69th International Astronautical Congress (IAC), Bremen, Germany, 1-5 October 2018.


IAC-18-A3.5.10

# Efficient planetary protection analysis for interplanetary missions

**Matteo Romano[a]*, Camilla Colombo[a], Jose Manuel Sánchez Pérez [b]**

[a] *Dept. of Aerospace Science and Technology, Politecnico di Milano, Milano, Italy*, matteo1.romano@polimi.it, camilla.colombo@polimi.it
[b] *European Space Operations Centre (ESOC), ESA, Darmstadt, Germany*, Jose.Manuel.Sanchez.Perez@esa.int
* Corresponding Author

**Abstract**

The process of verifying that interplanetary missions respect the planetary protection requirements must account for uncertainties in the design parameters of the mission and perform long numerical simulations to estimate the impact probability of the mission-related objects with celestial bodies that could develop extra-terrestrial life. This kind of analysis is usually done via Monte Carlo simulation, with high computational cost since the requirements also include high confidence levels of the probability estimate. In order to reduce the computational load of the simulation the line sampling method, already analysed in previous works, is used here in order to further characterise his numerical performance, by providing an approximate formula highlighting the dependency of the method from the level of probability and the shape of the impact regions in the uncertainty space, and by analysing how its accuracy changes for different shapes of the initial distribution. The observations made here will allow to identify in advance in which cases the method will perform better than the standard Monte Carlo according to the expected impact probability and the shape of the initial distribution.
**Keywords:** Planetary protection, Orbital propagation, Numerical integration, Monte Carlo sampling

## 1. Introduction

### *1.1 Planetary protection*

Planetary protection (PP) requirements set stringent constraints on the design of trajectories in the Solar System, since they aim to avoid the contamination of planets and moons where life could develop by limiting the probability of impact between spacecraft or launcher stages and these celestial bodies [1]. The process of verifying that spaceflight missions fulfil the requirements is typically performed via Monte Carlo (MC) methods and must account for uncertainties in the design parameters of the spacecraft, random failures, errors in the determination of its state, chaotic n-body dynamics and not modelled effect in the dynamics, which introduces numerical errors in the propagation of the trajectory. This is expensive in terms of numerical resources, since the requirements also include long time intervals (up to 100 years in most cases) and high confidence levels of the probability estimates, which increases the number of propagations to perform.

### *1.2 Proposed approach*

On the side of the statistical analysis, to reduce the computational load, the impact probability is estimated through the use of the Line Sampling (LS) method [2][4], as an alternative to the conventional Monte Carlo method, which propagates a large number of initial conditions directly sampled from an uncertainty distribution; instead, LS samples the initial distribution in a more efficient way, aimed to provide a probability estimation with a higher confidence level, or employing a lower number of samples to reach the desired accuracy level. In this work, the LS is further characterised by analysing how the shape of the initial uncertainty (expressed through a covariance matrix) affects numerical performance of the method; in addition, the number of LS runs necessary to reach the confidence level imposed by the PP requirements is estimated in advance by using an approximated analytical formula which was developed starting from the information already available in the literature.

On the other side, the orbital propagations are carried out by taking into account the characterisation of the close approaches with planetary bodies, by obtaining information about the dynamics using the eigenvalues of the Jacobian matrix of the equations of motion.

The techniques presented here have been implemented into SNAPPshot (tool suite for the verification of the compliance to planetary protection requirements initially developed at the University of Southampton in the framework of a study for ESA [5][6]). The tool follows a Monte Carlo approach, where the initial uncertainty (over the state or other design parameters of the spacecraft or launcher) is sampled into many initial conditions, that are then propagated to estimate the probability of impact (or orbital resonance) with other celestial bodies.

### *1.3 Manuscript content and outline*

The paper is organised as follows: Section 2 is dedicated to explaining the advances on the application of the LS method for MC analysis, and to the approach proposed to improve the accuracy and efficiency of the orbital propagation in terms of analysis of the close approaches; Section 3 considers the planetary protection analysis for the launcher upper stage of the Solo spacecraft to show the application of all the techniques, by defining test cases aimed to showing how different conditions affect the accuracy of the impact probability estimate to obtain a criterium to identify in advance when the LS will be more efficient than the standard MC; finally, Section 4 will summarise the main results and conclusions, and anticipate future developments.

## 2. New developments

### *2.1 Line Sampling*

The LS method was introduced in previous works [2][3], where a general explanation of the theory behind it was presented, together with the results of its application to different test cases, to show how the choice of this sampling method can improve the efficiency of the MC simulations for planetary protection analysis. In short, the main feature of this method is the analytical estimation of the probability, obtained by reducing the multi-dimensional integration problem across the uncertainty domain to many one-dimensional problems along lines following a reference direction that are used to sample the initial distribution; this direction is determined so that it points toward an impact region of the domain, and, if this is properly chosen, the method can considerably reduce the number of required system simulations with respect to a standard MC.

In this work, the method is further developed by introducing a way to estimate in advance the number of runs that are required to reach a desired confidence level for a given expected impact probability. This is done to mirror the functionality that is already built in the SNAPPshot tool for the standard MC analysis.

In the case of the LS, the literature already gives a qualitative estimation of its efficiency compared with the standard MC in terms of convergence rate [4]. A summary of it is reported in Section 2.1.1 to introduce the notation that will be used in Section 2.1.2 .

