{\boldsymbol{\theta}} g_{\boldsymbol{\theta}}(\boldsymbol{\theta})}{||\nabla_{\boldsymbol{\theta}} g_{\boldsymbol{\theta}}(\boldsymbol{\theta})||_2} \tag{7}$$

If not available analytically, the gradient can be estimated numerically. Another approach consists in obtaining an estimate by computing the normalized centre of mass of the failure domain. This is achieved by Monte Carlo

Markov Chain (MCMC), using as seed a point belonging to the failure region or close to it, and computing the mean of the N_s samples generated in the failure region. Once the important direction is identified, the LS method proceeds as follows [20]:

- Sample N_T vectors $\boldsymbol{\theta}$ from the normal multidimensional joint probability distribution.
- For each sample i, estimate its conditional one-dimensional failure probability $\hat{P}^i$. This task is solved performing the following operations (see also Fig. 3)
  - Project the vector $\boldsymbol{\theta}^i$ onto the straight line passing through the origin and perpendicular to $\boldsymbol{\alpha}$ to obtain the vector $\boldsymbol{\theta}^{i,\perp}$.
  - Write the parametric equation of samples along the important direction $\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i,\perp} + c\,\boldsymbol{\alpha}$
  - Compute the values of c_j^i, with $j = 1, 2$, for which the performance function is equal to zero. This step requires the evaluation of the performance function, which involves extra numerical propagations.
  - If the two values are coincident or no solution is found, then the ith one-dimensional probability $\hat{P}^i$ is equal to zero, otherwise, given the two solutions c_1^i and c_2^i, the probability becomes

$$\hat{P}^i(F) = P[c_2^i \leq N(0,1) \leq c_1^i] = \\ = \Phi(c_1^i) - \Phi(c_2^i)$$

    where $\Phi(c_j^i)$ is the standard normal cumulative distribution function, $N(0,1)$ is the standard normal distribution, with zero mean and unit standard deviation, and F is the failure event (in-orbit collision for conjunction analysis or NEO impact for impact monitoring).
- Compute the unbiased estimator $\hat{P}^{N_\mathrm{T}}(F)$ as the sample average of the independent conditional one-dimensional probability estimate

$$\hat{P}^{N_\mathrm{T}}(F) = \frac{1}{N_\mathrm{T}} \sum_{i=1}^{N_\mathrm{T}} \hat{P}^i(F) \tag{8}$$

  The variance of the estimator is given by

$$\sigma^2(\hat{P}^{N_\mathrm{T}}(F)) = \frac{1}{N_\mathrm{T}(N_\mathrm{T}-1)} \sum_{i=1}^{N_\mathrm{T}} (\hat{P}^i(F) - \hat{P}^{N_\mathrm{T}}(F))^2 \tag{9}$$

2.2.3. Subset Simulation

The Subset Simulation (SS) method is a Monte Carlo method based on the principle of computing small failure probabilities as the product of larger conditional probabilities [3, 8, 32]. Given a target failure event F, let $F_1 \supset F_2 \supset \ldots \supset F_n = F$ be a sequence of intermediate failure events, so that $F_k = \cap_{i=1}^k F_i$. Considering a sequence of conditional probabilities, then the failure probability can be written as

$$P(F_n) = P(F_1) \prod_{i=1}^{n-1} P(F_{i+1}|F_i) \tag{10}$$

where $P(F_{i+1}|F_i)$ represents the probability of F_{i+1} conditional to F_i.

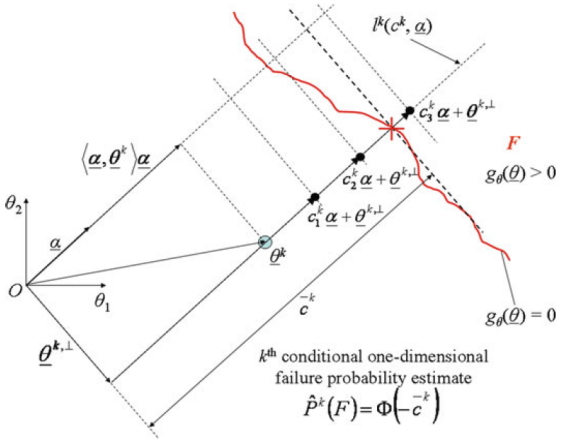

Figure 3: Scheme of the iterative procedure used to sample each line in the standard normal coordinate space. The failure region is labeled with F, with a single border highlighted as a red line (image courtesy of [31]).

The method is initialized using standard Monte Carlo to generate samples at the so called conditional level 0 (CL 0) starting from the available nominal state vectors and related uncertainties of the investigated objects. The number of samples generated at this level is maintained for each generated conditional level and it is referred to as N. Once the failure region F_1 is identified, a MCMC Metropolis Hastings algorithm is used to generate conditional samples in the identified intermediate failure region. Another intermediate region is then located, and other samples are generated by means of MCMC. The procedure is repeated until the failure region is identified. A scheme of the method is shown in Fig 4. In the presented approach, the intermediate failure regions are identified by assuming a fixed value of conditional probability $P(F_{i+1}|F_i) = p_0$. The identification of each conditional level, therefore, is strictly related to this value, and changes accordingly step by step, as explained in the followings. The resulting Subset Simulation algorithm follows the general description offered in [20] and goes through the following steps:

1. Set $i = 0$ and generate N samples $\boldsymbol{x}_0^{0,k}$, $k = 0, \ldots, N$ at conditional level 0 by standard Monte Carlo.

