Source: "http://www.redorbit.com/media/uploads/2013/01/shutterstock_105231248.jpg"

Estimation of impact probability of asteroids and space debris through Monte Carlo Line Sampling and Subset Simulation

KePASSA2017 – ESTEC/ESA

Matteo Romano, Matteo Losacco, Camilla Colombo, Pierluigi Di Lizia

Contents

- Problem of space debris and asteroids
  - Common approaches and alternative proposals
- Advanced Monte Carlo techniques
  - Line Sampling
  - Subset Simulation
- Application of the proposed methods
  - Definition of test cases
  - Results, comparisons and comments
- Conclusions and future developments

1. Common approach to the problem of space impacts
2. Proposed advanced methods

INTRODUCTION

Introduction

Common approach to the problem of space impacts

- Impacts of asteroids and space debris with planets threaten the safety of human activity in space
  - **Planetary protection** sets special requirements to avoid the contamination of celestial bodies due to man-made debris in interplanetary missions
  - Time periods under study generally span up to 100 years

- Estimating the impact probability for space objects with celestial bodies requires a **large number of long-term orbital propagations** with standard Monte Carlo Simulations, resulting in high computational cost
  - The amount of simulations required to estimate the probability within a given confidence level increases as the expected probability decreases

Introduction

Proposed advanced methods

More efficient sampling methods may increase the precision of the probability estimate, or reduce the amount of propagations and the computational cost:

- The **Line Sampling** method probes the impact region of the uncertainty domain by using lines instead of random points
  - The impact probability is estimated via analytical integration, resulting in a more accurate solution

- The **Subset Simulation** method computes the impact probability as the product of larger conditional probabilities
  - The method progressively identifies intermediate conditional levels moving towards the impact event, reducing the overall number of samples

1. Line Sampling
2. Subset simulation

ADVANCED MONTE CARLO METHODS

Advanced Monte Carlo methods

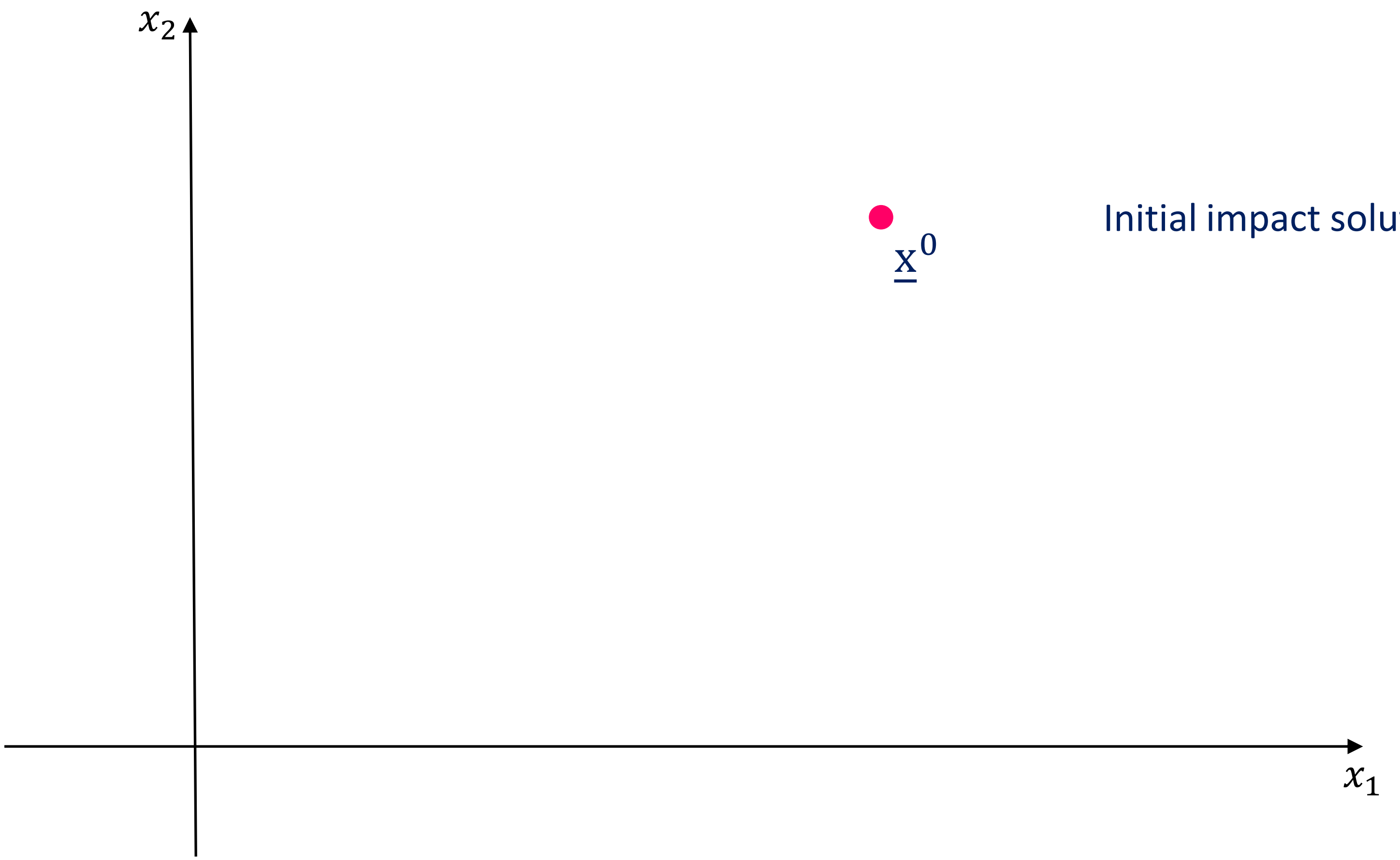
Line Sampling

The method follows 4 phases:

1. **Determination of the "reference direction"**
   Starting from a reference solution, a Markov Chain is used to find a direction pointing toward the impact region of the domain
2. **Mapping onto the standard normal space**
   Each sample is mapped from the physical coordinates to normalized ones, in order to associate a normal distribution to each of them
3. **Line Sampling**
   For each sample, a line following the reference direction is probed to identify the limits of the impact region
4. **Estimation of impact probability**
   Probability is analytically estimated as the average of integrals of unit normal distribution obtained along each line

