Satellites collision probabilities: switching from numerical integration to the adaptive splitting technique

Rudy Pastel[⋆,⋆⋆] *and Jérôme Morio*[⋆] *and François Legland*[⋆⋆].
[⋆]*Onera, Chemin de la Hunière BP 80100 FR-91123 Palaiseau Cedex, France*
[⋆⋆]*INRIA Rennes, Campus de Beaulieu 35042 Rennes, France*

Abstract

Spacecraft collision still does happen seldom, but the loss of a satellite can not be afforded: this high risk therefore has to be addressed carefully. To support the decision to start a collision avoidance maneuver, a dedicated tool is the probability of collision between the debris and the satellite [12]. Crude Monte Carlo (CMC) could be a way if it could cope with very small probabilities, say 10^{-6}, within the available simulation budget and time. The methodology nowadays in use is a numerical integration made tractable by physical hypothesis and numerical approximation [7]. We advocate the Adaptive Splitting Technique (AST), presented in [4], to our purpose as it avoids all the hypothesis needed for the numerical integration and clearly outperforms CMC w.r.t. rare events. However, AST requires tuning. We provide experience based empirical tuning rules to foster a wider spreading of the AST in the applied world.

1. Introduction

The February 10^{th} 2009, tough their configuration did not appear troublesome [12], active commercial satellite Iridium and out of order Russian satellite Cosmos collided. The impact produced number of smaller debris. Most of them can destroy any artificial space object, whether in use or not, they might encounter. Space debris *are* a threat to human space activity.

Dealing with space debris rises a number of questions. What to do with active satellites turning inactive? How to clean at least the frequently used orbits? How to design future satellite so that they pollute as less as possible when they go out of order? How to ensure the safety of active satellites? We focused on this last question[1].

The safest practice with respect to satellites is avoiding collision. Avoidance maneuvers are efficient but costly and reduce the possible usage time. However, there is no point in saving fuel if it ends up spread in space after a collision. Satellite safety responsible teams have to design a trade–off between fuel saving and collision avoiding. Avoidance maneuvers are decided, among other parameters, based on the estimated collision probability.

The Iridium–Cosmos collision was an unlikely event which probability cannot be calculated manually and therefore as to be estimated. It even qualifies as a rare event *i.e.* an event which probability is less than, say, $p = 10^{-6}$. Crude Monte Carlo (CMC) is unable to estimate such a low probability with a reasonable cost: about 10^8 simulations are needed to perform an estimate whose standard deviation is a tenth of p. This is not affordable and a specific tool is required.

According to our knowledge, the current methodology in the NASA [7] to estimate the collision probability between two orbiting objects is integrating a Rician probability density function (pdf) over a circular sub domain of the collision plan. Actually, the Gaussian uncertainty with respect to the real location and speed of the satellites when last observed *via* Earth based radars is deemed to remain Gaussian as the satellites go on their tracks, although the dynamics is nothing but linear. Said tracks are assumed, under specified hypothesis, to be straight lines in the encounter region so as to conveniently define a collision plan. The error ellipsoids are combined then projected on the collision plan and eventually numerically integrated.

[1]This paper is somewhat between a follow–up and a correction of [20]

However easily implemented and fast to calculate this method is, the two hypothesis it is based on are major drawbacks.

1. Uncertainty linear propagation.
2. Straight line tracks in the vicinity of the time of closest approach.

We find these hypothesis very strong and looked for an other way.

We therefore searched the existing rare event related literature to find an affordable, hypothesis free rare event dedicated probability estimation technique. This paper tells this quest. We first introduce the spacecraft collision issue via the Iridium-Cosmos case. We then try to estimate the collision probability through Crude Monte Carlo (CMC) in section 3. Next, Cross–entropy and the non parametric adaptive importance sampling, two rare event probability estimation techniques, are introduced and tried out as well. In section 5 we present the Adaptive Splitting Technique (AST), the special rare event probability estimation technique we advocate to our purpose, and show it can provide valuable accurate information in this framework and give some insight with respect to its tuning.

2. Spacecraft encounter and noised state measurement

On February 10, 2009, a commercial Iridium communications satellite and a defunct Russian satellite Cosmos collided, though their configuration was not reported dangerous [12]. In this section, the question "what was the probability it happened?" is formalised.

In order to understand the predicament an active spacecraft managing team can find itself in, we will describe the geometrical issue at hand and then the main source of randomness.