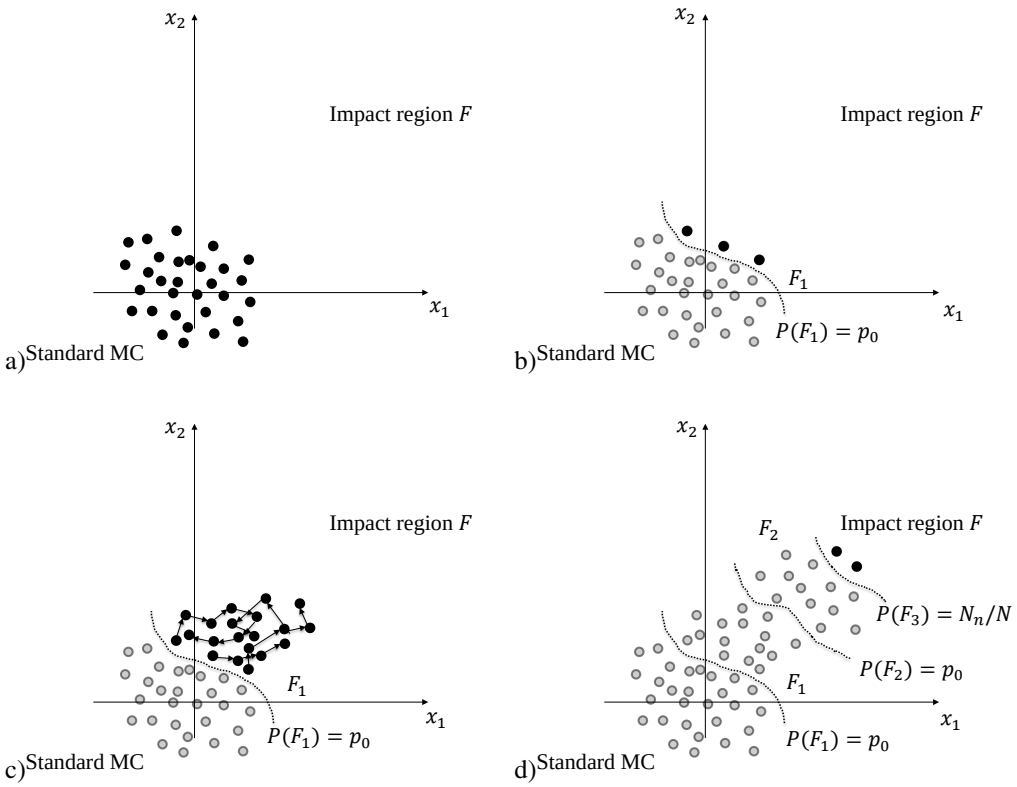

Figure 4: Subset Simulation process: a), initialization by standard MC, b), CL 1 identification, c), samples generation by means of MCMC, d), new iterations and impact region identification (image courtesy of [17]).

2. Propagate each sample for the selected time window and compute its performance function, defined as

$$g^i_{\boldsymbol{x}}(\boldsymbol{x}_0) = d(\boldsymbol{x}_0) - D_i \quad \begin{cases} < 0 \rightarrow \boldsymbol{x}_0 \in (i)\text{th CL} \\ > 0 \rightarrow \boldsymbol{x}_0 \notin (i)\text{th CL} \end{cases} \tag{11}$$

where D_i is the current threshold, whereas d is a function that provides, for each sample, its performance index. For conjunction analysis, the index is the Distance of Closest Approach, for NEO impact probability computation, this index represents the minimum planetocentric distance.

3. Sort the N samples in descending order according to their associated value of performance function.

4. Identify an intermediate threshold value D_{i+1} as the $(1-p_0)N$th element of the list. Define the $(i+1)$th conditional level as $F_{i+1} = \{d < D_{i+1}\}$. Considering how the threshold was defined, the associated conditional probability $P(F_{i+1}|F_i) = p_0$

5. If $D_{i+1} < D$, go the last step, otherwise select the last $p_0 N$ samples of the list $x_0^{i,j}$, $j = 0, ..., p_0 N$. By definition, these samples belong to the $(i+1)$th conditional level.

6. Using MCMC, generate $(1-p_0)N$ additional conditional samples starting from the previously selected seeds belonging to F_{i+1}. A sample is set to belong to F_{i+1} according to its performance function value.

7. Set $i = i + 1$ and return to step 2

8. Stop the algorithm.

The total number of generated samples is

$$N_{\text{T}} = N + (n-1)(1-p_0)N \tag{12}$$

where n is the overall number of conditional levels required to reach the failure region. Since the conditional probability is equal to p_0 for each level, the failure probability becomes:

$$\begin{aligned} P(F) = P(F_n) = P(F_n|F_{n-1})p_0^{n-1} = \\ = p_0^{n-1} N_{\text{F}}/N \end{aligned} \tag{13}$$

where N_{F} is the number of samples belonging to the failure region.

A Bayesian post-processor for SS (SS$^+$) is suggested in [32] to refine the computed failure probability and determine higher moments. If we define

$$n_l = \begin{cases} p_0 N & \text{if } l < n \\ N_{\text{F}} & \text{if } l = n \end{cases} \tag{14}$$

Table 1: Time, distance and velocity of closest approach for the Keplerian test cases and reference values for collision probability [20].

