

Estimating the Probability of Close Approaches Generated by an In-Orbit Fragmentation via Monte Carlo-Based Approach

Yeerang Lim^{1,a,*}, and Camilla Colombo^{1,b}

¹Department of Aerospace Science Technology, Politecnico di Milano, Via La Masa 34, 20133 Milan, Italy

^ayeerang.lim@polimi.it, ^bcamilla.colombo@polimi.it

Keywords: Probability Estimation, Conjunction Assessment (CA), Monte Carlo Simulation, Line Sampling (LS), Subset Simulation (SS)

Abstract. Advanced Monte Carlo-based approaches are investigated for conjunction assessment with a breakup event, namely Line Sampling and Subset Simulation, which enable to detect the same extent of probability with fewer samples. The proposed approaches are applied to a hypothetical fragmentation and compared with the standard Monte Carlo approach.

1. Introduction

Probability estimation of Close Approaches (CAs) caused by a fragmentation is one of the key components for orbital neighbourhood awareness conjunction assessment. While the distribution of the fragments is often derived from the NASA Standard Breakup Model, a Monte Carlo simulation with large number of samples is required from such a model for a conjunction risk assessment, especially when the probability is small.

On the other hand, advanced Monte Carlo methods, namely line sampling and subset simulation can be used. Both approaches were originally developed to estimate structural failure [1,2], similar to the CA assessment, the system failure probability is the fraction of the number of samples which lead to the failure region. From its nature, a large number of samples are required for standard MC method to guarantee the accuracy of the estimated probability, which causes high computational cost. Both line sampling and subset simulation reduce this problem, by transforming it to a one-dimensional problem in the standard normal space, or by reformulating it as a combination of multiple conditional probabilities. Smaller number of samples will be required as a result, to achieve the same level of accuracy. The same approaches were applied later to the conjunction analysis of near-Earth objects for planetary projection as well [3].

In this work, we propose a line sampling and subset simulation approach to estimate the probability of close approach against debris fragments, to reduce the number of required samples and enhance computational efficiency and improve accuracy in the estimation of the collision probability. Instead of the uncertainty covariance of the initial condition, the probability distribution function from the breakup model can be treated in the same manner. The sample fragmentations are numerically propagated using high-fidelity gravity model implemented in ESA GODOT [4]. The performance of the proposed approaches is compared with the standard Monte Carlo simulation under the same condition, and the feasibility and the limitations of the proposed approaches are addressed by numerical simulations.

2. Advanced Monte Carlo-Based Approaches

Line Sampling

The line sampling (LS) method is a Monte Carlo-based approach for the estimation of small probabilities. It was originally developed for the reliability analysis of failure in complex structural systems [5,6], and later applied to risk estimation for orbital conjunction analysis in combination with differential algebra [3,7]. Both problems have very low probability (generally in the range of

10^{-7} to 10^{-6}) and therefore require a large number of samples for probability estimation. The main feature of the Line Sampling method is the analytical estimation of the probability, obtained by reducing the multi-dimensional integration problem across the uncertainty domain to many one-dimensional integrals 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 increase the accuracy of the solution or reduce the number of required system simulations with respect to a standard MC. It consists of 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 α ; 3) the probing of the impact region along the lines following the reference direction; 4) the estimation of the impact probability [8].

The first step of the LS procedure is to map the physical coordinates $\mathbf{x} \in R^n$ onto the standard normal space $\boldsymbol{\theta} \in R^n$, whose components are all associated with a unit Gaussian distribution. The components of the vector $\boldsymbol{\theta}$ have the property of being independent and identically distributed random variables. In other words, each component has the same probability distribution as the others and all are mutually independent. Therefore, the joint PDF ϕ of these random parameters is equal to the product of all the single PDFs.

$$\phi(\boldsymbol{\theta}) = \prod_{j=1}^n \phi_j(\theta_j) \quad (1)$$

Thanks to this property, this transformation grants efficiency to the method, especially with high dimensions. It reduces the problem to a series of one-dimensional analytical evaluations, thus enabling a simplification of the computation of the probability. The direct and the inverse transformation, from the physical domain to the normalized one and vice versa, which preserve the joint cumulative distribution function (CDF) between the two spaces, are defined as below.