Advanced Monte Carlo methods

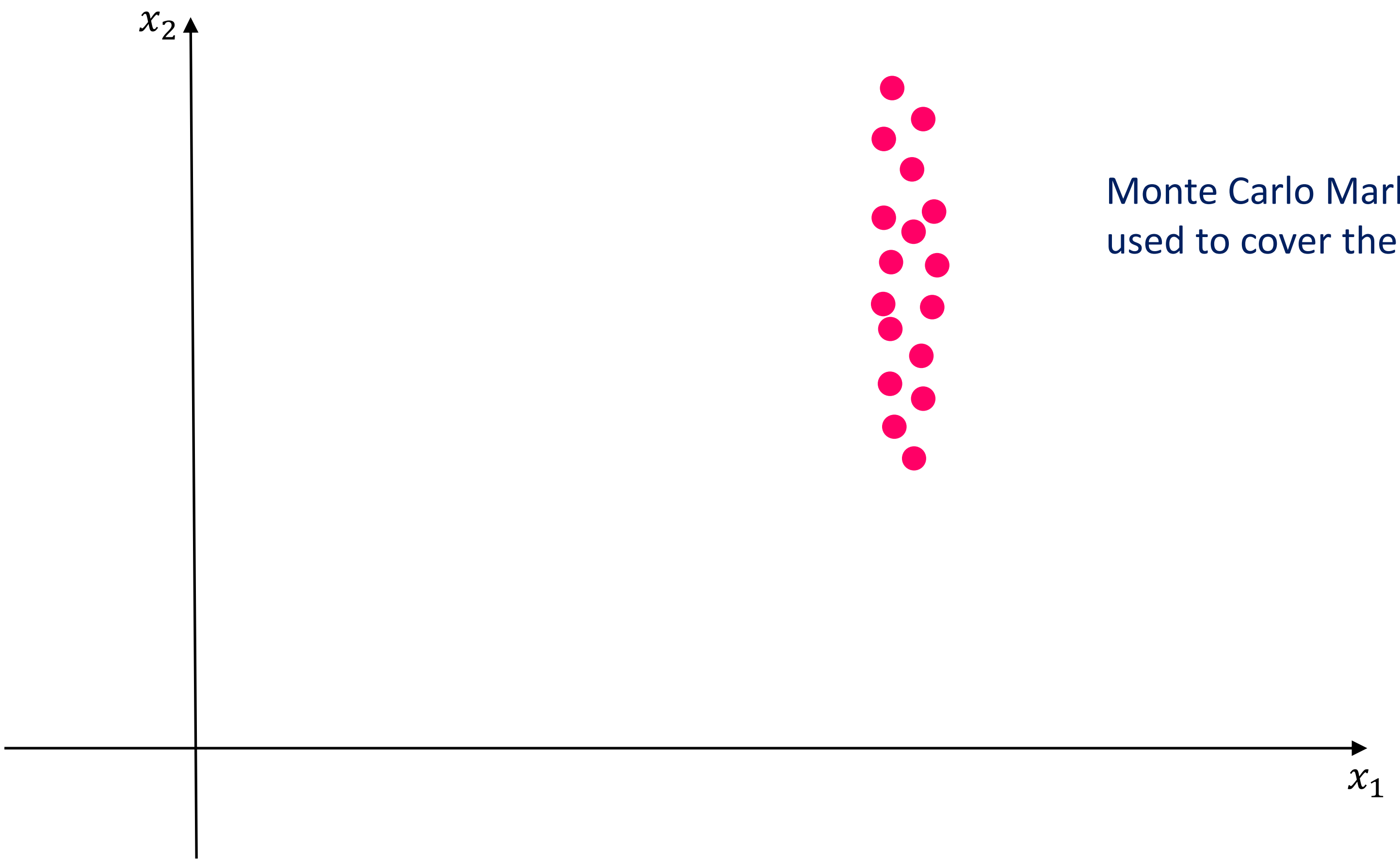
Line Sampling



Initial impact solution is found

Advanced Monte Carlo methods

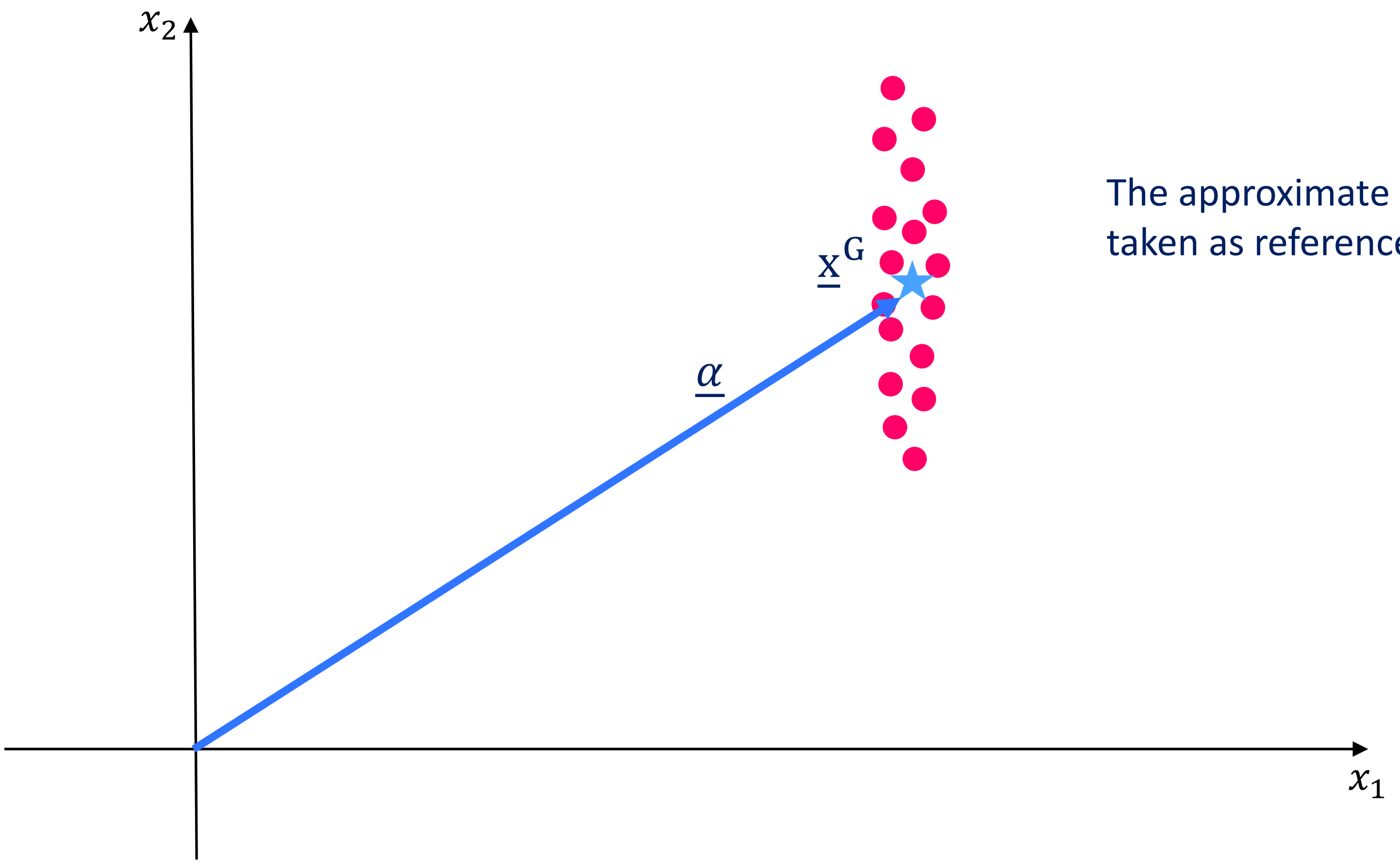
Line Sampling



Monte Carlo Markov Chain is used to cover the impact region

Advanced Monte Carlo methods