Test case	TCA (days)	DCA (m)	Δv_{TCA} (m/s)	D (m)	$\hat{P}_{\text{c}}$ (MC)	N_{T} [11]	$\hat{P}_{\text{c}}$ [2]	% err (-)
5	2.0	2.449	0.520	10	4.454e-2	2.30e6	4.440e-2	-0.32%
6	2.0	2.449	0.173	10	4.340e-3	2.50e7	4.324e-3	-0.36%
7	2.0	3.183	0.196	10	1.614e-4	6.71e8	1.580e-4	-2.13%

the first moment of the distribution of the failure probability becomes

$$E_{\text{SS}^+}[P] = \prod_{l=1}^{n} \frac{n_l + 1}{N + 2} \tag{15}$$

whereas the second moment is expressed by

$$E_{\text{SS}^+}[P^2] = \prod_{l=1}^{n} \frac{(n_l + 1)\,(n_l + 2)}{(N + 2)(N + 3)} \tag{16}$$

The variance of the estimator, therefore, can be computed as

$$\sigma^2(P) = E[P^2] - (E[P])^2 \tag{17}$$

Equations 15 and 17 will represent our reference for the analyses presented in this paper.

3. ORBITAL CONJUNCTION ANALYSIS

The estimation of the collision probability for in-orbit conjunction analysis requires the computation of a multivariate integral. Different methods exist for the computation of this multi-dimensional integral[1, 4, 21]. Most of these approaches rely on strong assumptions, assuming uncorrelated position uncertainties for the two objects, objects moving along straight lines at constant velocity during the conjunction, negligible uncertainties in velocity, and constant, multi-dimensional Gaussian uncertainties in position during the whole encounter for both objects. These assumptions produce accurate results when the relative motion between the satellite and the object is rectilinear and the conjunction occurs close to the initial epoch.

Methods that account for nonlinearities, which are usually relevant for GEO conjunctions, typically rely on analytical methods or Monte Carlo simulations. Despite being a general and flexible way to compute collision probability, the MC approach has the main drawback of requiring intensive computation. For this reason, MC methods are not suitable for daily collision probability computation, since the results can be obtained in a timely manner only relying on simplified orbital dynamics, such as two body propagators or SGP4/SDP4.

The approach illustrated in this paper was first presented in [20] and combines the previously described DA method with standard MC, LS and SS, obtaining the DAMC, DALS and DASS methods for orbital conjunction analysis. The basic idea is to use DA to obtain the Taylor expansion of the Time of Closest Approach (TCA) and the Distance of Closest Approach (DCA) of the two orbiting objects. The occurrence of close approaches is first identified using the nominal initial orbital states. Then, DA is used to propagate sets of initial conditions by computing the Taylor expansion of the final states $\boldsymbol{x}_f^1$ and $\boldsymbol{x}_f^2$ at the nominal TCA. The resulting polynomials are functions of both the final time and the initial uncertain state vectors $\boldsymbol{x}_0^1$ and $\boldsymbol{x}_0^2$. The polynomial map of the relative distance between the two objects is then computed through simple algebraic manipulations. By using partial polynomial inversion techniques and imposing the stationarity condition of the relative distance with respect to time, the Taylor expansions of TCA and DCA with respect to $\boldsymbol{x}_0^1$ and $\boldsymbol{x}_0^2$.

$$\begin{aligned} [t^*] &= t^* + \mathcal{M}_{t^*}(\delta\boldsymbol{x}_0^1, \delta\boldsymbol{x}_0^2) \\ [d^*] &= d^* + \mathcal{M}_{d^*}(\delta\boldsymbol{x}_0^1, \delta\boldsymbol{x}_0^2) \end{aligned} \tag{18}$$

are computed. Given any perturbed initial condition of the two objects $(\delta\boldsymbol{x}_0^1, \delta\boldsymbol{x}_0^2)$, the evaluation of the two Taylor polynomials provides the values of the TCA and DCA. This enables a drastic reduction of the computational cost by replacing standard propagations with multiple polynomial evaluations. At this stage, the sampling is performed relying either on MC, LS or SS, taking advantage of the availability of the resulting polynomial maps. In particular, for the DALS method, the availability of the DCA expansion is exploited for the computation of the important direction $\boldsymbol{\alpha}$, whereas all numerical propagations of the standard algorithm are replaced by numerical evaluations. A detailed explanation is offered in [20]. For DAMC and DASS, the advantages are limited to the replacement of numerical propagations.

We present here the results of the application of the proposed methods for three different test cases taken from [2], where simple Keplerian dynamics is used for performance comparison. Results and figures are taken from [20]. The comparison is done considering the number of propagated samples N_{T}, the standard deviation of the estimated collision probability σ, the coefficient of variations $\delta = \sigma/\hat{P}$, and the figure of merit Δ, defined as

$$\Delta = \delta\sqrt{N_{\text{T}}} \tag{19}$$

The selected figure of merit does not depend on the number of samples, since for Monte Carlo methods $\sigma \propto$

Table 2: Computed collision probability for the Keplerian test cases. The comparison is shown in terms of relative error with respect to reference $\hat{P}_c$, number of samples N_T, computational time t_{CPU}, coefficient of variations δ and figure of merit Δ [20].