#### *2.1.1 Theoretical formulation of the LS method*

Following the explanation and the notation presented in [4], the probability of the event F (which can be seen as the failure of a system or, in this case, an impact with a celestial body) can be expressed as the multidimensional integral in the form

$$P(F) = P(\mathbf{x} \in F) = \int I_F(\mathbf{x}) q_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \qquad (2.1)$$

where $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d$ is the vector of the uncertain variables of the system, $q_{\mathbf{x}}(\mathbf{x})$ is the multidimensional probability density function (pdf), $F$ is the subdomain of the variables $\mathbf{x}$ leading to the event of interest, defined by a performance function $g_{\mathbf{x}}(\mathbf{x})$ (which is lower than or equal to zero if $\mathbf{x} \in F$ and greater than zero otherwise) and $I_F(\mathbf{x})$ is an indicator function such that $I_F(\mathbf{x}) = 1$ if $\mathbf{x} \in F$ and $I_F(\mathbf{x}) = 0$ otherwise.

A coordinate transformation from the physical space to the standard normal space $T_{\mathbf{x0}} : \mathbf{x} \to \boldsymbol{\theta}$ brings as advantages the normalisation of the physical variables through the covariance matrix, and the possibility to express the multidimensional pdf as a product of $d$ unit Gaussian standard distributions $\phi_j(\theta_j)$:

$$\varphi(\boldsymbol{\theta}) = \prod_{j=1}^{d} \phi_j(\theta_j) \qquad (2.2)$$

With reference to Fig. 1, in the d-dimensional standard normal space, the domain $F$ is the subspace for which the samples $\boldsymbol{\theta} = (\theta_1, ..., \theta_d)^T$ satisfy a given property (e.g. an impact with a planet or a system failure). With the assumption that $\theta_1$ points in the direction of the sampling vector $\boldsymbol{\alpha}$ (this can always be assured by a suitable rotation of the coordinate axes), the subdomain $F$ can be also expressed as

$$F = \left\{ \boldsymbol{\theta} \in \mathbb{R}^d : \theta_1 \in F_1\left(\theta_1, ..., \theta_j, ..., \theta_d\right) \right\} \qquad (2.3)$$

with $F_1 \in \mathbb{R}^{d-1}$, in this way the region $F$ corresponds to the values of $\boldsymbol{\theta}$ such that the performance function $g_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ satisfies the relation $g_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = g_{\boldsymbol{\theta},-1}(\boldsymbol{\theta}_{-1}) - \theta_1 \geq 0$, where $\boldsymbol{\theta}_{-1} = (\theta_2, ..., \theta_d)^T \in \mathbb{R}^{d-1}$.

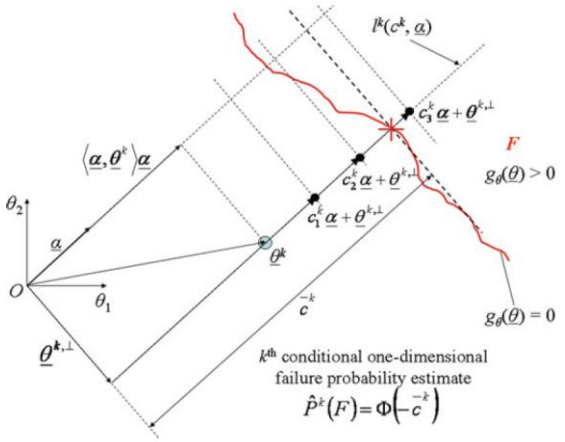



**Fig. 1 - Scheme representing the sampling procedure along a line (characterised by the $c^k$ parameter) to identify the border (red line) of the region $F$ of interest (image from [4]).**

Considering this change of variables and the definition in Eq. (2.3), the integral in Eq. (2.1) can be rewritten as

$$P(F) = \iint_d I_F(\mathbf{\theta}) \prod_{j=1}^{d} \phi_j(\theta_j) d\mathbf{\theta} = E\left[I_F(\mathbf{\theta})\right] \quad (2.4)$$

(where $E[\mathbf{X}]$ is defined as the expected value of the generic random variable $\mathbf{X}$ ) and manipulated as follows:

$$\begin{aligned} P(F) &= \int_d I_F(\mathbf{\theta}) \prod_{j=1}^{d} \phi_j(\theta_j) d\mathbf{\theta} \\ &= \int_{d-1} \ldots \int \left( \int I_F(\mathbf{\theta}_{-1}) \phi_1(\theta_1) d\theta_1 \right) \prod_{j=2}^{d} \phi_j(\theta_j) d\mathbf{\theta}_{-1} \\ &= \int_{d-1} \ldots \int \Phi\left(F_1(\mathbf{\theta}_{-1})\right) \prod_{j=2}^{d} \phi_j(\theta_j) d\mathbf{\theta}_{-1} \\ &= E_{\mathbf{\theta}_{-1}}\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] \end{aligned} \quad (2.5)$$

where $\Phi(A) = \int I_A(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x}$ is the definition of the Gaussian measure of A, where A is the subset of the random variables $\mathbf{x}$ which lead to a given result (e.g. an impact). In case of the standard MC (which could be considered a Point Sampling method, in relation with the LS), $\Phi\left(F_1(\mathbf{\theta}_{-1})\right)$ is a discrete random variable equal to $I_F(\mathbf{\theta})$ (meaning that $\Phi^2\left(F_1(\mathbf{\theta}_{-1})\right) = \Phi\left(F_1(\mathbf{\theta}_{-1})\right)$ is always true $\forall \mathbf{\theta}_{-1} \in \mathbb{R}^{d-1}$ ), while for the LS method $\Phi\left(F_1(\mathbf{\theta}_{-1})\right)$ is a continuous random variable where $F_1(\mathbf{\theta}_{-1}^k) = -\bar{c}^k$ (see Fig. 1, where the sampling procedure is represented highlighting the boundary of the region corresponding to the event F), meaning that $0 \le \Phi\left(F_1(\mathbf{\theta}_{-1})\right) \le 1$ and $0 \le \Phi^2\left(F_1(\mathbf{\theta}_{-1})\right) \le \Phi\left(F_1(\mathbf{\theta}_{-1})\right)$ are always true $\forall \mathbf{\theta}_{-1} \in \mathbb{R}^{d-1}$.