2.1 A basic geometry problem

Consider two satellites orbiting around the Earth in a Galilean frame of reference with our planet as origin and equipped with the Euclidean distance. This three-body problem will be considered a double two-body problem: each satellite interacts through gravity only with the Earth and not with the other satellite. Besides, the Earth and the satellites are assumed to be homogeneous spheres with radii d_E, d_1 and d_2. The collision distance is therefore $d_c = d_1 + d_2$. We wonder about the relative position of the two satellites: might they collide during a given time span $I = [t_s, t_e]$?

2.2 A convenient model of dynamics

To keep things simple, orbital mechanics being not our topic, we will use Kepler mechanics. One can use more advanced models such as SGP4 [17] if wanted. The discussed probability estimation methodologies are independent of the method.

At time t, the satellites will be represented by their states $\vec{s}_1(t)$ and $\vec{s}_2(t)$ i.e. their positions $\vec{r}_1(t)$ and $\vec{r}_2(t)$ and their speeds $\vec{v}_1(t)$ and $\vec{v}_2(t)$ such that $\vec{s}_i = (\vec{r}_i, \vec{v}_i)$.

In our setting, the speeds evolve according to the same well-known Ordinary Differential Equations defining the two body problem

$$\forall t, \frac{d\vec{v}_i}{dt} = -a\frac{\vec{r}_i}{r_i^3} \tag{1}$$

where a is a positive constant given by physics.

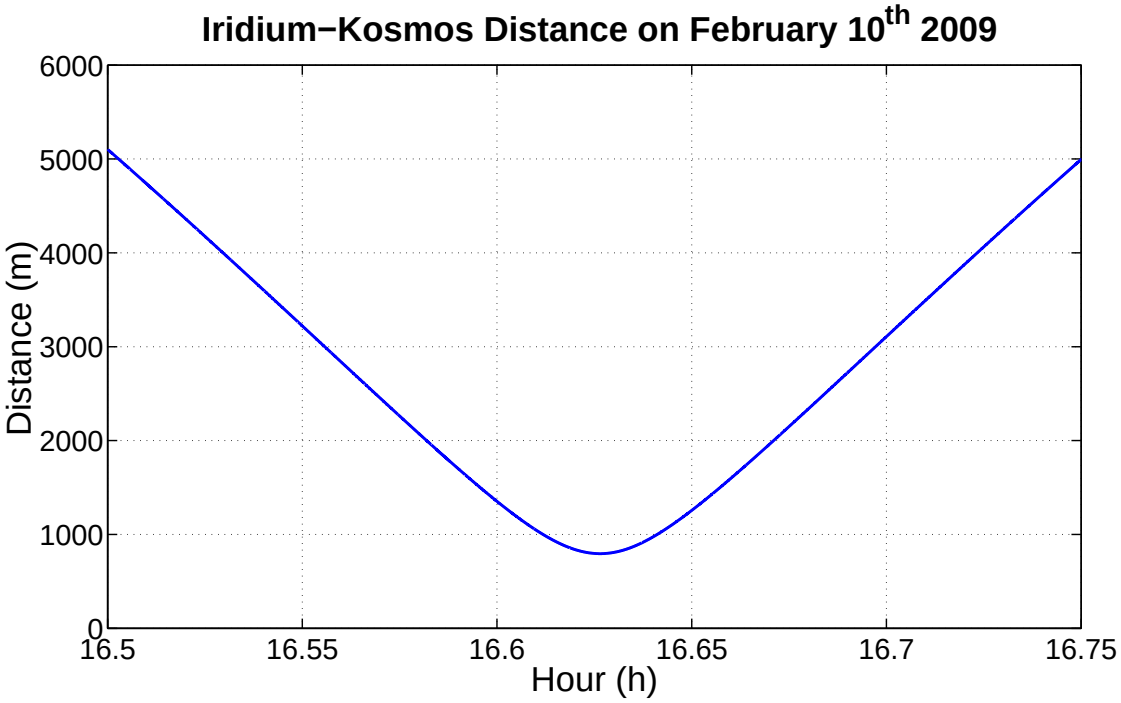
This ordinary differential equation (*ode*) is analytically solved in many textbooks and its solution depends continuously and in a bijective fashion on the given so called initial conditions: its value $\vec{s}_i^m$ at t_i^m, the measurement time, through Φ the ode's resolvant i.e. its solution map:

$$i \in \{1, 2\}, \forall t \in I, \vec{s}_i = \Phi(\vec{s}_i^m, t_i^m, t) \tag{2}$$

At this point, there is a natural way to clear out the collision issue using

$$\delta = \min_{t \in I}\{\|\vec{r}_2 - \vec{r}_1\|(t)\} \tag{3}$$

Figure 1: Iridium-Comos distance on collision day according to TLE.