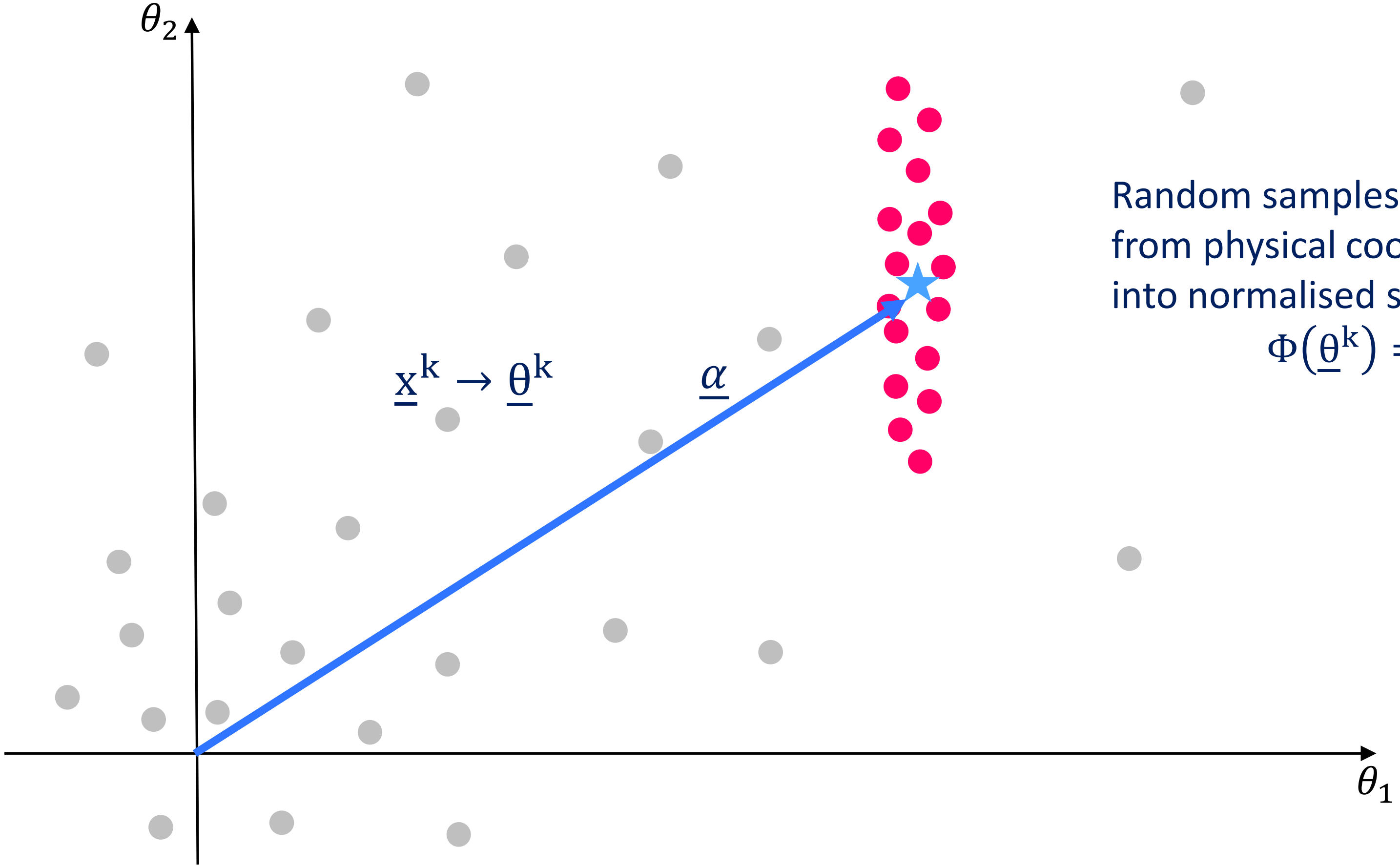
Line Sampling



The approximate barycentre is taken as reference direction

Advanced Monte Carlo methods

Line Sampling

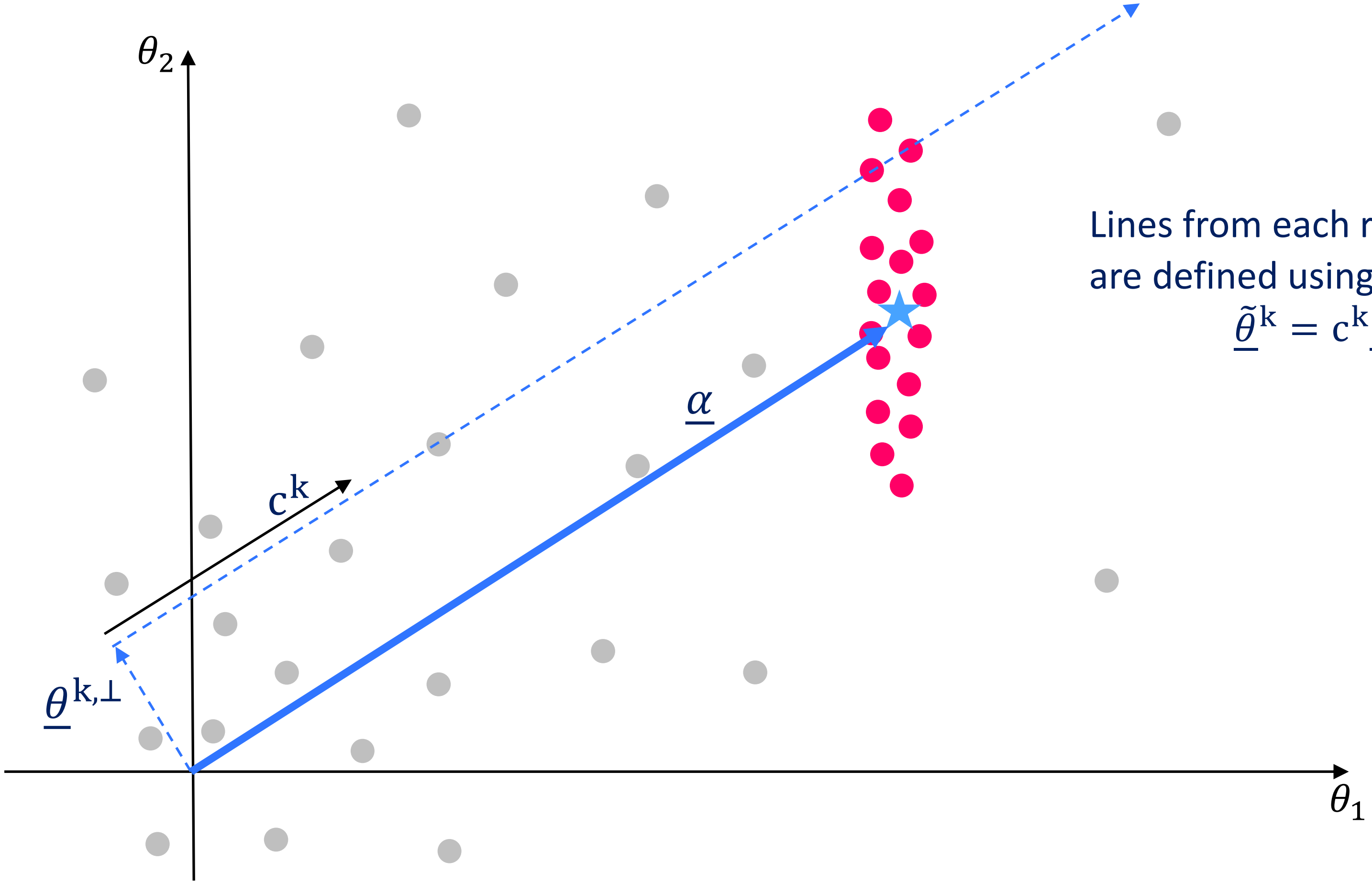


Random samples are transformed from physical coordinate space into normalised standard space