Test case	Method	$\hat{P}_c$ (-)	% err (-)	N_T	t_{CPU} (s)	δ (-)	Δ (-)
5	MC	4.452e-2	-0.05%	1.0e5	4.75	1.465e-2	4.63
	DAMC	4.459e-2	+0.11%	1.0e5	0.67	1.464e-2	4.63
	DALS	4.451e-2	-0.07%	5.0e3	2.53	7.662e-4	0.05
	DASS	4.450e-2	-0.09%	2.0e4	0.13	2.738e-2	3.87
6	MC	4.339e-3	-0.01%	1.0e6	43.21	1.515e-2	15.14
	DAMC	4.350e-3	+0.24%	1.0e6	6.67	1.513e-2	15.13
	DALS	4.341e-3	+0.03%	5.0e3	2.58	1.484e-3	0.11
	DASS	4.328e-3	-0.27%	4.0e4	0.27	3.586e-2	7.17
7	MC	1.615e-4	+0.04%	2.7e7	1155.36	1.514e-2	78.68
	DAMC	1.612e-4	-0.15%	2.7e7	179.34	1.516e-2	78.76
	DALS	1.621e-4	+0.41%	5.0e3	1.43	1.936e-2	1.37
	DASS	1.626e-4	+0.72%	6.0e4	0.43	4.580e-2	11.22

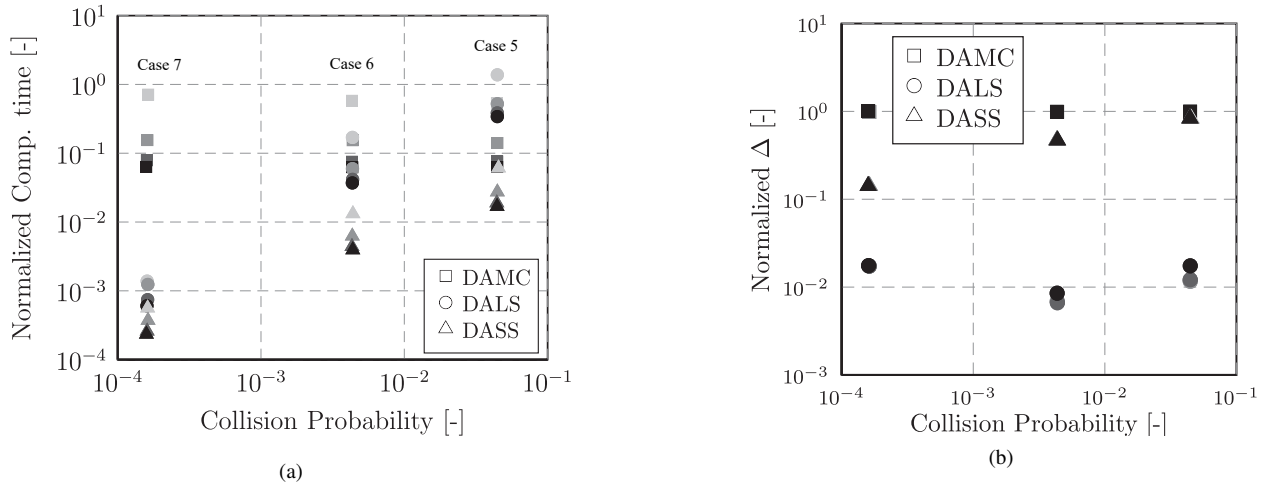

Figure 5: Normalized computational time t_{CPU} (a) and figure of merit Δ (b) of DAMC, DALS and DASS for the three considered test cases vs collision probability for different expansion orders. Markers are colored using a grey scale, where black is used for the expansion order $k = 1$ and light gray for $k = 4$ [20].

$1/\sqrt{N_T}$. It is designed to enable the comparison of different methods in terms of accuracy and number of samples required to reach that accuracy level. The three considered test cases are labeled as test case 5, 6 and 7, and are characterized by linear relative motion between the two objects, motion at the boundary of linear relative motion and nonlinear relative motion respectively.

Table 1 shows the characteristics of the three selected cases in terms of time, distance, and relative velocity at the closest approach Δv_{TCA}, along with the reference value for the collision probability as obtained with standard Monte Carlo simulation. The value of N_{T} used for MC simulations was selected according to [11]. For each trial, two sets of initial conditions are sampled from each initial covariance matrix and the associated DCA is computed in the proximity of the nominal TCA. The last two columns of the table show the collision probability obtained using Alfano's formula and its associated percentage relative error to the reference.

Table 2 shows the results obtained with the three proposed DA-based approaches. The results are expressed in terms of expected collision probability $\hat{P}_{\text{c}}$, relative error with respect to reference, number of required samples N_{T}, computational time t_{CPU}, coefficient of variation δ and figure of merit Δ. For all simulations, an expansion order k for DA propagation equal to 3 was considered. As can be seen, for all three test cases, the computed collision probability values are in good agreement with the

Table 3: Nominal equinoctial parameters and related uncertainty covariance matrix for asteroid 2017 RH_{16} at 6475 MJD2000 (September 24, 2017).

	a (AU)	P_1 (-)	P_2 (-)	Q_1 (-)	Q_2 (-)	l (deg)
State	0.8752	-0.1867	-0.4014	-0.0020	0.0050	319.9653
Cov.	9.7910e-08	1.2616e-08	3.6644e-08	-1.1067e-09	-2.2278e-10	-2.2745e-06
	1.2616e-08	1.8745e-09	5.2329e-09	-1.3999e-10	-3.7340e-11	-2.1989e-07
	3.6644e-08	5.2329e-09	1.4767e-08	-4.0882e-10	-1.0114e-10	-7.0073e-07
	-1.1067e-09	-1.3999e-10	-4.0882e-10	1.2541e-11	2.4080e-12	2.6480e-08
	-2.2278e-10	-3.7340e-11	-1.0114e-10	2.4080e-12	9.2142e-13	2.6346e-09
	-2.2745e-06	-2.1989e-07	-7.0073e-07	2.6480e-08	2.6346e-09	7.4365e-05

reference values. A first consideration on the use of differential algebra can be made by comparing the performance of MC and DAMC for all three selected test cases. By analyzing the required computational time, indeed, it is evident that, as the expected probability decreases, the required number of samples increases, and the advantages guaranteed by the introduction of the polynomial map become more and more evident. A second consideration can be done by comparing the results of DAMC with DALS and DASS. As can be seen, for all three test cases, both methods guarantee good accuracy levels with a number of samples that is at least one order of magnitude lower than what required by MC and DAMC. This difference increases as the expected collision probability decreases. This fact has obviously a direct positive consequence on the value of the figure of merit Δ.