$t \in I \mapsto \|\vec{r}_2 - \vec{r}_1\|(t)$ experimental convexity, figure 1, makes δ available through numerical optimisation and its associated test:

$$\xi(\vec{s}_1^m, t_1^m, \vec{s}_2^m, t_2^m, I) = \begin{cases} 1 & \text{if} \quad \delta \leq d_c \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

eventually closes the deal. Things would be all that easy and deterministic, had randomness not barged in.

2.3 Random measurements lead to uncertainty

In real life, the states are not monitored around-the-clock: they are merely measured from times to times, by a radar. The Two Line Elements (TLE) provided by NORAD sum up this information and feed the models with the $(\vec{s}_i^m, t_i^m)$ pairs. However, TLEs are inaccurate and their inaccuracy is unknown. This uncertainty is our very issue. To cope with this and to better reflect the reality, we added independent, identically distributed (iid) noises $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$, with density $f_{\vec{\epsilon}}$, to the models' inputs s_i^m.

$$\vec{E} = \begin{pmatrix} \vec{\epsilon}_1 \\ \vec{\epsilon}_2 \end{pmatrix} \qquad\qquad f_{\vec{E}}(\vec{e}) = f_{\vec{\epsilon}}(\vec{\varepsilon}_1) \times f_{\vec{\epsilon}}(\vec{\varepsilon}_2) \tag{5}$$

The collision issue can not be answered in a cut-and-dried way anymore as it has to be rephrased in a probabilistic fashion itself: what is the probability of collision between the two satellites?

Via the random counterpart of our deterministic geometrical problem

$$\begin{array}{rcccc} i \in \{1,2\}, \forall t \in I, & \vec{S}_i & = & \Phi(\vec{s}_i^m + \vec{\epsilon}_i, t_i^m, t) = (\vec{R}_i(t), \vec{V}_i(t)) \\ & \Delta & = & \min_{t \in I}\{\|\vec{R}_2 - \vec{R}_1\|(t)\} \\ & \Xi & = & \xi(\vec{s}_1^m + \vec{\epsilon}_1, t_1^m, \vec{s}_2^m + \vec{\epsilon}_2, t_2^m, I) \end{array} \tag{6}$$

this question is equivalently stated as

$$\mathbb{P}[\{ \text{ The satellites collide during } I\}] = \mathbb{P}[\Delta < d_c] = \mathbb{E}[\Xi] \tag{7}$$

We now face a plain expectation estimation problem.

3. The Crude Monte Carlo

As an explicit analytical way to calculate $\mathbb{E}[\Xi]$ is unlikely to be found, one will most likely make do with an estimation. Crude Monte Carlo (CMC) is a very convenient and reliable way to reach it, if the sought probability is not too low, indeed: CMC can not handle rare event probability estimation at a reasonable cost.

3.1 A basic set of tools and notations

CMC's estimators for Ξ's expectation $\mathbb{E}[\Xi]$ and variance $\mathbb{V}[\Xi]$ are defined respectively as μ, the empirical mean and σ^2, the empirical variance of n *iid* tests, $\Xi^i, i \in \{1, \cdots, n\}$:

$$\mu(\Xi, n) \equiv \frac{1}{n}\sum_{i=1}^{n} \Xi^i \qquad\qquad \sigma^2(\Xi, n) \equiv \frac{1}{n}\sum_{i=1}^{n} (\Xi^i - \mu(\Xi, n))^2 \tag{8}$$

As the empirical mean of *iid* random variables, $\mu(\Xi, n)$ is a random variable as well and we hope its variance is as small as possible with respect to its estimated mean. To measure this, m *iid* $\mu(\Xi, n)$ throws are made in order to calculate $\mu(\Xi, n)$'s empirical relative deviation (*Erd*) estimator ρ

$$\rho(\mu\,(\Xi, n)\,, m) \quad \equiv \quad \frac{\sigma(\mu(\Xi, n), m)}{\frac{1}{m}\sum_{i=1}^{m} \mu(\Xi, n)^i} \tag{9}$$

as a way to measure its accuracy as the ratio of its standard deviation over its empirical mean.