The consequence of these properties is visible when considering the definition of variance of an estimator for the two methods. An estimator $\hat{P}(F)$ of the probability $P(F)$ as expressed in Eq. (2.5) can be computed as

$$\hat{P}(F) = \frac{1}{N_T} \sum_{k=1}^{N_T} \Phi\left(F_1(\mathbf{\theta}_{-1}^k)\right) \quad (2.6)$$

where $\mathbf{\theta}^k, k = 1, \ldots, N_T$ are independent and identically distributed samples in the standard normal coordinate space. Given the generic definition of variance for $P(F)$ following Eq. (2.5) as

$$\begin{aligned} &\sigma^2\left(P(F)\right) \\ &= \int \left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right) - P(F)\right]^2 \varphi(\mathbf{\theta}_{-1}) d\mathbf{\theta}_{-1} \\ &= E_{\mathbf{\theta}_{-1}}\left[\Phi^2\left(F_1(\mathbf{\theta}_{-1})\right)\right] - E_{\mathbf{\theta}_{-1}}^2\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] \\ &= E_{\mathbf{\theta}_{-1}}\left[\Phi^2\left(F_1(\mathbf{\theta}_{-1})\right)\right] - P(F)^2 \end{aligned} \quad (2.7)$$

the variance of the estimator $\hat{P}(F)$ is defined as

$$\begin{aligned} &\sigma^2\left(\hat{P}(F)\right) \\ &= \sigma^2\left(P(F)\right) / N_T = \sigma^2\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] / N_T \end{aligned} \quad (2.8)$$

meaning that the variance of the estimator directly depends on the variance of the random variable $\Phi\left(F_1(\mathbf{\theta}_{-1})\right)$. Consequently

$$\begin{aligned} &\sigma^2\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] \\ &= E_{\mathbf{\theta}_{-1}}\left[\Phi^2\left(F_1(\mathbf{\theta}_{-1})\right)\right] - E_{\mathbf{\theta}_{-1}}^2\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] \\ &\le E_{\mathbf{\theta}_{-1}}\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] - E_{\mathbf{\theta}_{-1}}^2\left[\Phi\left(F_1(\mathbf{\theta}_{-1})\right)\right] \\ &= P(F)\left(1 - P(F)\right) = \sigma^2\left[I_F(\mathbf{\theta})\right] \end{aligned} \quad (2.9)$$

A coefficient of variation (c.o.v.) $\delta = \sqrt{\sigma^2(\hat{P}(F))} / P(F)$ can be defined as a measure of the efficiency of the sampling method, with lower values of $\delta$ meaning a higher efficiency of the method in converging to the exact value of the probability. Eq. (2.9) demonstrates that the c.o.v. of estimator in Eq. (2.6) as given by the LS method is always smaller than the one given by the standard MC, implying that the convergence rate of the LS is always faster than, or as fast as, that of the standard MC.

### *2.1.2 New developments*

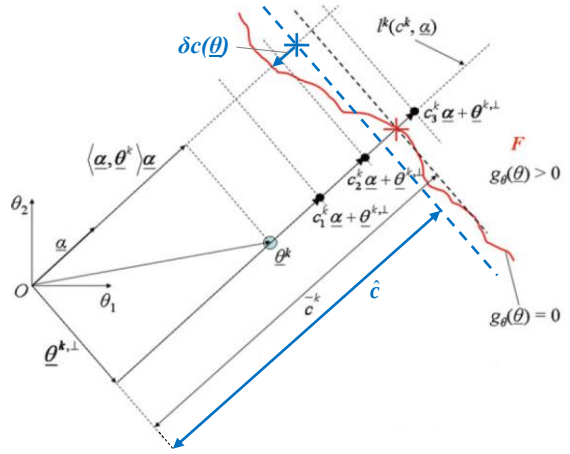



**Fig. 2 - Scheme representing the approximations used to express the variance of the LS method as a function of the probability estimate.**

While in the case of the standard MC Eq. (2.9) is easy to treat, since $E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\right]=E_{\boldsymbol{\theta}_{-1}}\left[\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)\right]$, in the LS case $E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\right]$ is a continuous variable, defined through the integral

$$E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\right] = \int_{d-1}\ldots\int\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\prod_{j=2}^{d}\phi_j(\theta_j)d\boldsymbol{\theta}_{-1} \tag{2.10}$$

which cannot be easily manipulated analytically due to the presence of $\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)$. For this reason, it is chosen to express this term with an approximation.