$$\begin{aligned} \Phi(\boldsymbol{\theta}) &= F(\mathbf{x}) \\ \boldsymbol{\theta} &= \Phi^{-1}[F(\mathbf{x})] \\ \mathbf{x} &= F^{-1}[\Phi(\boldsymbol{\theta})] \end{aligned} \quad (2)$$

Then the reference direction α is determined, as the direction of a normalized “centre of mass” of the impact region in [3,8].

$$\alpha = \frac{1}{N_s} \sum_{u=0}^{N_s} \frac{\boldsymbol{\theta}^u}{\|\boldsymbol{\theta}^u\|} \quad (3)$$

This region is approximated by applying the Metropolis-Hastings algorithm [9,10] as below, to generate a Markov Chain Monte Carlo (MCMC) lying entirely in the impact subdomain starting from an initial condition within it.

- from the current element $\mathbf{x}_u$, ξ is generated by randomly sampling a user-defined "proposal" pdf $p^*(x_u)$;
- compute the ratio $r = p(\xi)/p(\mathbf{x}_u)$;
- set $\tilde{\mathbf{x}} = \xi$ with probability $\min(1, r)$ and set $\tilde{\mathbf{x}} = \mathbf{x}_u$ with the remaining probability $1 - \min(1, r)$, where $\tilde{\mathbf{x}}$ is the candidate for the next element of the chain;

- d. check whether the candidate $\tilde{\mathbf{x}}$ lies in the region of interest I or not: if $\tilde{\mathbf{x}} \in I$, accept it as the next sample $\mathbf{x}_{u+1} = \tilde{\mathbf{x}}$; else, reject it and take the current sample as the next sample $\mathbf{x}_{u+1} = \mathbf{x}_u$.

After determining the reference sampling direction, N_T initial conditions $\mathbf{x}^k$ are randomly drawn from the nominal uncertainty distribution and then mapped to standard normal coordinates as $\boldsymbol{\theta}^k$ using the transformation above. For each sample in the standard normal space, a line starting from $\boldsymbol{\theta}^k$ and parallel to $\boldsymbol{\alpha}$ is defined according to the parameter c^k , such that

$$\begin{aligned}\boldsymbol{\theta}^k &= c^k \boldsymbol{\alpha} + \boldsymbol{\theta}_{\perp}^k \\ \boldsymbol{\theta}_{\perp}^k &= \boldsymbol{\theta}^k - \langle \boldsymbol{\alpha}, \boldsymbol{\theta}^k \rangle \boldsymbol{\alpha}\end{aligned}\quad (4)$$

where $\langle \boldsymbol{\alpha}, \boldsymbol{\theta}^k \rangle$ is the scalar product. In this way, the problem is reduced to a series of one-dimensional evaluations associated with each sample, with c^k also normally distributed in the standard space.

The standard domain is then explored along each line with the aim of identifying the impact region of an uncertainty set. For the problem under study, the occurrence of the events is expressed by the value of the non-dimensional performance function

$$g_{\theta}(\boldsymbol{\theta}(c)) = Y(c)/R_{imp} - 1 \quad (5)$$

where Y is a function that maps the value of the parameter c to a performance index, an R_{imp} is the selected critical distance. By setting Y as the minimum distance from the body of interest in the given time of window, while R_{imp} as the radius around the body to define an impact, the performance can be written as

$$g_{\theta}(\boldsymbol{\theta}(c)) \begin{cases} < 0 \rightarrow \text{Impact} \\ = 0 \rightarrow \text{Limit condition} \\ > 0 \rightarrow \text{No impact} \end{cases} \quad (6)$$

There will be a maximum of two intersections between each line and the impact region, meaning that two values of c^k exist for each standard normal random sample $\boldsymbol{\theta}^k$. Once the value $(\bar{c}_1^k, \bar{c}_2^k)$ are known for all the sampling lines, the unit Gaussian CDF provides each random initial condition $\boldsymbol{\theta}^k$ with the conditional impact probability $\hat{P}^k(I)$, which can be computed as a one-dimensional integrals based on the properties of the normalized coordinate space.