$$\Phi(\underline{\theta}^k) = F(\underline{x}^k)$$

Advanced Monte Carlo methods

Line Sampling

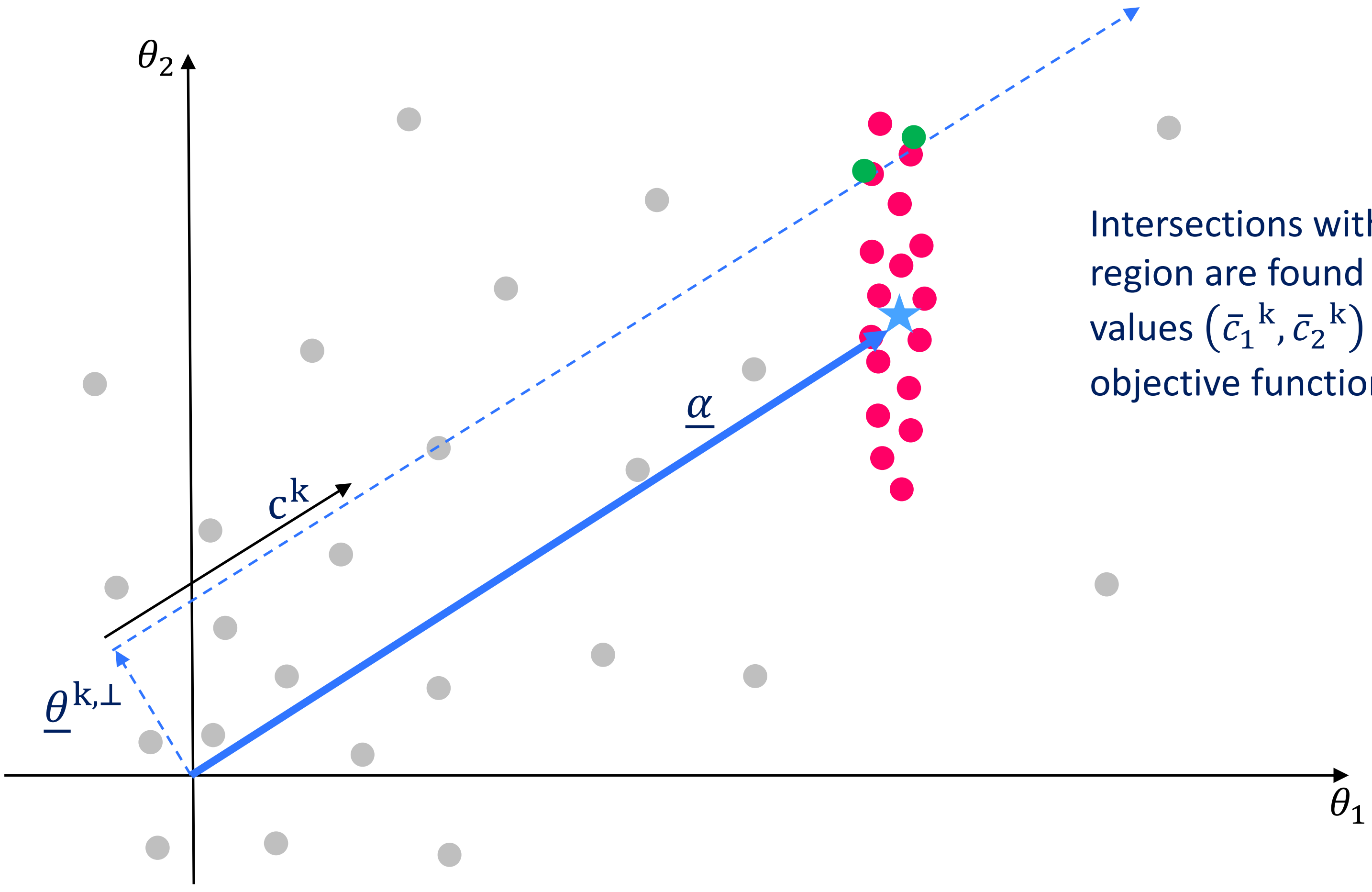


Lines from each random sample are defined using parameter c^k

$$\underline{\tilde{\theta}}^k = c^k \underline{\alpha} + \underline{\theta}^{k,\perp}$$

Advanced Monte Carlo methods

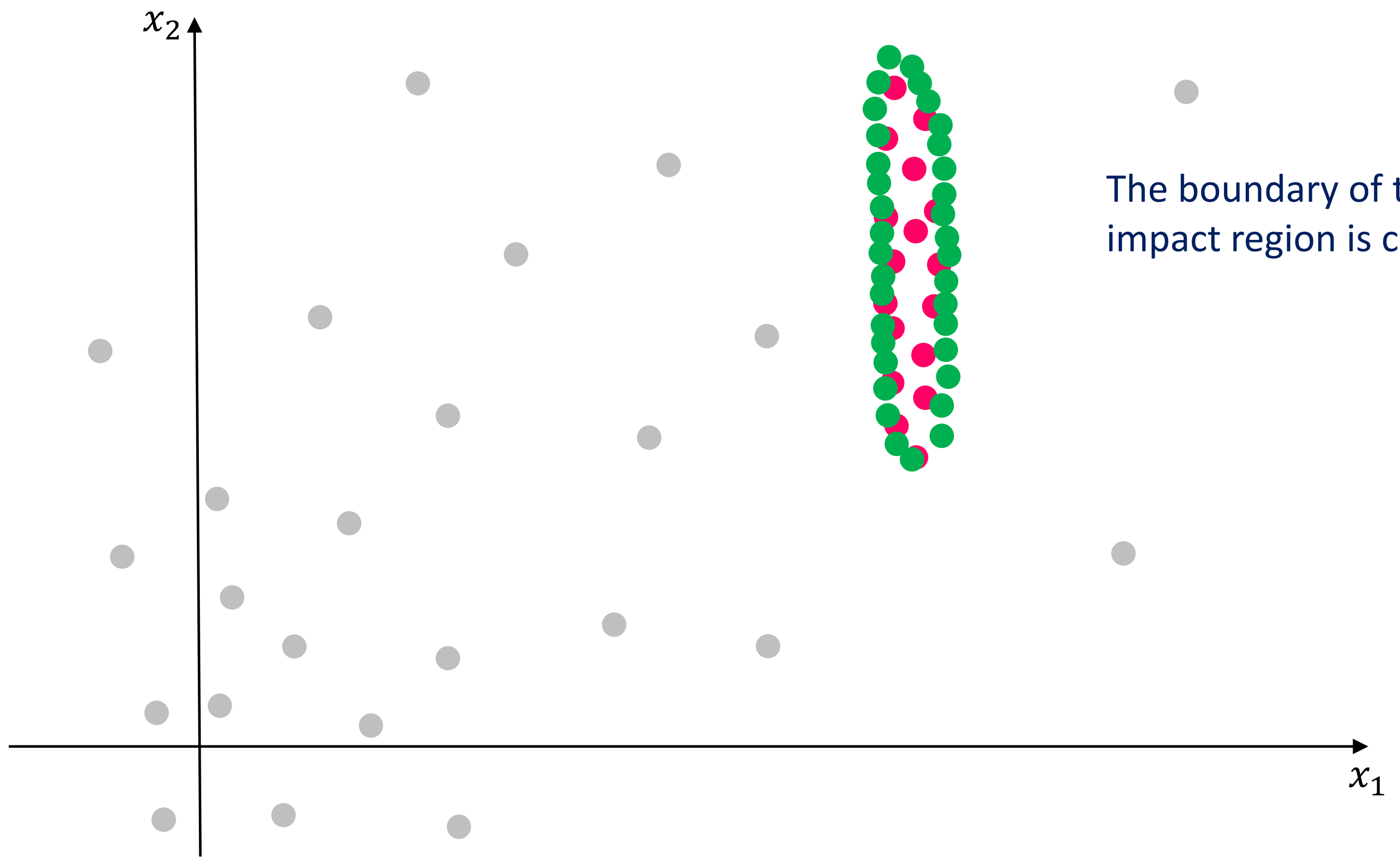
Line Sampling



Intersections with impact region are found with limits values $(\bar{c}_1{}^{k}, \bar{c}_2{}^{k})$ where an objective function $Y(c^k) = 0$