The performance of all three DA-based methods is strongly influenced by the selected expansion order. As the expansion order increases, the accuracy obtained in the definition of the polynomial maps increases, but the required computational time increases as well. Figure 5a shows the computational time required by the three DA-based methods, normalized with respect to the computational time of standard MC, as a function of the collision probability, for expansion orders from 1 to 4. A greyscale is used, with black for $k = 1$ and light gray for $k = 4$. Different markers are used for the three methods: squares for DAMC, circles for DALS and triangles for DASS. Let us concentrate the analysis on the performance of DALS and DASS. As can be seen, for each method, the required computational time increases as the expansion order increases, and the advantages with respect to standard MC increase as the expected collision probability decreases. In particular, DALS becomes progressively more and more appealing with decreasing collision probabilities.

Figure 5b shows the figure of merit Δ, normalized for each test with the value of the standard Monte Carlo method, as a function of the collision probability. The same criteria used in Fig. 5a are used here. The normalized unitary coefficient of variations is equal to 1 for the DAMC since the same number of samples of standard MC is used. The lower value is achieved with DALS, which is two orders of magnitude lower than DAMC. The efficiency of DASS increases for lower probabilities. As can be seen, the effect of the expansion order in this case is negligible. Overall, both DALS and DASS offer a valuable alternative tool with respect to standard MC for conjunction analysis.

4. NEO IMPACT PROBABILITY COMPUTATION

Present day approaches for robust detection and prediction of planetary encounters and potential impacts by NEO mainly refer to linearized models [9] or full nonlinear orbital sampling [29, 10, 19]. Among nonlinear methods, the Line of Variations (LOV) method [19] represents the reference technique for impact probability computation of NEO.

The preferred approach to detecting potential impacts depends on the uncertainty in the estimated orbit, the investigated time window and the dynamics between the observation epoch and the epoch of the expected impact [14]. Linear methods are preferred when linear approximations are reliable for both the orbit determination and uncertainty propagation. When these assumptions are not valid, one must resort to more computationally intensive techniques: among these, Monte Carlo methods are the most accurate but also the most computationally intensive, whereas the LOV method guarantees compute times 3-4 orders of magnitude lower than those required in MC simulations, though the LOV analysis may grow quite complex after it has been stretched and folded by multiple close planetary encounters, leaving open the possibility of missing some pathological cases [14].

Two different approaches are described in this paper: for the general problem of NEO impact probability computation, the presented LS and SS methods are directly applied, providing an efficient tool for impact monitoring. The special case of resonant return analysis is faced using a dedicated combination of ADS and IS. For each method, a single test case is presented. The propagations are carried out in Cartesian coordinates with respect to the J2000 reference frame centered in the Solar System Barycenter (SSB), with the inclusion of the