3.2 Why CMC can not cope with rare events

Using the independence of the Ξ^i and the fact that $\Xi^2 = \Xi$ for it is a binary 1 or 0 mapping, one can write the following easily.

$$\mathbb{E}[\mu(\Xi, n)] = \mathbb{E}[\Xi] \qquad\qquad \mathbb{V}[\mu(\Xi, n)] = \frac{\mathbb{E}[\Xi](1 - \mathbb{E}[\Xi])}{n} \tag{10}$$

If the sought probability $\mathbb{E}[\Xi]$ is 10^{-4}, to be accurate up to a tenth of the real value i.e. to have

$$\sqrt{\mathbb{V}[\mu(\Xi, n)]}/\mathbb{E}[\Xi] \leq 10^{-1}$$

almost $n \approx 10^6$ points are needed!

Most of the time though, in a real context, there is no way one can generate that many samples. Yet, this degree of accuracy is becoming a standard, for amounts of money at stake are huge (the Iridium satellite program is worth at least 200 M$) and safety standards more and more demanding.

3.3 A huge CMC estimation as reference

A huge Monte Carlo estimate was done to serve as a reference in our case. It benefited using fast to evaluate Kepler dynamics instead of SGP4 and an unreasonable more than a week calculation time on four computers.

$$n = 77 \cdot 10^6 \qquad\qquad \mathbb{P}\left[\Delta \leq d_c\right] \approx 1.15 \cdot 10^{-6} = \mathbf{p} \tag{11}$$

4. An overview of Importance Sampling

The Importance Sampling (IS) is an attempt at modifying the CMC so as to deal properly with a given rare event case [9, 1, 2] and more generally stands as a variance reduction technique. It takes advantage on the integral representation of the expectation of the random output of any transfer function ϕ with a random

input $X \sim f_X$ and conjugates it with the available extra information to propose a hopefully better suited input space sampling random variable $Z \sim f_Z$ called auxiliary random variable.

$$\mathbb{E}\left[\phi\left(X\right)\right] = \int_{\mathbb{X}} \phi\left(x\right) f_X\left(x\right) \mathrm{d}x = \int_{\mathbb{X}} \phi\left(z\right) \frac{f_X\left(z\right)}{f_Z\left(z\right)} f_Z\left(z\right) \mathrm{d}z = \mathbb{E}\left[\phi\left(Z\right) \frac{f_X\left(Z\right)}{f_Z\left(Z\right)}\right] \tag{12}$$

Formally, the only requirement to assert the existence of $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ via the Radon-Nikodym theorem is that probability density function f_X is absolutely continuous w.r.t. f_Z, *i.e.* $\forall \mathbf{A} \in \mathcal{X}, \left[\int_{\mathbf{A}} f_X = 0\right] \Rightarrow \left[\int_{\mathbf{A}} f_Z = 0\right]$ This way f_Z merely generates the very same events as f_X but with another probability, purposely making rare events more frequent.

The underlying hope is that with a good choice of f_Z, the IS estimator, which is unbiased, *can* yield reduced variance with respect to CMC, keeping its convenient implementation. Actually, once f_Z is chosen, it suffices to perform a CMC estimation of $\mathbb{E}\left[\phi\left(Z\right) \frac{f_X(Z)}{f_Z(Z)}\right]$. Eventually, it all boils down to choosing an appropriate auxiliary density.

4.1 Optimal auxiliary density design

Being of Monte-Carlo type, the Importance Sampling expectation estimator $\widehat{Y}_n$ is a random variable. Its performance depends on how appropriately chosen the auxiliary density is.

The ultimate goal is minimising $\mathbb{V}\left[\widehat{Y}_n\right]$, making it at least less than what could be achieve through CMC. Calculations provide an optimal auxiliary density.

$$\left[\mathbb{V}\left[\widehat{Y}_n\right] \text{ is minimal.}\right] \Leftrightarrow \left[\forall x \in \mathbb{X}, f_Z^*\left(x\right) = \frac{\left|\phi\right|\left(x\right) f_X\left(x\right)}{\mathbb{E}\left[\left|\phi\right|\left(X\right)\right]}\right] \tag{13}$$

Flabbergastingly, sampling according to f_Z^* directly comes as both impossible and pointless as it requires a normalizing constant ($\mathbb{E}\left[\left|\phi\right|\left(X\right)\right]$) as difficult to calculate as the very sought value. This, nonetheless gives an hint about about how to design the auxiliary density f_Z: as similar to f_Z^* as possible.

4.2 Cross–Entropy and Non parametric Adaptive Importance Sampling

Cross–Entropy (CE) and Non parametric Adaptive Importance Sampling (NAIS) are two ways of building a proxy to f_Z^*. CE defines f_Z^{CE} as the closest element to f_Z^*, in the sens of the Kullback-Leibler divergence, within a parametric pdf family [3, 8, 21]. The (NAIS) strategy is two folded: first build f_Z^{NAIS}, a proxy to f_Z^*, via a K probability kernel mixture, then use it to estimate the desired probability [19, 23].