$$\hat{P}^k(I) = \hat{P}[\bar{c}_1^k < N(0,1) < \bar{c}_2^k] = \int_{\bar{c}_1^k}^{\bar{c}_2^k} \phi(u) du = \Phi(\bar{c}_2^k) - \Phi(\bar{c}_1^k) \quad (7)$$

The total probability of the event $\hat{P}(I)$ can be approximated as follows.

$$\hat{P}(I) = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{P}^k(I) \quad (8)$$

Subset Simulation

The subset simulation (SS) method is an adaptive stochastic simulation method to compute efficiently small failure probabilities [3,7]. The approach was originally developed to address structural failure, but it was also used in different research areas in reliability, e.g. to address the failure probability of thermal-hydraulic passive system [5]. The basis of the method is to compute the probability as a product of larger conditional probabilities. Given a sequence of intermediate failure regions $F_1 \supset F_2 \supset \dots \supset F_m = F$, the failure probability becomes

$$P(F) = P(F_m) = P(F_1) \prod_{l=1}^{m-1} P(F_{l+1} | F_l) \quad (9)$$

where $P(F_{l+1} | F_l)$ indicates the probability of F_{l+1} conditional to F_l . The method is initialized using a standard MC simulation to generate samples at conditional level 0. Once the failure region F_1 is determined and the probability $P(F_1)$ computed, a Monte Carlo Markov Chain (MCMC) algorithm is used to generate samples conditional to the failure region F_1 . The next intermediate failure region F_2 is then located and other samples are generated with MCMC. The process can be repeated till the failure region corresponding to objects collision is identified. Similar to the LS method, the MCMC Metropolis-Hastings algorithm is to generate conditional samples in the new region [10]. The resulting SS algorithm can be performed through the following steps [7].

- set $l = 0$ and generate N sample vectors $\mathbf{x}_0^{0,k}$ by standard MC simulation.
- propagate each sample for the selected time window, and a list of possible target events is obtained. And compute the values of the performance function $g_x^l(\mathbf{x}_0)$ for the N samples $\mathbf{x}_0^{l,k}$.
- sort the N samples in ascending order, according to their values of the performance function g_x^l . The samples closer to the failure region will be at the bottom of the list.
- identify the intermediate threshold value D_{l+1} from the $(1 - p_0)N$ th value of the sorted list. Define the $(l + 1)$ th conditional level as $F_{l+1} = \{d_{\Delta t_e} < D_{l+1}\}$. Considering how the threshold is defined, the associated conditional probability $P(F_{l+1} | F_l) = p_0$
- if $D_{l+1} < D_{\lim}$, where $D_{\lim}$ is the target threshold distance, go to the last step. Otherwise select the last $p_0 N$ samples of the list. By definition, these samples belong to the $(l + 1)$ th conditional level.
- using MCMC, generate $(1 - p_0)N$ additional condition samples starting from the previously selected seeds belonging to F_{l+1} . A sample is set to belong to F_{l+1} according to the following performance function.
- set $l = l + 1$ and return to step b.
- stop the algorithm.

The total number of generates samples is

$$N_T = N + (m - 1)(1 - p_0)N \quad (10)$$

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

$$\hat{P}(F) = \hat{P}(F_m) = P(F_m | F_{m-1}) p_0^{m-1} = p_0^{m-1} N_F / N \quad (11)$$

where N is the total number of samples at each conditional level and N_F is the number of samples at conditional level m , whose distance from the target is lower than D_{lim} .

3. Simulation Studies

The proposed LS and SS approaches are applied to a breakup scenario in Table 1, together with the standard Monte Carlo, for both explosion and non-catastrophic collision cases.

Table 1 Breakup simulation scenario

Parent orbit @ 2025-01-01 00:00:00					
SMA [km]	Ecc	Inc [deg]	RAAN [deg]	AoP [deg]	TA [deg]
8000.8008	0.0001	98.0	50.0	0.0	180.0
Target orbit @ 2025-01-01 00:00:00					
SMA [km]	Ecc	Inc [deg]	RAAN [deg]	AoP [deg]	TA [deg]
8248.307	0.03	80.0	50.0	0.0	188.0408

Conjunction assessment results after 2 hours from the breakup are summarized in Table 2. Taking into account the fact that each Monte Carlo simulation is different, it is noticeable that LS gives the same order of probability with a one hundredth of samples. If the standard deviations of the computed probabilities are considered, advanced Monte Carlo-based approaches provide smaller error covariances, meaning the LS is more reliable than the standard Monte Carlo when the number of samples is the same, or even smaller.