Advanced Monte Carlo methods

Line Sampling



The boundary of the impact region is covered

Advanced Monte Carlo methods

Line Sampling

- The limit values are used to compute one-dimensional impact probability estimates $\widehat{P}^k(I)$ along each line as
$$\widehat{P}^k(I) = \Phi\left({\bar{c}_2}^k\right) - \Phi\left({\bar{c}_1}^k\right)$$
where Φ represents the CDF of the unit normal distribution
- The total probability and the associated variance are estimated as
$$\widehat{P}(I) = \frac{1}{N_T}\sum_{k=1}^{N_T}\widehat{P}^k(I)$$
$$\hat{\sigma}^2\left(\widehat{P}(I)\right) = \frac{1}{N_T(N_T-1)}\sum_{k=1}^{N_T}\left(\widehat{P}^k(I) - \widehat{P}(I)\right)^2$$

Advanced Monte Carlo methods

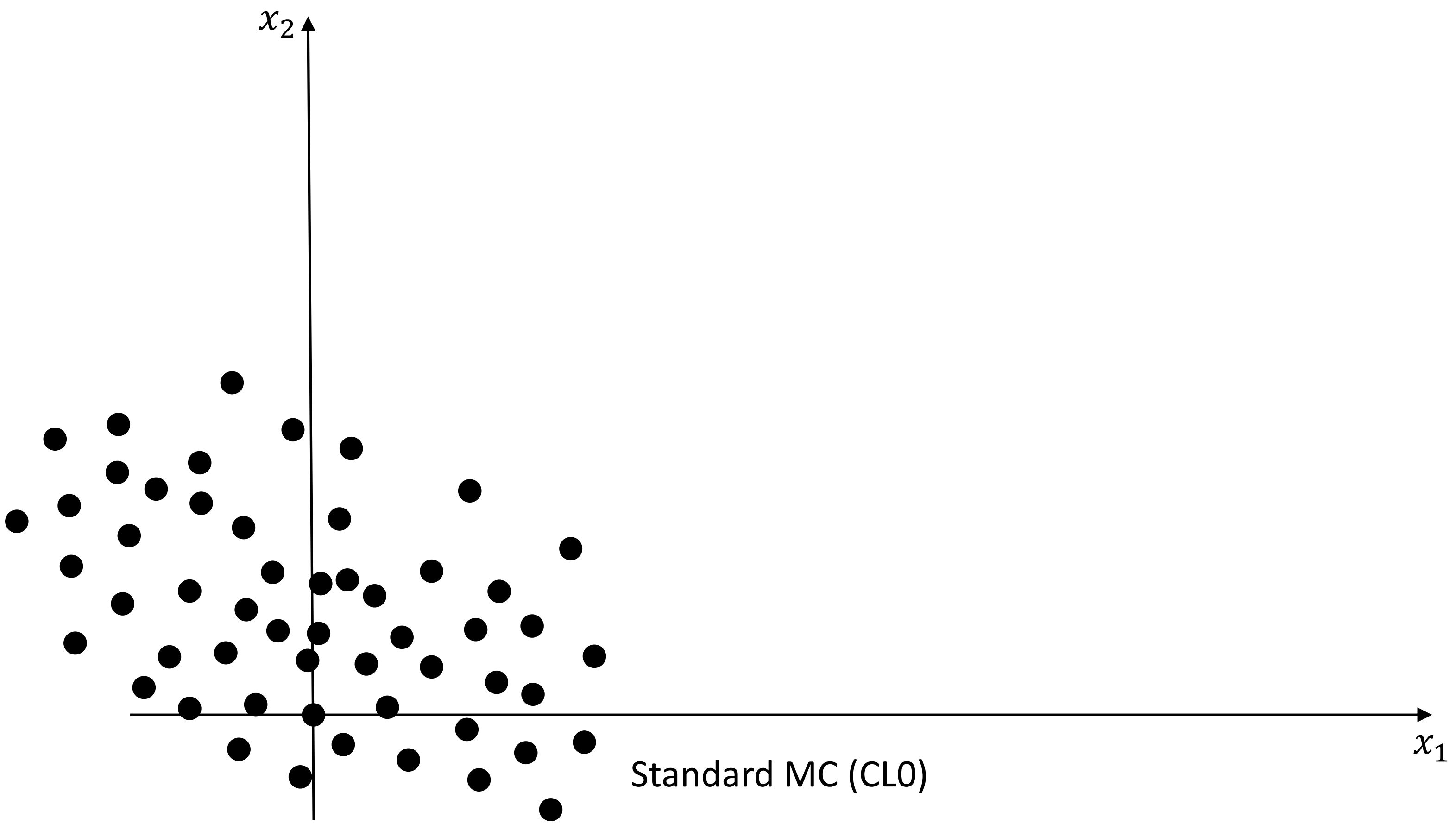
Subset Simulation

The method uses a series of MC Markov Chains in order to cover the initial uncertainty domain and move toward the impact region:

1. Generate N samples at conditional level (CL) 0 by standard Monte Carlo
2. Propagate each sample and identify its minimum distance from the Earth, and sort them according to their distance from Earth
3. Identify samples belonging to the next conditional level
   - Consider fixed values of conditional probability and identify the threshold as the (1-p_0)Nth element of the list
4. Generate (1- $p_0 N$) new samples belonging to that CL by means of Monte Carlo Markov Chain
5. Go back to step 2