We failed to use CE efficiently. We first used CE with the family of Gaussian distributions with diagonal variance matrices. Random minimal distance between satellites Δ was divided by a factor two to ten with respect to CMC, but no collision was simulated. CE was then used with the family of Gaussian distribution as whole, *i.e.* without any variance matrix structure hypothesis but symmetry and positivity. This led to small to no improvement at all.

As far as NAIS is concerned, experience showed us it is cumbersome and inaccurate when the sampling space dimension is over 10 [18], so it was not used.

The issue with CE is the choice of an adapted parametric pdf family. Being devoid of any analytical insight about the δ mapping, we are clueless about what is an appropriate choice. In the NAIS case, the same issue applies with respect to the the kernel shape and is topped with an increased calculation burden. The following method, Adaptive Importance Sampling, solves these two problems, yet deals efficiently with rare events.

5. The Adaptive Splitting Technique

So as to estimate the unlikely collision probability, a rare event dedicated technique is needed. We want it able to deal with a black box mapping and not too calculation consuming. The Adaptive Splitting Technique (AST) seems a good candidate.

The Splitting Technique (ST) is a rare event dedicated technique, a good introduction to which can be found in [15] and [16]. In order not to perform quite intractable analysis or integration w.r.t. ϕ, a ST variant that would adapt to the information gathered on the fly is used: the Adaptive Splitting Technique (AST) as explained in [6] for a dynamic case and in [5, 11] for the static case, such as the one here at hands.

5.1 An intuitive approach to AST

The basic idea here is *divide and conquer*: instead of estimating the very low probability directly, the work is divided in estimating a sequence of easier probabilities and eventually calculating the sought value as a plain product. This is the very purpose of this Bayesian formulation of our problem.

$$\mathbb{P}[\Delta \leq d_c] = \prod_{i=1}^{K} \mathbb{P}[\Delta \leq l_i | \Delta \leq l_{i-1}] \tag{14}$$

where $l_0 = \infty \geq l_1 \geq \cdots \geq l_K = d_c$ form a decreasing sequence of thresholds to be defined later. Hopefully, we have just reformulated our hard to estimate expectation as the product of easy to estimate conditional expectations!

The Adaptive Splitting Technique (AST) is a way of making this wish come true in an iterative three step way. Start with a sample of iid throws of Δ known to be under threshold l_i. With $i = 1$, this is only performing a plain CMC simulation. Then, and until the threshold is less than d_c, do as follows:

1. Define l_{i+1} as a the p–quantile of the current sample with $0.75 \leq p \leq 0.8$ [14, 10].
2. Resample uniformly among the realisations under the new threshold.
3. Use the selected points to sample new points conditionally to being under l_{i+1}.

When the threshold is lower than d_c, conclude that

$$\mathbb{P}[\Delta \leq d_c] \approx \prod_{i=1}^{K-1} p_i \times p^{\bullet}_{K-1} \tag{15}$$

where $p^{\bullet}_{K-1}$ is the estimated probability of collision given that the minimum relative distance is less than l_{K-1}.

The general idea of AST is simple but step 3 needs to be detailed.

5.2 Reversible kernel resampling

Step 3 is achieved thanks to a $f_{\vec{E}}$-reversible Markov kernel $M(\cdot,\cdot)$. $M(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{R}^d,\ M(x,\cdot) : \mathbb{R}^d \to \mathbb{R} \text{ is a density function.} \tag{16}$$

$\forall x \in \mathbb{R}^d$, $M(x,\cdot)$ stands as a x–specific random way to propose another $\mathbb{R}^d$ point. Let us now impose a constraint on M so as to respect the probability law. M is said to be a $f_{\vec{E}}$-reversible Markov kernel if

$$\forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d,\ f_{\vec{E}}(x)M(x,y) = f_{\vec{E}}(y)M(y,x) \tag{17}$$

This equation is known to physicists as the detailed balance equation [13]. It means that if from a $f_{\vec{E}}$ set, you use M to generate another, then

- the new set is distributed according to $f_{\vec{E}}$ as well: this is the invariance property.
- statistically, no one can say which set generated the other: this is the reversibility property[2].

[2] Reversibility $\Rightarrow$ Invariance.