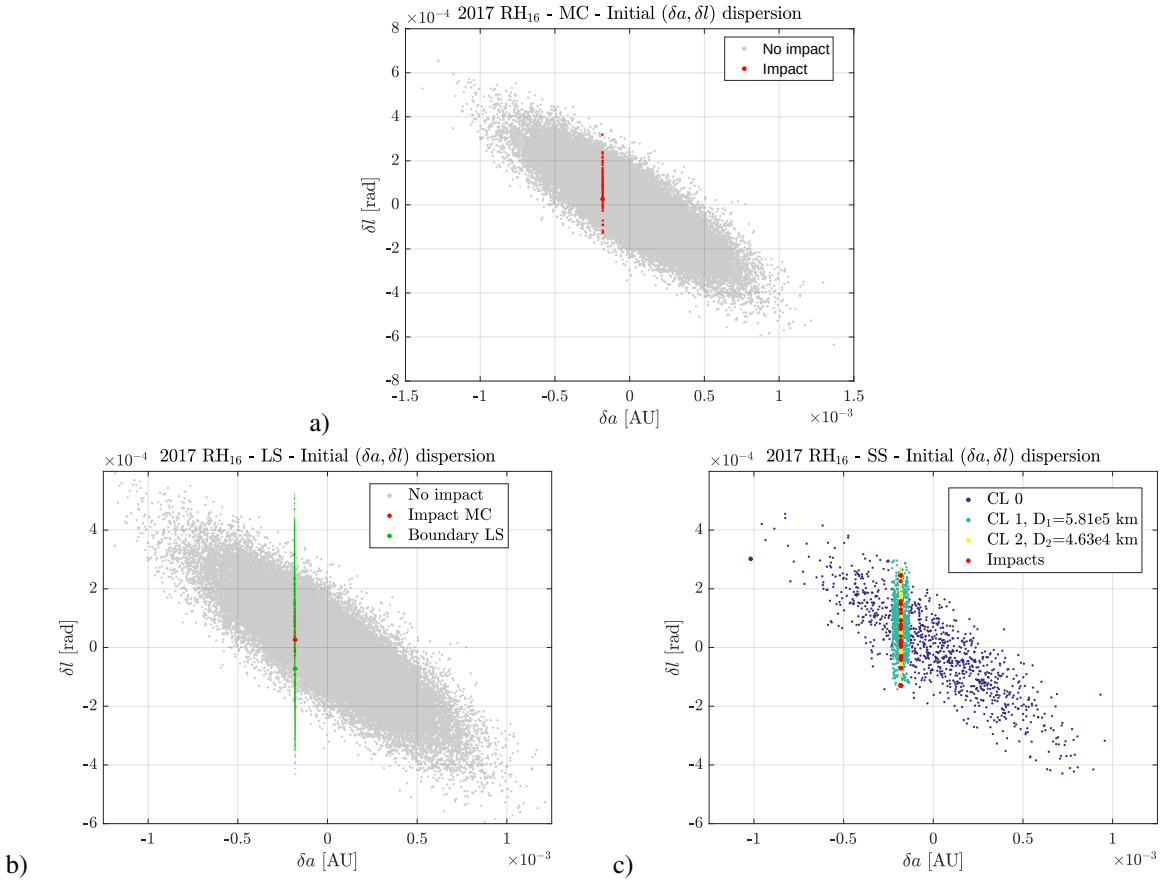

Figure 6: Visualization of the initial dispersion in the uncertainty space ($\delta\alpha$,δl) in the case of asteroid 2017 RH$_{16}$: a) initial conditions leading to impact obtained via standard MC, b) boundaries of the subdomain identified via LS, c) samples per conditional level obtained with SS.

Table 4: Application of Line Sampling and Subset Simulation to the case of asteroid 2017 RH$_{16}$ against standard Monte Carlo simulation.

	N_{IC}	N_T	$\hat{P}$ (-)	$\hat{\sigma}$ (-)	δ (-)	Δ (-)
MC	5e4	5e4	1.42e-3	1.68e-4	1.19e-1	26.61
LS (ref)	1e3	8245	1.56e-3	7.19e-5	4.60e-2	4.18
LS (σ^{MC})	164	1557	1.43e-3	1.70e-4	1.19e-1	4.70
LS (N_T^{MC})	6250	5.02e4	1.52e-3	2.85e-5	1.89e-2	4.19
SS (ref)	1e3	2.8e3	1.45e-3	2.24e-4	1.55e-1	8.19
SS (σ^{MC})	2e3	5.6e3	1.43e-3	1.57e-4	1.10e-1	8.27
SS (N_T^{MC})	1.8e4	5.04e4	1.42e-3	5.20e-5	3.66e-2	8.22

gravitational contributions of the Sun, all the major planets, and the Moon, including relativity effects [28]. All the physical constants (gravitational parameters, planetary radii, etc.) and ephemerides are obtained from the JPL Horizons database via the SPICE toolkit (`https://naif.jpl.nasa.gov/naif/`). All propagations are carried out in dimensionless units (with the scaling length and time respectively equal to 1 AU and 1 solar year) using the adaptive Dormand-Prince Runge-Kutta scheme of 8th order (RK78), with absolute and relative tolerances both equal to 1e-12.

The first test case presented is asteroid 2017 RH$_{16}$. Asteroid 2017 RH$_{16}$ will encounter the Earth at short distance in 2026, with a currently expected impact probability of 1:689. Table 3 shows the initial conditions in terms

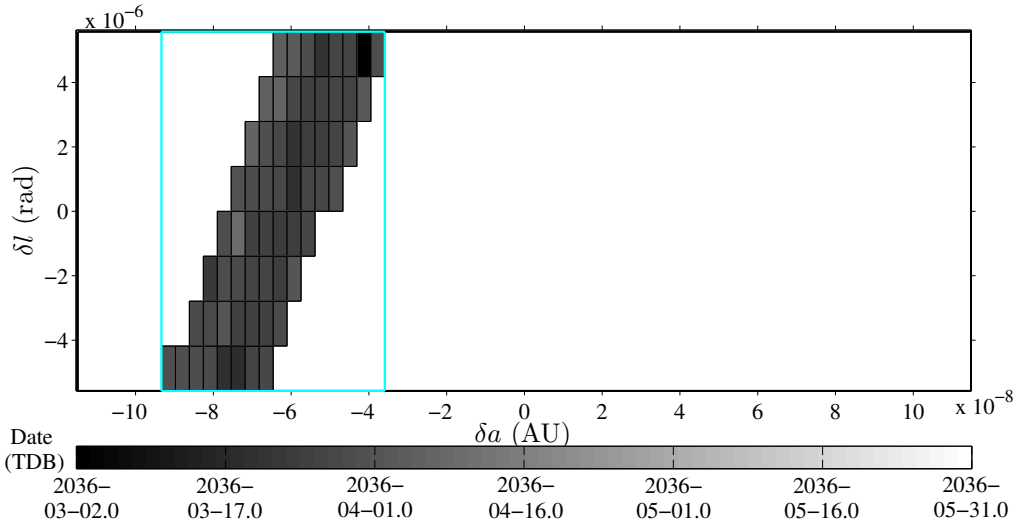

Figure 7: Projection of the generated subsets onto the $a - l$ plane of the initial conditions for test case of asteroid 99942 Apophis (ADP propagation, order 5, tolerance 1e-10, N_{max} 10, 5σ domain). In light blue, boundaries of the ISD [17].

Table 5: Apophis equinoctial parameters and related uncertainties on June 18, 2009 00:00:00 (TDB).

	Nominal value	σ	
a	0.922438242375914	2.29775e-8	AU
P_1	-0.093144699837425	3.26033e-8	-
P_2	0.166982492089134	7.05132e-8	-
Q_1	-0.012032857685451	5.39528e-8	-
Q_2	-0.026474053361345	1.83533e-8	-
l	88.3150906433494	6.39035e-5	◦

of nominal equinoctial parameters and related uncertainties, as obtained from the Near Earth Objects Dynamic Site[1]. This case was used to test the presented LS and SS method against the computation of the impact probability value in 2026.