The definition of $\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)$ given in Eq. (2.5) can be further expanded as

$$\begin{aligned}\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right) &= \int_{-\infty}^{\infty} I_F(\boldsymbol{\theta}_{-1})\phi_1(\theta_1)d\theta_1 \\ &= \int_{\bar{c}(\boldsymbol{\theta}_{-1})}^{\infty}\phi_1(\theta_1)d\theta_1 = 1-\Phi\left(\bar{c}(\boldsymbol{\theta}_{-1})\right)=\Phi\left(-\bar{c}(\boldsymbol{\theta}_{-1})\right)\end{aligned} \tag{2.11}$$

with $\bar{c}(\boldsymbol{\theta}_{-1})$ defined as the border of the region $F$ displayed in Fig. 2 as a red line. $\bar{c}(\boldsymbol{\theta}_{-1})$ is then expanded as $\bar{c}(\boldsymbol{\theta}_{-1})=\hat{c}+\delta c(\boldsymbol{\theta}_{-1})$, with the first term defined as an "average" value of $\bar{c}(\boldsymbol{\theta}_{-1})$ (represented as a dashed blue line in Fig. 2) such that $P(F)=E_{\boldsymbol{\theta}_{-1}}\left[\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)\right]=\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)=\Phi\left(-\hat{c}\right)$, and the second term as a variation with respect to this average value.

The hypothesis is made that $\delta c(\boldsymbol{\theta}_{-1})$ represents a small variation with respect to the average value $\hat{c}$, as in the case of a quasi rectilinear border of the region $F$ orthogonal to the sampling direction $\boldsymbol{\alpha}$. Under this hypothesis, the integral in Eq. (2.11) can be rewritten as

$$\begin{aligned}&\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right) \\ &= \int_{\bar{c}(\boldsymbol{\theta}_{-1})}^{\infty}\phi_1(\theta_1)d\theta_1 = \int_{\hat{c}+\delta c(\boldsymbol{\theta}_{-1})}^{\infty}\phi_1(\theta_1)d\theta_1 \\ &= \int_{\hat{c}}^{\infty}\phi_1(\theta_1)d\theta_1 - \int_{\hat{c}}^{\hat{c}+\delta c(\boldsymbol{\theta}_{-1})}\phi_1(\theta_1)d\theta_1 \\ &\cong \Phi\left(-\hat{c}\right)-\phi(\hat{c})\delta c(\boldsymbol{\theta}_{-1})\end{aligned} \tag{2.12}$$

resulting in

$$\begin{aligned}&E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\right] \\ &\cong E_{\boldsymbol{\theta}_{-1}}\left[\left(\Phi\left(-\hat{c}\right)-\phi(\hat{c})\delta c(\boldsymbol{\theta}_{-1})\right)^2\right] \\ &= E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(-\hat{c}\right)-2\Phi\left(-\hat{c}\right)\phi(\hat{c})\delta c(\boldsymbol{\theta}_{-1})+\phi^2(\hat{c})\delta c^2(\boldsymbol{\theta}_{-1})\right] \\ &= E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(-\hat{c}\right)\right]-E_{\boldsymbol{\theta}_{-1}}\left[2\Phi\left(-\hat{c}\right)\phi(\hat{c})\delta c(\boldsymbol{\theta}_{-1})\right]+E_{\boldsymbol{\theta}_{-1}}\left[\phi^2(\hat{c})\delta c^2(\boldsymbol{\theta}_{-1})\right] \\ &= P(F)^2-P(F)\cdot 2\phi(\hat{c})E_{\boldsymbol{\theta}_{-1}}\left[\delta c(\boldsymbol{\theta}_{-1})\right]+\phi^2(\hat{c})\cdot E_{\boldsymbol{\theta}_{-1}}\left[\delta c^2(\boldsymbol{\theta}_{-1})\right]\end{aligned} \tag{2.13}$$

Taking expression (2.13) into account, and defining in a compact way $\Delta c(\boldsymbol{\theta}_{-1})=E_{\boldsymbol{\theta}_{-1}}\left[\delta c(\boldsymbol{\theta}_{-1})\right]$, the variance given by the LS in Eqs. (2.7) and (2.9) becomes

$$\begin{aligned}&\sigma^2\left(P(F)\right)=\sigma^2\left(\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)\right) \\ &= E_{\boldsymbol{\theta}_{-1}}\left[\Phi^2\left(F_1(\boldsymbol{\theta}_{-1})\right)\right]-P(F)^2 \\ &\cong -P(F)\cdot 2\phi(\hat{c})\Delta c(\boldsymbol{\theta}_{-1})+\phi^2(\hat{c})\cdot E_{\boldsymbol{\theta}_{-1}}\left[\delta c^2(\boldsymbol{\theta}_{-1})\right] \\ &\cong -P(F)\cdot 2\phi(\hat{c})\Delta c(\boldsymbol{\theta}_{-1}) \\ &\le P(F)\left(1-P(F)\right)=\sigma^2\left[I_F(\boldsymbol{\theta})\right]\end{aligned} \tag{2.14}$$

Highlighting the new terms in Eq. (2.15)

$$\begin{aligned}&\sigma^2\left(\Phi\left(F_1(\boldsymbol{\theta}_{-1})\right)\right)\cong -P(F)\cdot 2\phi(\hat{c})\Delta c(\boldsymbol{\theta}_{-1}) \\ &\le P(F)\left(1-P(F)\right)=\sigma^2\left[I_F(\boldsymbol{\theta})\right]\end{aligned} \tag{2.15}$$

this means that a new estimation for the worst covariance given by the LS method (nominally, from Eq. (2.9), equal to the one given by the standard MC) was obtained, which takes into account the probability level through the term $\phi(\hat{c})$, and the shape of the region $F$ and the direction of sampling through the term $\Delta c(\boldsymbol{\theta}_{-1})$. When the approximation of small $\delta c(\boldsymbol{\theta}_{-1})$ is valid (that is, when the region $F$ has a regular shape and is distributed across the initial uncertainty, and the sampling direction is chosen properly so that it points toward it) and the probability level is low, the term $\phi^2(\hat{c})\cdot E_{\boldsymbol{\theta}_{-1}}\left[\delta c^2(\boldsymbol{\theta}_{-1})\right]$ is also small, and we can say that the variance given by the LS is below a value $f\left(P(F),\Delta c(\boldsymbol{\theta}_{-1})\right)$ such that