Eventually, assume $X_i \sim f_{\vec{E}|i} = \frac{\mathbf{1}_{\Delta \leq l_i} f_{\vec{E}}}{\int \mathbf{1}_{\Delta \leq l_i} f_{\vec{E}}}$ and define

$$\psi_i(X_i) = \begin{cases} M(X_i, \cdot)\text{'s proposal (throw), } Y_i & \text{if } \Delta(Y_i) \leq l_i \\ X_i & \text{otherwise} \end{cases} \tag{18}$$

Thanks to reversibility, ψ_i is distributed according to $f_{\vec{E}|i}$ as well. We now can grow and inflate a $f_{\vec{E}|i}$–sample set thanks to the $f_{\vec{E}|i}$ available points via a $f_{\vec{E}}$-reversible Markov kernel.

However, the iterative nature of the algorithm is such that points share a common genealogy. That, at the end of the day, can translate into increased estimator variance. We have no theoretically proven way to avoid that yet. The intuition is that using functional composition[3] i.e. $\psi_i^{\circ(\omega)}$, leads to lower and lower variance as ω increases. It was shown in [22] that under mild conditions, $\omega > 1$ cannot increase variance and might even help. One can hence iterate ψ_i at will or based on a stopping time e.g. until 90% of the points moved from their original position.

5.3 The AST algorithm

Let us now state the AST algorithm.

Algorithm 5.1 (Adaptive Splitting Technique)**.** *So as to estimate $\mathbb{P}[\Delta \leq d_c]$, proceed as follows.*

1. *Generate η iid throws of Δ.*
2. *Set $\kappa = 1$.*
3. *Calculate the empirical p-quantile l_κ.*
4. *While $l_\kappa \geq d_c$, do:*
   (a) *Select throws under l_κ and discard others.*
   (b) *Replace discarded points resampling uniformly with replacement among selected points.*
   (c) *Apply $\psi_\kappa^{\circ(\omega)}$ to all the points.*
   (d) *Increment κ by one: $\kappa = \kappa + 1$.*
   (e) *Calculate the empirical p-quantile l_κ.*
5. *Estimate $p^\bullet_{\kappa-1}$.*
6. *Conclude that $\mathbb{P}[\Delta \leq d_c] \approx p^{\kappa-1} \times p^\bullet_{\kappa-1}$*

η, p, ψ and ω are fixed before hand. The number of quantiles i.e. κ's ultimate value, is random. The AST estimator will be denoted

$$\widetilde{Y} + \nu(d_c, \eta, p, M, \omega) \tag{19}$$

and the total number of generated points is

$$N = \eta \times (1 + ((\kappa - 1) \times \omega)) \tag{20}$$

6. Experimental Results

Let us now proceed to the application, describing the noise model, the actual AST parameters, especially the Markov kernel tuning, and comparing the results delivered by CMC and AST. Eventually, some insight about the AST results sensibility with respect to its parameters will be given.

Refer to table 1 for the definition of $f_{\vec{e}}$, $M(x, dy)$ and the tuning parameter α.

[3] Notation convention: $f^{\circ(n+1)} = f^{\circ(n)} \circ f$ and $f^{\circ(0)} = Id$.

6.1 Markov kernel tuning

Given a point $X \sim \mathcal{N}(0_6, D^2)$, the Markov kernel proposal is

$$Y \sim \frac{\alpha X + W}{\sqrt{1+\alpha^2}} \text{ where } W \text{ and } X \text{ are } iid \tag{21}$$

To choose $\alpha > 0$, there is no theoretical result yet. According to [4], one should make big steps at the beginning and make smaller and smaller steps, as the thresholds l_i decrease. The chosen heuristic is setting initially $\alpha = 1$ and adapt it in the course of the algorithm:

$$\alpha = \begin{cases} \alpha \times \theta & \text{if over 50\% of accepted transitions} \\ \alpha/\theta & \text{if under 50\% of accepted transitions} \end{cases}, \text{ where } \theta > 0. \tag{22}$$

We experimentally found out it leads to estimates with lesser variance than fixing α.

6.2 Comparison between CMC and AST

Using parameters in table 1, the results[4] in table 2 were obtained with a reasonable simulation budget. They show that AST estimated more accurately than CMC as its estimators has a way lesser relative variance as it was *divided by 5*, and for a very similar cost as it consumed on average the same amount of points.