Figure 6 shows the results of the numerical simulations obtained with standard MC, LS and SS, in terms of samples distribution on the semi-major axis- true longitude ($a - l$) plane of initial conditions. Figure 6(a) shows the results of standard MC, with impacting samples represented in red. Fig. 6(b) shows the results obtained with LS. Grey dots represent samples drawn from the initial distribution that do not lead to impact, whereas green dots are the initial conditions along the boundaries of the impact regions resulting from the LS application. Figure 6(c), finally, shows the evolution of the conditional samples obtained with SS. The method was applied by employing 1000 samples per conditional level, a fixed conditional probability equal to 0.2, and an auxiliary distribution centered in the current sample and with the same magnitude of the original one. Blue dots represent samples drawn at CL0 by standard MC, whereas different colors are used for each conditional level. The algorithm rapidly identifies the impact region, and after three conditional levels, the current threshold distance becomes lower than the Earth radius, and the algorithm stops.

An overview of the obtained performance is shown in Table 4. For both LS and SS, three results are presented: the results at convergence, the results obtained using a number of samples granting the same accuracy level of standard MC (σ^{MC}), and the results obtained performing the same number of propagations of standard MC ($N_{\mathrm{T}}^{\mathrm{MC}}$). The comparison between the methods is performed by analyzing the number of random initial conditions N_{IC}, the total number of orbital propagations N_{T}, the impact probability estimate $\hat{P}$, the standard deviation $\hat{\sigma}$ of $\hat{P}$, the coefficient of variation δ of the probability estimate and the figure of merit Δ. The overall number of propagations N_{T} was selected as a term of comparison for the computational burden of the methods. As can be seen, both methods allows us to improve the performance of standard MC both in terms of achievable accuracy and computational cost. That is, the same accuracy level of standard MC simulation can be obtained by performing a number of propagations that is one order of magnitude lower for both LS and SS, or a higher accuracy level can be obtained by propagating the same number of samples used for standard MC. This result is confirmed by the value of the figure of merit Δ.

The critical case of impact probability computation of NEO during resonant returns is here faced combining ADS and IS. The method, referred to as ADP-IS, was presented in [17] and can be divided in two phases. During the first phase, the epoch of the resonant return is estimated with DA, and the uncertainty set is propagated in time by means of ADS. A tailored pruning action is introduced during the propagation, and only subsets that are involved in the resonant return of interest are maintained throughout the propagation. At the end of the first

[1] `http://newton.dm.unipi.it/neodys/`

Table 6: Comparison between standard MC and ADP-IS method for the test case of the resonant return of asteroid 99942 Apophis [17].

	N_{T}	t_{CPU} / sample (s)	$\hat{P}$ (-)	$\hat{\sigma}$ (-)
MC	1.0e6	1.2	2.20e-5	4.71e-6
ADP-IS	341804	0.012	1.90e-5	6.81e-6

phase, several subdomains are generated only in specific regions of the initial uncertainty set. At this point, the sampling phase starts. An ISD is shaped on the basis of the generated subsets, and the IS phase is started. Samples are propagated to the epoch of the expected resonant return, and impacting samples are recorded. The procedure terminates when the variations in the estimated impact probability becomes lower than a selected threshold.

The method is here presented for the critical case of asteroid 99942 Apophis. Table 5 shows the nominal initial state and associated uncertainties σ for Apophis on June 18, 2009 expressed in terms of equinoctial parameters $\boldsymbol{p} = (a, P_1, P_2, Q_1, Q_2, l)$, considering a diagonal covariance matrix. Data were obtained from the Near Earth Objects Dynamic Site in September 2009. Asteroid Apophis will have a close encounter with Earth on April 13, 2029 with a nominal distance of 3.8e4 km. According to the selected initial conditions, though an impact in 2029 can be ruled out, the perturbations induced by the encounter open the door to resonant returns in 2036 and 2037. The aim is therefore to apply the presented method to provide an estimate for the impact probability at the epoch of the first resonant return, in 2036.

Figure 7 shows the results of the first propagation phase in terms of subdomains distribution on the $a - l$ plane. The method was applied considering an uncertainty set of 5σ size, an expansion order of the ADS propagation equal to 5, a maximum number of splits equal to 10, and a tolerance for the splitting procedure equal to 1e-10. Each subdomain is coloured according to the epoch at which its DA propagation was stopped. The subdomains distribution represents the starting point for the second phase, based on IS. First, an ISD including all the generated subsets is defined. Figure 7 shows in light blue the boundaries of this ISD as projected onto the $a - l$ plane. Then, the sampling phase starts: samples are generated on the basis of the ISD, each sample is associated, if possible, to a specific subset, and it is propagated to the epoch of minimum geocentric distance. All impacting samples are stored, and an estimate for the impact probability is obtained. Table 6 shows the results for the considered case, and a comparison with standard MC. Given the different propagation window for samples of the two method, an equivalent computational time per sample is introduced here as a term of comparison. As can be seen, the combination of ADS and IS guarantees a significant reduction in the required computational time, still providing good accuracy levels.

5. CONCLUSIONS

This paper presented an overview of past and present results obtained applying advanced uncertainty propagation and sampling techniques to the challenging tasks of in-orbit conjunction analysis and NEO impact probability computation. All presented approaches offer great benefits in terms of computational effort with respect to standard Monte Carlo approach, while granting a competitive accuracy level. Our research activity is currently devoted to the enhancement of the presented NEO impact probability computation tools, with the aim of widening the range of test cases and obtaining a more robust tool for impact monitoring.