$\sigma^2\left(P(F)\right)_{LS} \le f\left(P(F),\Delta c(\boldsymbol{\theta}_{-1})\right) \le \sigma^2\left(P(F)\right)_{MC}$, thus increasing the convergence rate of LS with respect to standard MC. On the contrary, when the approximation does not hold (that is in cases with high probability levels, non-optimal sampling direction, or badly shaped impact regions), $f\left(P(F),\Delta c(\boldsymbol{\theta}_{-1})\right)$ grows toward the covariance level of the MC.

### *2.2 Fly-by detection*

#### *2.2.1 Explanation*

As already pointed out in previous works [2], close approaches with planetary bodies critically influence the accuracy of the numerical propagation, due to an increase of the nonlinearity of the dynamics with respect to the interplanetary phase of the propagation. This effect is stronger for very close fly-bys and was observed to affect in similar ways any integration method that was examined.

For this reason, a technique that uses information from the dynamics to identify numerically a fly-by condition has been developed as a criterion that can be evaluated automatically during the integration. The Jacobian of the equations of motion (expressed via its eigenvalues and their derivatives) is used to detect when the propagated object is approaching a planet, looking at both the relative position (already accounted for when considering distance-based criteria such as Sphere of Influence SOI radius) and the relative velocity between the planet and the object. This method was already introduced in a previous work [2], while here it will be explained in detail and applied to more test cases.

The equations of motion of the barycentric restricted n-body problem can be written as

$$\mathbf{a} = \ddot{\mathbf{r}} = -\sum_{j=0}^{n} \mu_j \frac{\mathbf{r}-\mathbf{r}_j(t)}{\left\|\mathbf{r}-\mathbf{r}_j(t)\right\|^3} \tag{2.16}$$

with the Jacobian matrix defined as

$$\mathbf{J} = \frac{d\mathbf{f}}{d\mathbf{x}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{G} & \mathbf{0} \end{bmatrix} \tag{2.17}$$

where $\mathbf{x}$ is the state vector, containing position and velocity vectors $\mathbf{r}$ and $\mathbf{v}$), $\mathbf{I}$ is the identity matrix, and $\mathbf{G}$ results from the derivation of the gravitational accelerations defined in Eq. (2.16):

$$\mathbf{G} = \frac{\partial \mathbf{a}}{\partial \mathbf{r}} = \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} = \sum_{j=0}^{n} \mathbf{G}_j \tag{2.18}$$

where the index j=0 corresponds to the main attractor of the system. The set of eigenvalues of the complete Jacobian are given by

$$\lambda = \det(\mathbf{GI} - \lambda\mathbf{I}) = \det(\mathbf{G} - \lambda\mathbf{I}) \tag{2.19}$$

with $\Lambda$ being the eigenvalue with the maximum absolute value. In this case, as an approximation, only the contributions to the Jacobian given by each planet alone are considered:

$$\lambda_j = \det(\mathbf{G}_j - \lambda\mathbf{I}), \quad j = 1,...,N \tag{2.20}$$

with $\lambda_j$ being the set of eigenvalues given by the contribution of j-th planet, and $\Lambda_j$ the maximum eigenvalue of such set (it is clear that $\lambda \neq \sum_j \lambda_j$ ). Debatin et al. [7] propose a simplified expression for $\Lambda$ that can be estimated as

$$\Lambda_j = \frac{2\mu_j}{\left\|\mathbf{r}-\mathbf{r}_j\right\|^3} \tag{2.21}$$

With this criterion, not only the single eigenvalues are considered, but also their derivative in time:

$$\dot{\Lambda}_j = \left| 2\mu_j \frac{3\left(\mathbf{r}-\mathbf{r}_j\right)^T\left(\mathbf{v}-\mathbf{v}_j\right)}{\left\|\mathbf{r}-\mathbf{r}_j\right\|^5} \right| \tag{2.22}$$