As a matter of fact, AST outperforms CMC widely when it comes to estimating a rare event probability with a limited budget. For instance, using table 1 but setting

Table 1: Parameters

$$\text{TLEs} : 01/13/09 \quad D = 10^2 Diag(4,4,10,4,4,7) \quad f_{\vec{\epsilon}} \sim \mathcal{N}(0_6, D^2) \quad M(x, dy) \sim \mathcal{N}(\frac{\alpha x}{\sqrt{1+\alpha^2}}, \frac{1}{1+\alpha^2} D^2) \tag{23}$$

$$d_c = 10^4 \qquad n3 \cdot 10^5 \qquad m = 100 \qquad \theta = 1.1 \tag{24}$$

$$\eta = 1250 \qquad p = 0.75 \qquad \omega = 5 \tag{25}$$

$$\eta = 500 \qquad p = 0.20 \qquad \omega = 25 \tag{26}$$

one can have the following results

$$\eta = 1.0 \cdot 10{-}6 \qquad p = 42.94\% \qquad \omega = 89375 \qquad \omega = 4.4\% \tag{27}$$

that CMC cannot deliver.

Table 2: AST-CMC comparison

	Mean estimate	*Erd*	Mean simulation number	*Erd*
CMC	$1.1 \cdot 10^{-6}$	1.7258	300000	0
AST	$1.9 \cdot 10^{-6}$	0.3232	309060	0.0232

6.3 AST sensitivity

To test AST's sensitivity to its parameters, we changed some parameters to see their impact on the estimated value. Results are presented in table 3.

[4] Erd stands for Empirical relative deviation *i.e.* empirical standard deviation over mean ratio as explained at 9.

- Column–wise
  - Column 1 shows that if the "shaking parameter" θ does not evolve, regardless of how many times the transition kernel is applied, AST fails to provide an estimation. A possible explanation is that a static θ can not adapt itself so as to efficiently explore the input space.
  - Column 1.05 and 1.10 show that the more the transition kernel is applied, the closer to $\mathbf{p}$ the estimation and the lower the variance but the costlier the estimation. It is impossible to decide whether the variance decrease is due to the budget increase or a better usage of it.
- Row–wise
  - All three rows suggest that a gradual variation of α leads to a smaller variance.
  - Comparison across rows is difficult because then both N and ω change.

The best trade–off seems to have θ close to but different from 1 and as many kernel application as possible and, as could be expected, high η *i.e.* as many simulated points as possible.

Table 3: Kernel variance and iteration number influence with many cloud particles.
Experience parameters $\begin{matrix} \mathrm{d_{col}} & \eta & m & p & \kappa_{\max} \\ 10^4 & 2000 & 100 & \frac{4}{5} & 300 \end{matrix}$.

ω \ θ	1			1.05			1.1		
5	869.2	2992	$5.25 \cdot 10^{-26}$	205.0	657.5	0.40	180.5	640.1	0.6
	3.3	0	98.9	5.1	1.8	25.1	3.5	1.8	23.9
10	1784.8	5982	$9.45 \cdot 10^{-26}$	376.8	1235.6	0.96	340.2	1232	1.01
	5.4	0	79.6	5.8	1.6	19.9	5.9	1.5	20.7
20	3470.2	11962	$1.8 \cdot 10^{-25}$	748.3	2449.6	1.06	787.2	2450.8	1.07
	3.46	0	56.5	3.8	1.3	15.1	3.9	1.4	18.0

Legend

- ω is how many times the transition kernel is used.
- θ is the transition kernel variance.
- $\kappa_{\max}$ is the maximum threshold number. When κ reaches this value, whatever the threshold value, the estimated probability is $\mathbb{P}[\Delta \leq d_c] \approx p^{\kappa} \times p_{\kappa}^{\bullet}$.
- Results are presented as follows: $\begin{matrix} \mu\left(T, m\right) & \mu\left(N, m\right) \cdot 10^{-3} & \mu\left(\widetilde{Y}, m\right) \cdot 10^{6} \\ \rho\left(T, m\right) & \rho\left(N, m\right) & \rho\left(\widetilde{Y}, m\right) \end{matrix}$
- ρ is expressed in percentages.

7. Conclusion

So as to estimate probability collision at reasonable cost without relying on restrictive hypothesis, we searched the rare event probability estimation literature, as Crude Monte Carlo could not deliver. According to our experiments, the Adaptive Splitting Technique (AST) is a better choice with this respect than Cross–Entropy or Non Parametric Importance Sampling, two other rare event dedicated techniques. We advocate AST to this purpose.

The Adaptive Splitting Technique, however, requires tuning. We provided experimental insight about how to do it: a shaking parameter close to one but not equal to one and many transition kernel applications. These empirical results remain to be backed up with theoretical results.