ACKNOWLEDGMENTS

M. Losacco and M. Romano gratefully acknowledge professors E. Zio, N. Pedroni and F. Cadini from Politecnico di Milano for their introduction to the Line Sampling and Subset Simulation methods. The part of work about the Line Sampling was funded by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 679086 COMPASS).

REFERENCES

1. Akella, M. R. and Alfriend, K. T. (2000). Probability of collision between space objects. *Journal of Guidance, Control, and Dynamics*, 23(5):769–772.

2. Alfano, S. (2009). Satellite conjunction Monte Carlo analysis. *Advances in Astronautical Sciences*, 134:2007–2024.

3. Au, S.-K. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by Subset Simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.

4. Berend, N. (1999). Estimation of the probability of collision between two catalogued orbiting objects. *Advances in Space Research*, 23(1):243–247.

5. Berz, M. (1986). The new method of TPSD algebra for the description of beam dynamics to high orders. techreport AT-6:ATN-86-16, Los Alamos National Laboratory.

6. Berz, M. (1987). The method of power series tracking for the mathematical description of beam dynamics. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, A258(3):431–436.

7. Berz, M. (1999). *Modern Map Methods in Particle Beam Physics*. Advances in Imaging and Electron Physics. Academic Press, 1 edition.

8. Cadini, F., Avram, D., Pedroni, N., and Zio, E. (2012). Subset Simulation of a reliability model for radioactive waste repository performance assessment. *Reliab. Eng. Syst. Safe.*, 100:75–83.

9. Chodas, P. W. (1993). Estimating the impact probability of a minor planet with the Earth. *Bulletin of the American Astronomical Society*, 25:1236.

10. Chodas, P. W. and Yeomans, D. K. (1999). Could asteroid 1997 XF11 collide with Earth after 2028? In *AAS/Division of Dynamical Astronomy Meeting*, volume 31 of *Bulletin of the American Astronomical Society*, page 1227.

11. Dagum, P., Karp, R., Luby, M., and Ross, S. (2000). An optimal algorithm for Monte Carlo estimation. *SIAM J. Comput.*, 29(5):1484–1496.

12. Di Lizia, P. (2008). *Robust Space Trajectory and Space System Design using Differential Algebra*. phdthesis, Politecnico di Milano.

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

14. Farnocchia, D., Chesley, S. R., Milani, A., Gronchi, G. F., and Chodas, P. W. (2015). Orbits, long-term predicitions, and impact monitoring. In Michel, P., DeMeo, F. E., and Bottke, W. F., editors, *Asteroids IV*, pages 815–834. University of Arizona, 1 edition.

15. Klinkrad, H., Alarcon, J. R., and Sanchez, N. (2005). Collision avoidance for operational ESA satellites. In *Proc. of the 4th European Conference on Space Debris*.

16. Kolchin, E. R. (1973). *Differential Algebra and Algebraic Groups*. Academic Press.

17. Losacco, M., Di Lizia, P., Armellin, R., and Wittig, A. (2018). A differential algebra-based importance sampling method for impact probability computation on Earth resonant returns of Near Earth Objects. *MNRAS*, 479(4):5474–5490.

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

19. Milani, A., Chesley, S. R., Chodas, P. W., and Valsecchi, G. B. (2002). Asteroid close approaches: Analysis and potential impact detection. In Bottke Jr., W. F., Cellino, A., Paolicchi, P., and Binzel, R. P., editors, *Asteroids III*, pages 55–69. The University of Arizona Press, 1 edition.

20. Morselli, A., Armellin, R., Di Lizia, P., and Bernelli Zazzera, F. (2014). A high order method for orbital conjunctions analysis: Monte Carlo collision probability computation. *Advances in Space Research*, 55(1):311–33.

21. Patera, R. P. (2001). General method for calculating satellite collision probability. *Journal of Guidance, Control, and Dynamics*, 24(4):716–722.

22. Rasotto, M., Morselli, A., Wittig, A., Massari, M., Di Lizia, P., Armellin, R., Valles, C. Y., and Ortega, G. (2016). Differential algebra space toolbox for nonlinear uncertainty propagation in space dynamics. In *6th International Conference on Astrodynamics Tools and Techniques (ICATT)*.

23. Risch, R. H. (1969). The problem of integration in finite terms. *Transactions of the American Mathematical Society*, 139:167–189.

24. Risch, R. H. (1970). The solution of the problem of integration in finite terms. *Bulletin of the American Mathematical Society*, 76(3):605–608.

25. Ritt, J. F. (1932). *Differential Equations From the Algebraic Standpoint*. American Mathematical Society.

26. Ritt, J. F. (1948). *Integration in finite terms: Liouville's theory of elementary methods*. Columbia University Press.

27. Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.*, 23(3):470–472.

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

29. Yeomans, D. K. and Chodas, P. W. (1994). Predicting close approaches of asteroids and comets to Earth. In Gehrels, T., Matthews, M. S., and Schumann, A. M., editors, *Hazards Due to Comets and Asteroids*, page 241.

30. Zio, E. (2013). *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. Springer Series in Reliability Engineering. Springer, 1 edition.

31. Zio, E. and Pedroni, N. (2009). Subset Simulation and Line Sampling for advanced Monte Carlo reliability analysis. In *Proceedings of the European Safety and RELiability (ESREL) 2009 Conference*, pages 687–694.

32. Zuev, K. M., Beck, J. L., Au, S.-K., and Katafygiotis, L. S. (2012). Bayesian post-processor and other enhancements of Subsets Simulation for estimating failure probabilities in high dimensions. *Computers and Structures*, 92-92:283–296.