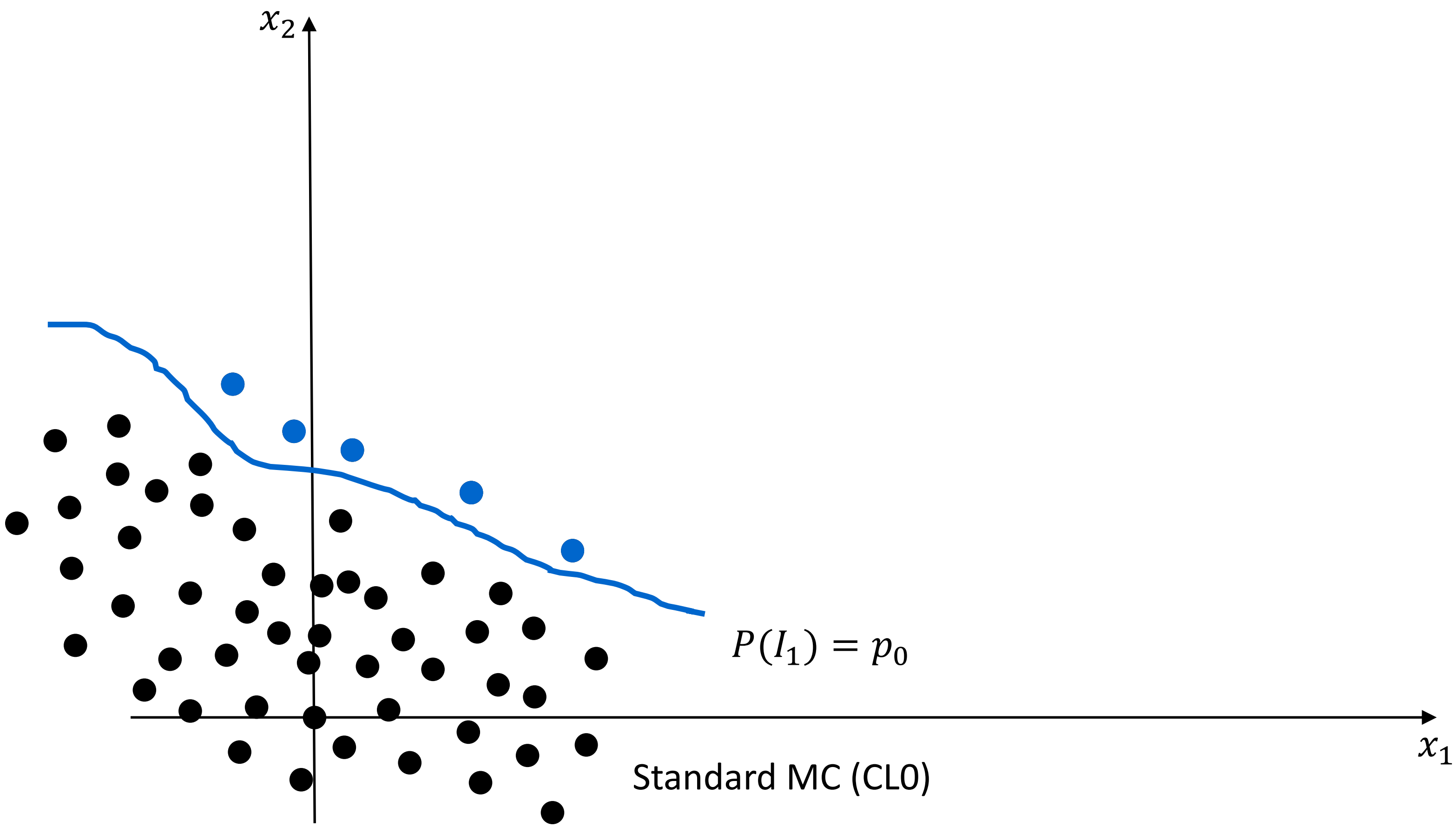
Advanced Monte Carlo methods

Subset Simulation



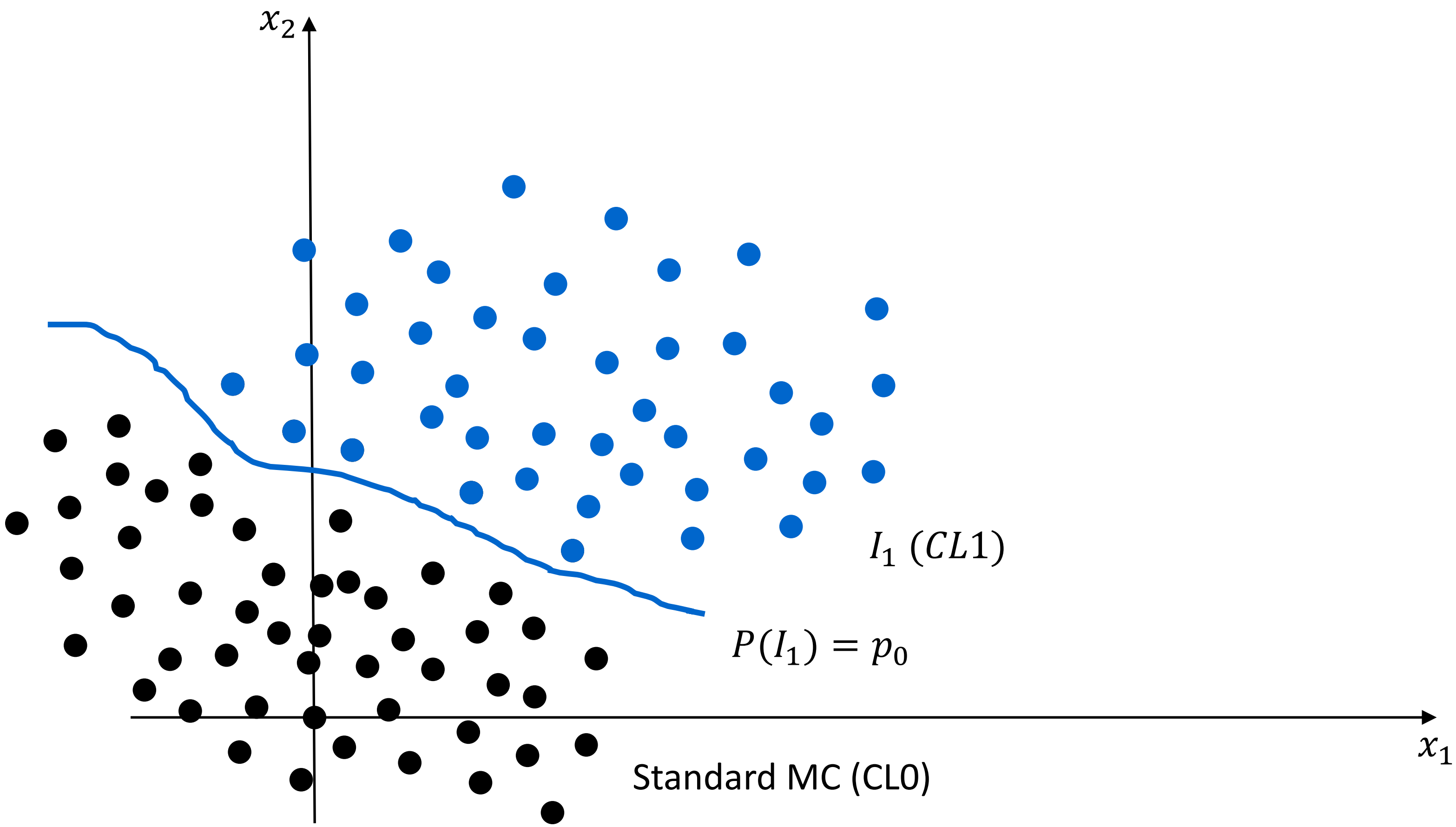
Advanced Monte Carlo methods

Subset Simulation



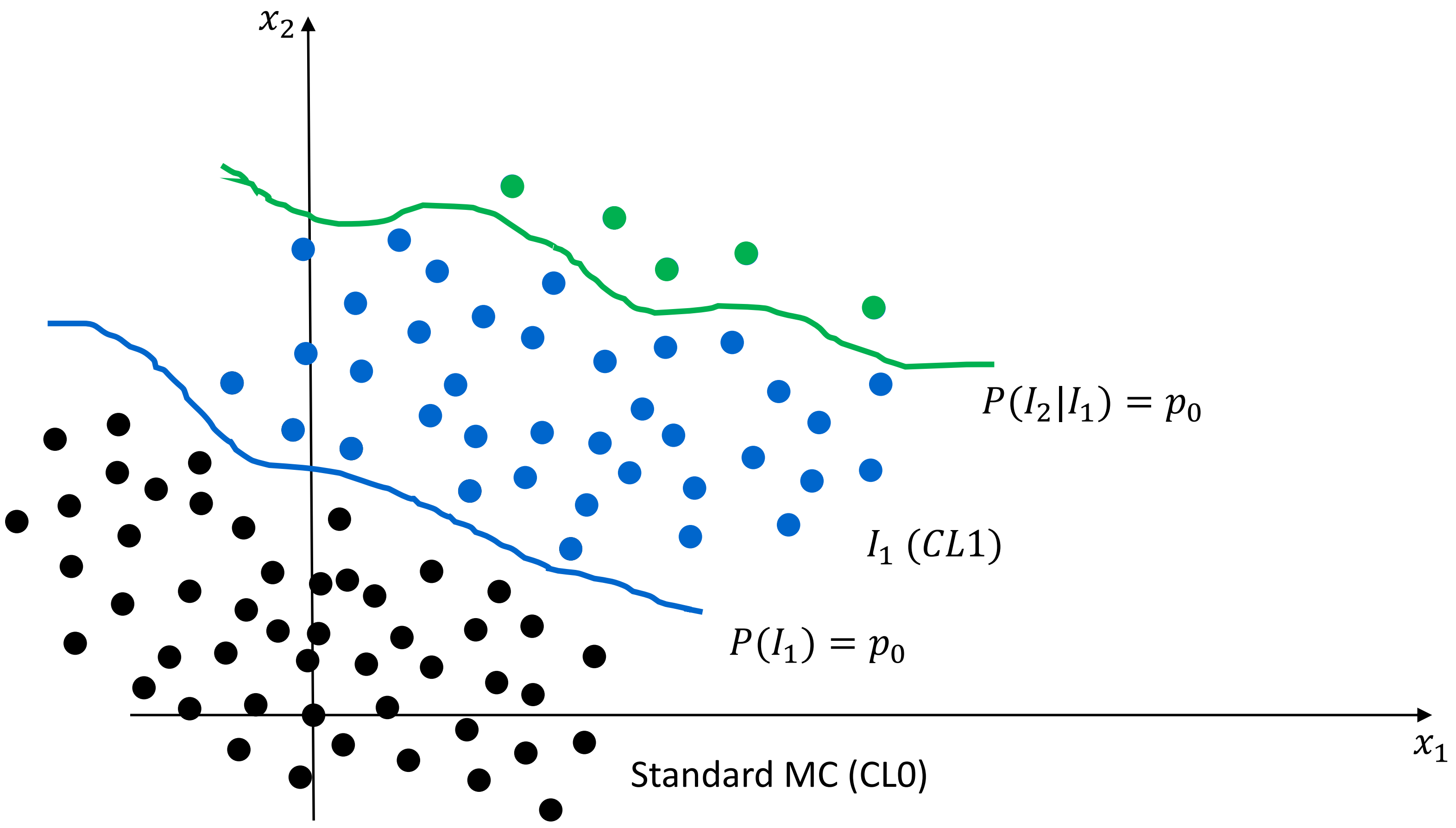
Advanced Monte Carlo methods

Subset Simulation



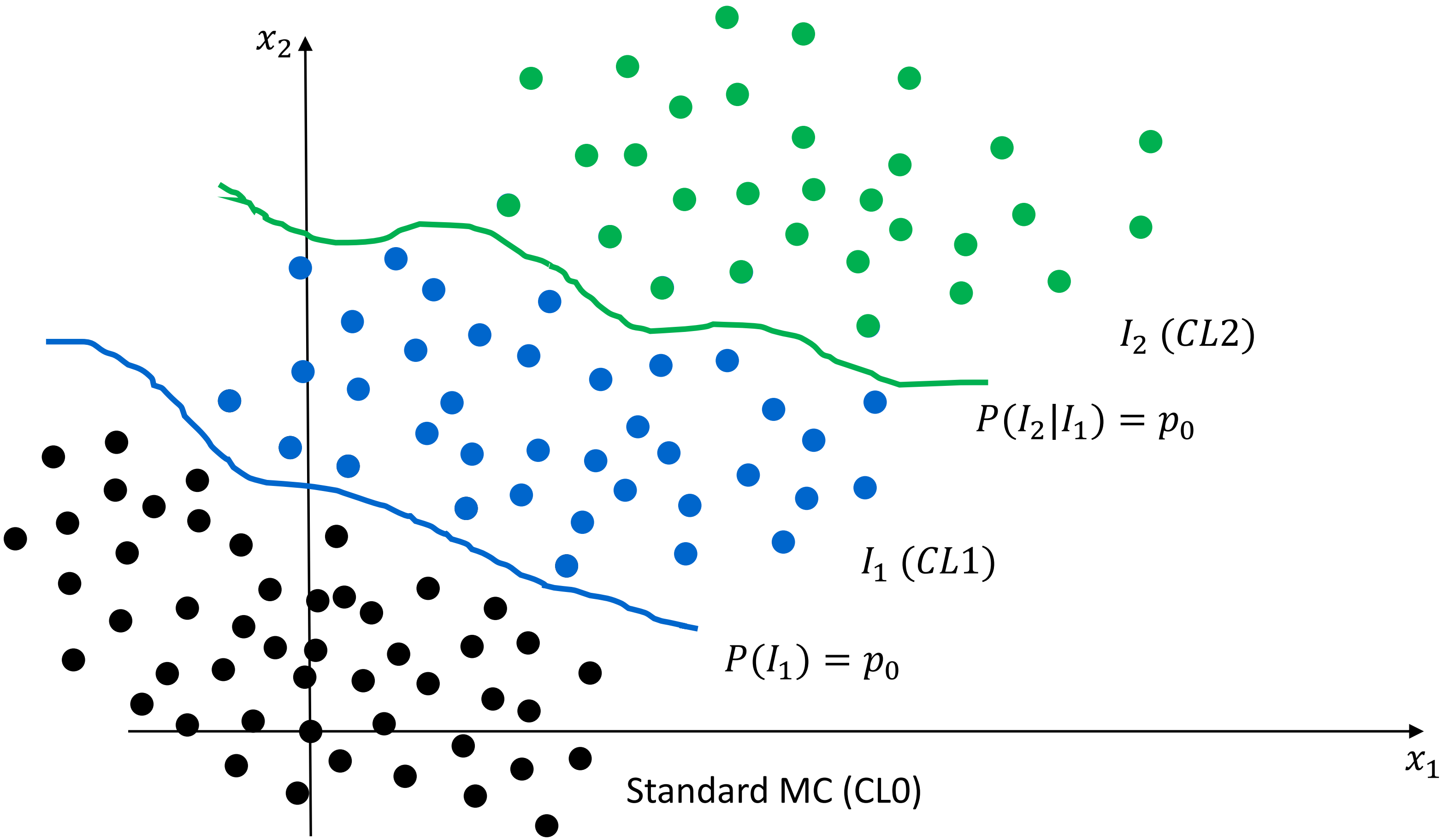
Advanced Monte Carlo methods

Subset Simulation



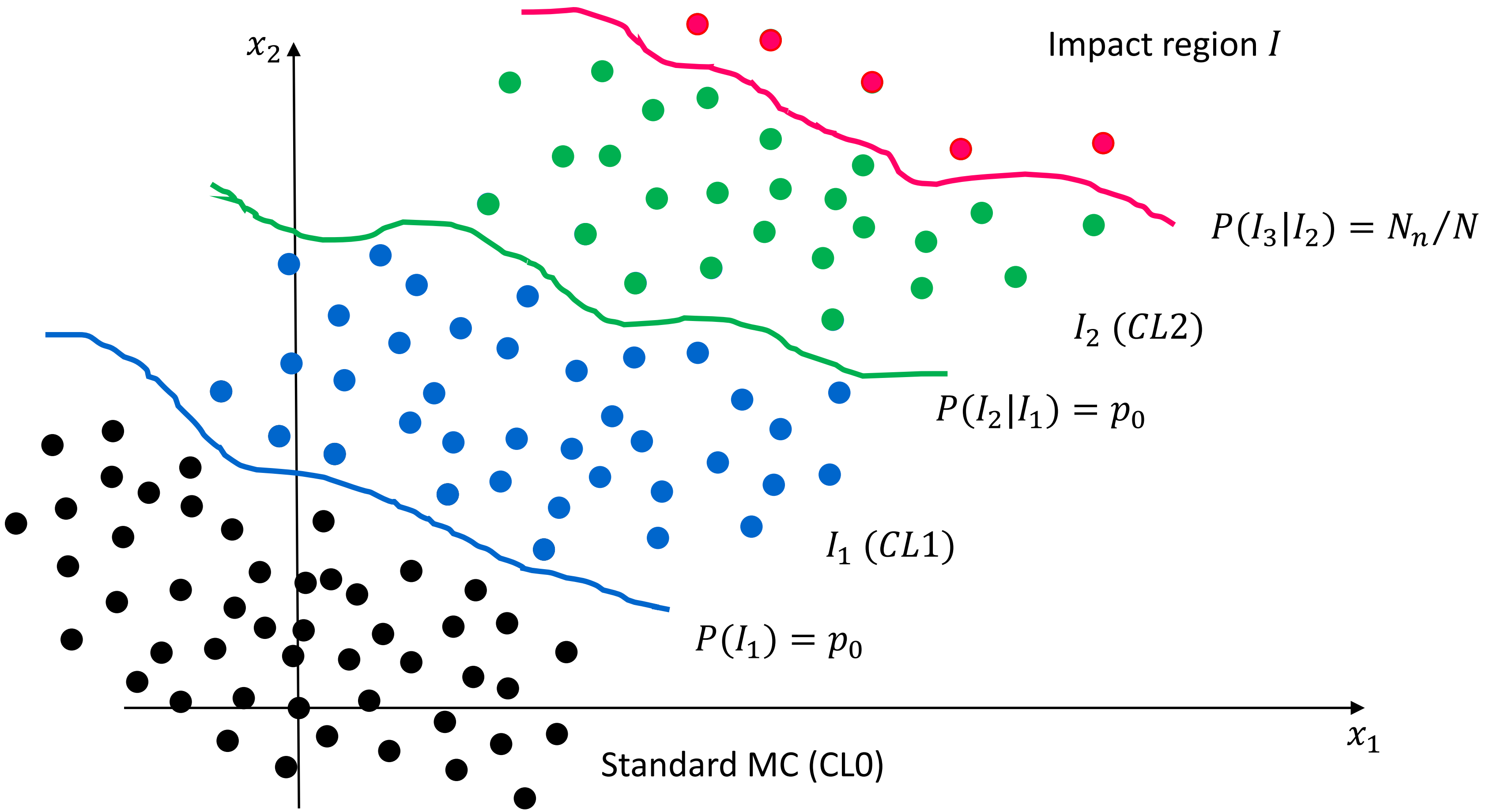
Advanced Monte Carlo methods

Subset Simulation



Advanced Monte Carlo methods

Subset Simulation



Advanced Monte Carlo methods

Subset Simulation