Nonetheless, for it can be conveniently used with any black–box mapping and requires not hypothesis about it, we reckon AST is a worthy choice when it comes to estimating rare event probability or extreme quantile in an industrial context.

References

[1] I. Beichl and F. Sullivan. The importance of importance sampling. *Computing in science and engineering*, Approximating the Permanent via Importance Sampling with Application to the Dimer Covering Problem:pp.71–73, Mars/Avril 1999.

[2] P H Borcherds. Importance sampling: an illustrative introduction. *Computing in science and engineering*, pages 405–411, Avril 2000.

[3] Z. I Botev, D.P. Kroese, and T. Taimre. Generalized cross-entropy methods with applications to rare-event simulation and optimization. *Simulation*, 11:785 – 806, 2007.

[4] Frédéric Cérou, Pierre Del Moral, Teddy Furon, and Arnaud Guyader. Rare event simulation for a static distribution. Technical report, `http://hal.inria.fr/inria-00350762/fr/`, 2009.

[5] Frédéric CEROU and Arnaud GUYADER. Adaptive particle techniques and rare event estimation. *ESAIM:Proceedings*, 19:65–72, 2007.

[6] Frédéric Cérou and Arnaud Guyader. Adaptive multilevel splitting for rare event analysis. *Stochastic Analysis and Applications*, 25(2):417–443, 2007. Available on line.

[7] F.Kenneth Chan. *Spacecraft Collision Probability*. The Aerospace Corporation & AIAA, 2008.

[8] T. Homem de Mello and R.Y. Rubinstein. *Rare event estimation for static models via cross-entropy and importance sampling*. John Wiley, New York, 2002.

[9] Mark Denny. Introduction to importance sampling in rare-event simulations. *European Journal of Physics*, pages 403–411, Juillet 2001. mark.denny@baesystems.com.

[10] Arnaud Guyader, Nicolas W. Hengartner, and Eric Matzner-Løber. Iterative Monte Carlo for extreme quantiles and extreme probabilities. Technical report, March 2010. Paper and slideshow and presentation available on line.

[11] R. Pastel J. Morio and F. le Gland. An overview of importance splitting for rare event simulation. *European Journal of Physics*, 31:1295–1303, 2010.

[12] T.S. Kelso. Analysis of the Iridium 33-Cosmos 2251 collision. Technical report, 2009. Available on line.

[13] Martin J. Klein. Principle of detailed balance. *Phys. Rev.*, 97(6):1446–1447, Mar 1955.

[14] Agnés Lagnoux. Rare event simulation. *PEIS*, 20(1):45–66, January 2006. Available on line.

[15] Pierre L'Écuyer, François Le Gland, Pascal Lezaud, and Bruno Tuffin. Splitting methods. In Gerardo Rubino and Bruno Tuffin, editors, *Monte Carlo Methods for Rare Event Analysis*, chapter 3, pages 39–61. John Wiley & Sons, Chichester, 2009.

[16] François LEGLAND. *Filtrage bayésien optimal et approximation particulaire*. Les Presses de l'Ecole Nationale Supérieure de Techniques Avancées, 2008. Available on line in French.

[17] Nicholas Zwiep Miura. *Comparison and design of Simplified General Perturbation Models (SGP4) and code for NASA Johnson Space Center, ORBITAL DEBRIS PROGRAM OFFICE*. PhD thesis, Faculty of California Polytechnic State University,, Available Online, May 2009.

[18] J. Morio, R. Pastel, and F. Le Gland. Estimation de probabilités et de quantiles rares pour la caractérisation d une zone de retombée d un engin. *Journal de la Société Française de Statistique*, To be published after revision.

[19] Jan C. Neddermeyer. Computationally efficient nonparametric importance sampling. *Journal of the American Statistical Association*, 104(486):788–802, 2009.

[20] Rudy Pastel, Jérôme Morio, and François Le Gland. Estimating satellite versus debris collision probabilities via the adaptive splitting technique. In *Proceedings of the 3rd International Conference on Computer modeling and simulation, Mumbai,India*, 2011.

[21] R.Y. Rubinstein and D.P. Kroese. *The Cross-Entropy Method: A Unified Approach to Monte Carlo Simulation*, volume Randomized Optimization and Machine Learning. Springer Verlag, 2004.

[22] Luke TIERNEY. Markov chains for exploring posterior distributions. *Annals of Statistics*, 22:1701–1728, December 1994. Available on line.

[23] Ping Zhang. Nonparametric importance sampling. *Journal of the American Statistical Association*, 91(434):1245–1253, September 1996.