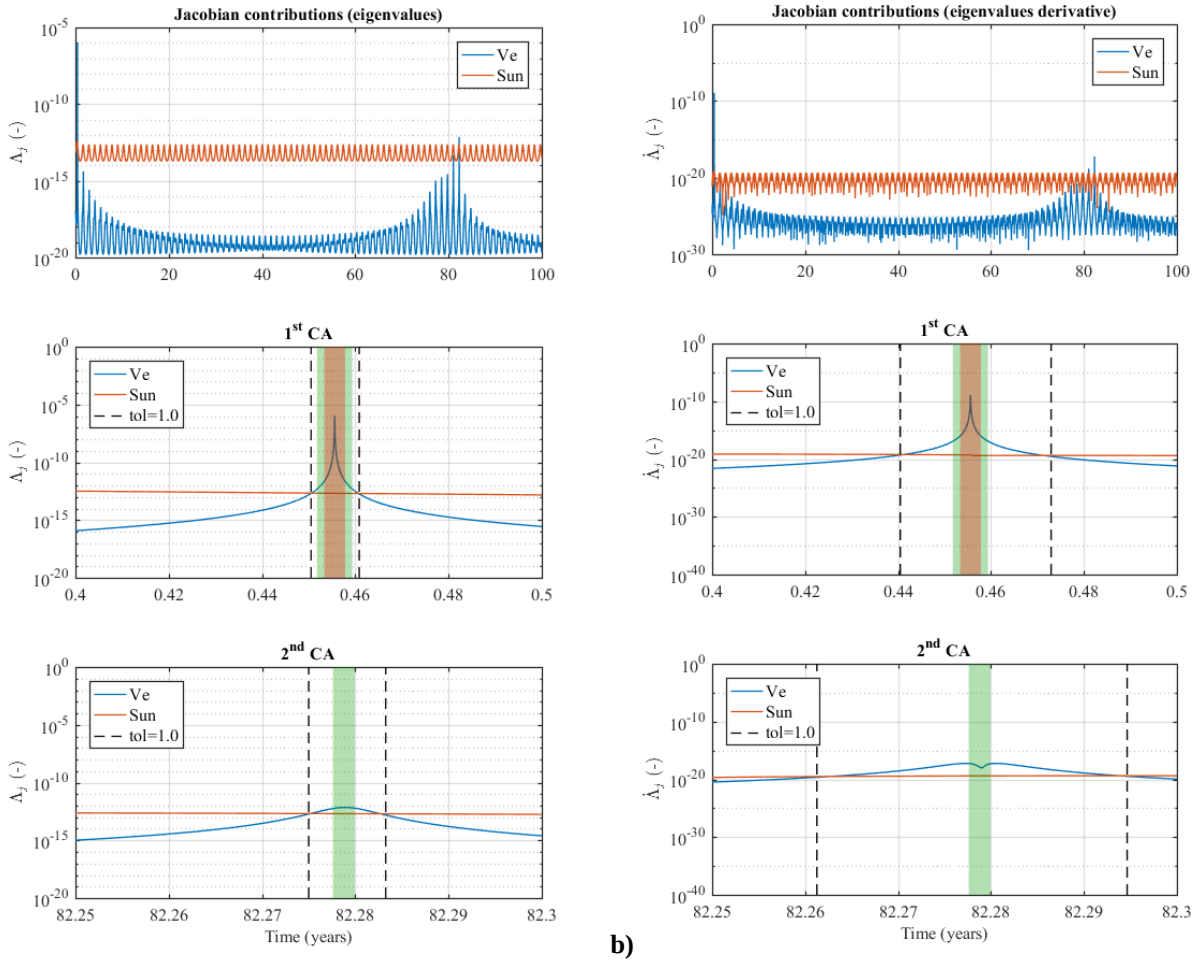
The value of the eigenvalue contributions given by the single planets are compared with the one given by the main attractor (the Sun in the case of an interplanetary trajectory), and the same is done for their derivatives. A fly-by event is identified when one or both ratios in Eq. (2.23) reaches a given threshold:

$$\begin{aligned} \Lambda_j / \Lambda_0 &\ge \varepsilon_1 \\ \dot{\Lambda}_j / \dot{\Lambda}_0 &\ge \varepsilon_2 \end{aligned} \tag{2.23}$$

For the application of this method, one or both expressions in (2.23) (ratio of the values and ratio of the derivatives of the eigenvalues) can be used, separately or together, as shown in the next section.

#### *2.2.2 Example*

The application of this method can be seen in Fig. 3, which shows the case of multiple close approaches between the launcher upper stage od Solo and Venus. In both cases the variations of the eigenvalues and of their derivatives are compared with the crossing of the SOI and Hill sphere of Venus, to show the differences between the two criteria.




**a)** **b)**

**Fig. 3 - Variation in time of the eigenvalues corresponding to the single contribution of the Sun and Venus during the propagation of the nominal trajectory of Solo's launcher upper stage: values (a) and derivatives (b) of the eigenvalues, compared with the crossing of SOI (red area) and Hill sphere (green area) of Venus. The vertical dashed lines refer to the epochs where the ratio between the eigenvalues of Venus and the Sun is equal to 1.0.**

Fig. 3 shows the variation in time of the eigenvalues relative to Venus and the Sun and their associated derivative (respectively (a) and (b)) during the propagation of the nominal trajectory of the launcher upper stage of Solo. In both cases, the comparison is done on the 100 years propagation (top), with a focus on the 1st close approach with Venus during the first year of the mission (centre), and on a second close approach 80 years later (bottom). In both graphs, different information is reported: the red and green areas represent, respectively, the crossing of the SOI and the Hill sphere of Venus; the vertical dashed lines indicate the epochs where the ratios defined in Eq. (2.23) are both equal to 1.0 (only one value is used for both ratios for simplicity).

The plots show that both close approaches can be successfully identified using the definitions defined previously. In particular, a threshold value of 1.0 for the ratios correctly identifies the not only the 1st CA (where an actual crossing of the SOI occurs), but also the 2nd CA, which happens at a larger distance from Venus, with no SOI crossing. This is possible due to the information about the relative velocity between the propagated body and the planet contained in the derivative in Eq. (2.22). Notice also that a tolerance equal to 1.0 allows to determine initial and final epochs for the close approach in a broader sense than the ones defined by the SOI and Hill sphere crossings, particularly for the derivative ratio, meaning that a lower value can be also used in the case of interplanetary trajectories.