Table 2 Probability of close approach < 5 [km]

	Explosion	Collision
Standard MC ($N = 100,000$)	6.3282e-4	3e-5
LS ($N = 1000$)	2.5442e-4	2.7482e-5

A couple of points shall be added. First, the essential idea of the LS is based on the Gaussian distribution of the samples, which is the assumption used in the domain transformation. The distribution of the fragments from the breakup model, on the other hand, is not Gaussian, which explains the difference between two probabilities. Moreover, we suspect that this distribution becomes even less Gaussian along the propagation. It is expected the LS cannot find the $(\bar{c}_1^k, \bar{c}_2^k)$ at a certain point. Second, we decided to discard the SS at the moment, not because it is not able to calculate the probability, but the value is strongly dependent on the distribution used in the MCMC, and it is not clearly figured out yet which distribution shall be used. The SS is still a considerable approach when the objective of the application is to detect possible close approaches, however the proposed distribution part should be investigated further for probability estimation.

4. Conclusion

Line sampling (LS) and subset simulation (SS) are applied to the conjunction assessment with a breakup event, which are alternatives to calculate a small probability with less random samples. While the approaches detect 1e-4 to 1e-5 probability of close approaches, we also notice some limitations, which are worth investigating further. One possible future work is the sensitivity analysis of the LS with respect to propagation time, to check when it starts not finding the roots in the standard normal domain. Another idea is to think of a way to keep the samples distribution Gaussian as possible, e.g. by introducing Gaussian mixture model. Last but not least, Markov

Chain in the SS could be examined further, to decide which proposed distribution provides the most realistic probability.

Acknowledgement

The research in this work has received funding from the European Space Agency, ESA Contract No. 4000146083/24/D/SR.

References

- [1] G.I. Schuëller, H.J. Pradlwarter, P.S. Koutsourelakis, A critical appraisal of reliability estimation procedures for high dimensions. *Probab. Eng. Mech.* 19 (2004) 463-474. <https://doi.org/10.1016/j.pro bengmech.2004.05.004>
- [2] S.-K. Au, J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation. *Probab. Eng. Mech.* 16 (2001) 263-277. [https://doi.org/10.1016/S0266-8920\(01\)00019-4](https://doi.org/10.1016/S0266-8920(01)00019-4)
- [3] M. Romano, M. Losacco, C. Colombo, P. Di Lizia, Impact probability computation of near-Earth objects using Monte Carlo line sampling and subset simulation. *Celest. Mech. Dyn. Astr.* 132 (2020) 42. <https://doi.org/10.1007/s10569-020-09981-5>
- [4] Information on <https://godot.io.esa.int/godotpy/>
- [5] E. Zio, N. Pedroni, Subset simulation and line sampling for advanced Monte Carlo reliability analysis, in *Proceedings of the European Safety and RELiability (ESREL) Conference* (2009).
- [6] H. Pradlwarter, M. Pellissetti, C. Schenk, G. Schueller, A. Kreis, S. Fransen, A. Calvi, and M. Klein, Realistic and efficient reliability estimation for aerospace structures, *Computer Methods in Applied Mechanics and Engineering*. 194 (2005) 1597-1617. <https://doi.org/10.1016/j.cma.2004.05.029>
- [7] A. Morselli, R. Armellin, P. Di Lizia, F.B. Zazzera. A high order method for orbital conjunctions analysis: Monte Carlo collision probability computation. *Adv. in Space Research*. 55 (2015) 311-333. <https://doi.org/10.1016/j.asr.2014.09.003>
- [8] M. Romano, Orbit propagation and uncertainty modelling for planetary protection compliance verification, Doctoral thesis, Politecnico di Milano (2020). Supervisors: Colombo, C., and Sánchez Pérez, J. M.
- [9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equations of state calculations by fast computing machines, *J. Chem. Phys.* 21 (1953). 1087-1092. <https://doi.org/10.1063/1.1699114>
- [10] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*. 57.1 (1970). 97-109. <https://doi.org/10.1093/biomet/57.1.97>