## 3. Planetary protection analysis

### *3.1 Test case definition*

The techniques presented in the previous sections were implemented into the SNAPPshot tool [5][6] and used to perform a Planetary Protection (PP) analysis for the Atlas V upper stage of ESA's SolO (Solar Orbiter) mission (according to the October 2018 launch option [8]. The analysis will focus on the fly-by with Venus the is expected during the first year of the mission with the given launch option. Although this planet has no explicit planetary protection requirements, the Venus fly-by represents an interesting case to test the LS technique. Initial data are taken from [5], with initial conditions and covariance matrix expressed in Cartesian coordinates.

A series of test cases is defined by modifying the covariance matrix used for the simulation to reproduce the effect of different shapes of the initial uncertainty distribution on the performance of the LS method. In particular, it is chosen to perform a transformation composed of a rotation into the local-vertical-local-horizontal frame, followed by a squeezing transformation in the along-track direction expressed by the matrix **F** defined as

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{sq} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{sq} \end{bmatrix}, \quad \mathbf{F}_{sq} = \begin{bmatrix} 1/f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1/f \end{bmatrix} \quad (3.1)$$

where the elements of the rotated covariance matrix referring to the along-track direction (of both the position and the velocity) are increased by a factor $f$, while the components in the radial and normal-to-plane directions are reduced by the same factor, in order to preserve the total volume of the uncertainty distribution.

This choice was made following the observations reported in [3], where the method was successfully applied to different test cases involving the propagation of NEOs, as in those cases the initial uncertainty distribution appeared highly elongated in the along-track direction of the orbit.

### *3.2 Dynamical model and propagation setup*

The propagations are carried out in Cartesian coordinates with respect to an EME2000 reference frame centred in the Solar System Barycentre (SSB), with the inclusion of the gravitational contributions of the Sun, all the major planets, and the Earth's moon. Most of the physical constants (gravitational parameters, planetary radii, etc.) are obtained from the JPL Horizons database via the SPICE toolkit*.

Propagations are stopped according to 3 conditions: the maximum time is reached; an impact with one of the included celestial bodies occurs; an escape from the SOI of the Sun occurs.

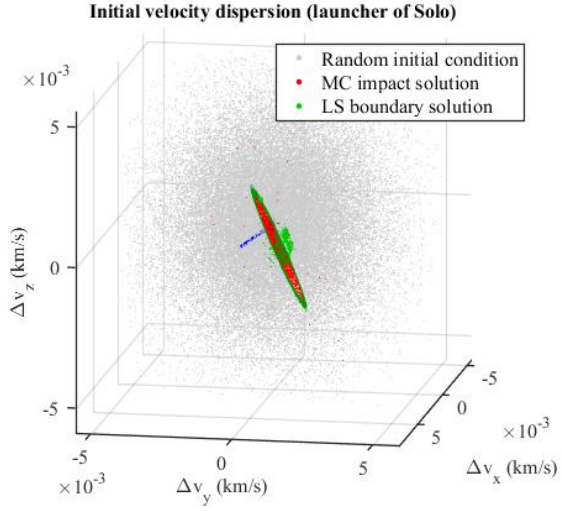
The propagations are carried out with the use of an 8th order Dormand-Prince RK method, with an embedded scheme to adapt the time-step (already available in SNAPPshot) with absolute and relative tolerances of $10^{-12}$.

### *3.3 Results*

In this Section the results of the application of LS to the selected test case for different shapes of the initial uncertainty, obtained by setting different values for the squeezing factor of the covariance matrix.

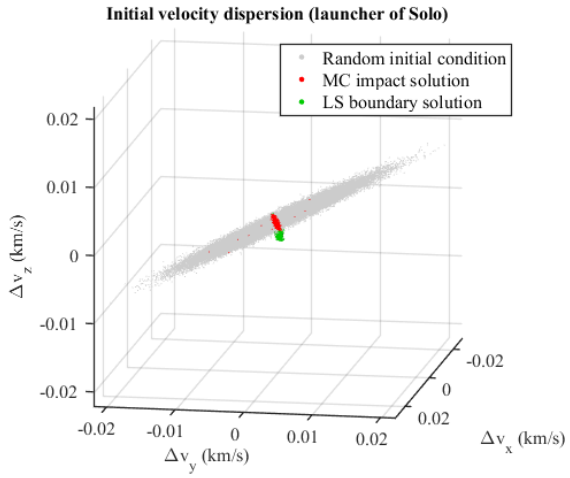
Fig. 4 shows the uncertainty distributions in two different cases: the unmodified one in Fig. 4a, and one elongated in the along-track direction using a squeezing factor *f=16* in Fig. 4b, with the impact region (found via standard MC) highlighted in red, and its boundary. The results of the corresponding simulations are reported in Table 1 and Table 2 respectively, in terms of number of random samples, number of orbital propagations, impact probability and the relative standard deviation (which is used as a measure of the accuracy of the methods, with smaller values corresponding to a higher accuracy).

Note that in the LS more propagations are performed than in standard MC due to the numerical iterations necessary to identify the zeroes of the performance function that defines the border of the impact region, as already specified in [2]. In particular, in all cases presented here 10 iterations were used in order to identify the border to ensure a correct identification of the boundary.




**a)**

* https://naif.jpl.nasa.gov/naif/




**b)**

**Fig. 4 – Representations of the initial velocity dispersions for the launcher upper stage of Solo mission, shaped differently according to a squeezing factor *f=1* (no squeezing) in (a), and *f=16* in (b). The initial conditions leading to an impact with Venus (and identified via standard MC simulation) are shown in red, while the boundary of the impact region (computed via LS) is shown in green. The blue arrow represents the sampling direction.**

| | $N_{samples}$ | $N_{prop}$ | $\hat{P}(I)$ | $\hat{\sigma}$ |
|---|---|---|---|---|
| **MC** | 54114 | 54114 | 4.34e-2 | 8.76e-4 |
| **LS** | ~46000 | ~460000 | 4.63e-2 | 4.41e-4 |

**Table 1 – Results of the application of standard MC and LS to the test case with unmodified initial distribution (squeezing factor *f=1*), as in Fig. 4a.**

| | $N_{samples}$ | $N_{prop}$ | $\hat{P}(I)$ | $\hat{\sigma}$ |
|---|---|---|---|---|
| **MC** | 54114 | 54114 | 2.01e-2 | 6.15e-4 |
| **LS** | ~46000 | ~500000 | 2.01e-2 | 2.15e-4 |

**Table 2 – Results of the application of standard MC and LS to the test case with modified initial distribution (squeezing factor *f=16*), as in Fig. 4b.**

From Fig. 4 one can see that the impact regions with Venus in the two cases are identified by the same initial conditions, but while in case (a) the impact region is lumped and all contained inside the uncertainty distribution, in case (b) the impact region goes from side to side of the distribution, thus representing one of the favourable cases already shown in [3]. This is confirmed by the results reported in Table 1 and Table 2, showing that the value of standard deviation given by the LS, already lower than the one of the MC, decreases when the distribution has an elongated shape.

Similar considerations can be made by considering Fig. 5, where the variation of the values of impact probability and the associated standard deviation for more values of the squeezing factor *f*. It shows that even a low elongation of the initial distribution can decrease the value of the standard deviation, thus improving the accuracy of the LS.

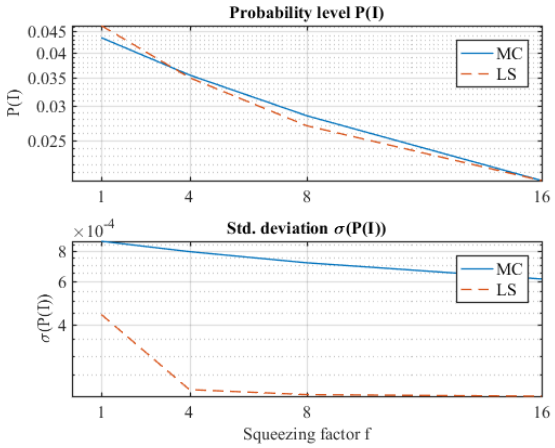



**Fig. 5 – Variation of the impact probability with Venus (top) and of the associated standard variation (bottom) with the variation of the squeezing factor *f*, comparing standard MC and LS.**

## 4. Conclusions

The work presented here describes some developments of the state of research about the LS method, by providing a better understanding of the performance of the method with respect to the one of the standard MC simulations. This is done both theoretically, by providing an approximated formula that highlights the dependency of the method both from the level of impact probability (as already proven by the existing literature and the previous works related to the method) and from the shape of the impact region, and numerically, by providing a test cases devised to show how the accuracy of the LS depends on the shape of the initial uncertainty distribution. The information gained from this work can be used to identify in advance in which cases the LS (compared with the standard MC) will be more efficient (in terms of number of random samples needed to reach a given confidence level) depending on the expected impact probability and the shape of the initial distribution.

Future work to further improve tools for PP analysis will focus on the extension of the LS algorithm to the case of multiple impact events with different bodies and

the improvement of the preliminary analysis to identify impact conditions. Aside from sampling methods, different ways and parameterisations to express the initial uncertainties will be explored to make the sampling more efficient, together with the direct propagation of uncertainties.

**Acknowledgements**

The work performed for this paper has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 679086 – COMPASS), and from the European Space Agency (ESA) though a Networking/Partnering Initiative (NPI).

**References**

[1] Kminek G., *ESA planetary protection requirements, Technical Report ESSB-ST-U-001*, European Space Agency, February 2012.

[2] Romano M., Colombo C., Sánchez Pérez J. M., *Verification of planetary protection requirements with symplectic methods and Monte Carlo Line Sampling*, 68th International Astronautical Congress (IAC) 2017, Adelaide, Australia, September 25-29, 2017 (IAC-17-C1.9.5).

[3] Romano M., Losacco M., Colombo C., Di Lizia P., *Estimation of impact probability of asteroids and space debris through Monte Carlo Line Sampling and Subset Simulation* (unpublished manuscript).

[4] Zio E., *The Monte Carlo Simulation Method for System Reliability and Risk Analysis, 1st edn., Springer Series in Reliability Engineering*, Springer-Verlag London (2013).

[5] Colombo C., Letizia F., Van den Eynde J., Jehn R., *SNAPPshot, ESA planetary protection compliance verification software. Final report*, ESA 2016, ESA ref. ESA-IPL-POM-MB-LE-2015-315.

[6] Letizia F., Colombo C., Van den Eynde J., Armellin R., Jehn R., *SNAPPSHOT: Suite for the numerical analysis of planetary protection*, 6th International Conference on Astrodynamics Tools and Techniques (ICATT), 14-17 Mar. 2016, Darmstadt.

[7] Debatin F., Tilgner A., Hechler F., *Fast numerical integration of interplanetary orbits*, In Second International Symposium on Spacecraft Flight Dynamics, 1986.

[8] European Space Agency, *Solar Orbiter, Exploring the Sun-heliosphere connection*, Definition Study Report, ESA/SRE(2014)11, July 2011.