- Procedure is stopped when the threshold for the current conditional level is lower than the minimum distance set for the impact condition
- Given a sequence of intermediate impact regions $I_1 \supset \ldots \supset I_n = I$, the impact probability is estimated as the product of larger conditional probabilities

$$\hat{P}(I) = \hat{P}(I_1) \prod_{i=1}^{n-1} \hat{P}(I_{i+1}|I_i) = {p_0}^{n-1} N_n / N$$

  - $\hat{P}(I_{i+1}|I_i)$ – probability of I_{i+1} conditional to I_i
  - p_0 – predefined conditional probability (equal for every level)
  - n – number of conditional levels
  - N – number of samples per conditional level
  - N_n – number of impacting samples at the last conditional level

1. Case 1: definition, results and comparison
2. Case 2: definition, results and comparison

APPLICATION OF THE PROPOSED METHODS

Application of the proposed methods

Premise

The comparison among the different techniques is performed by analysing the following parameters:

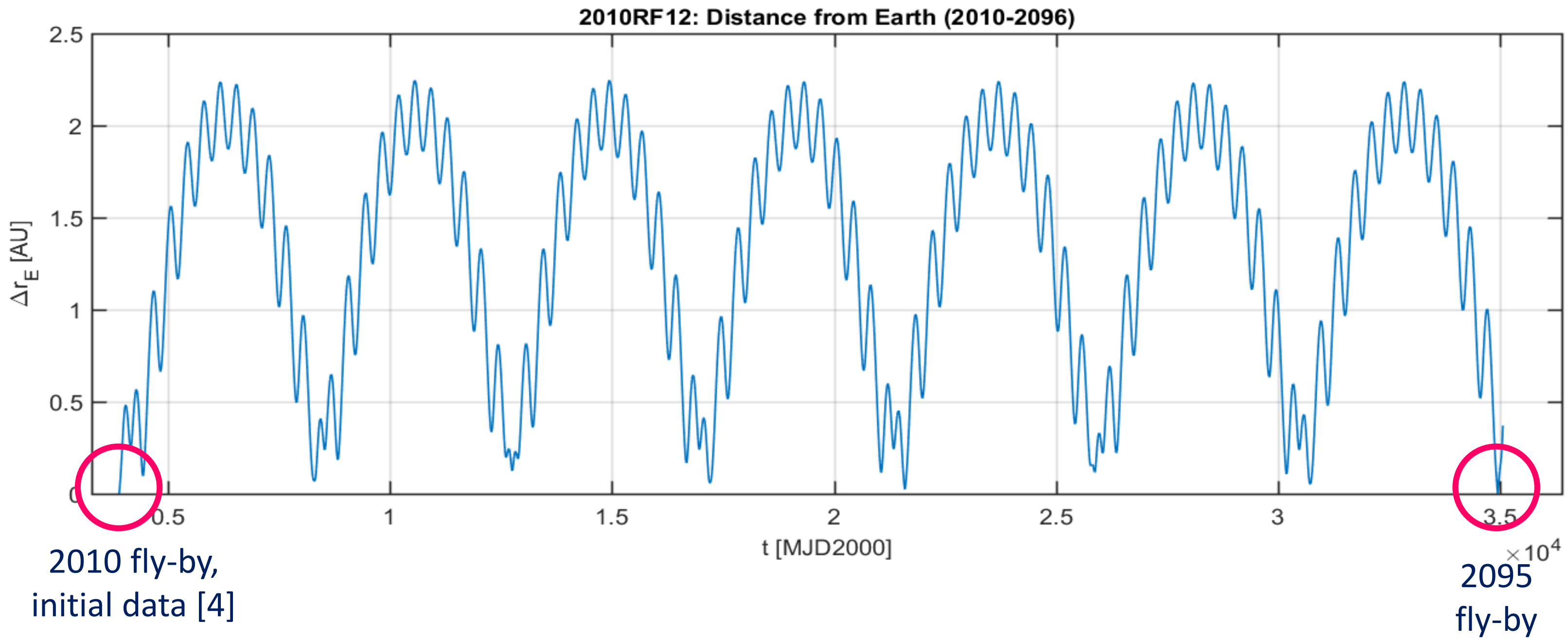
- Number of initial random samples N_S
- Total number of propagations N_P
- Impact probability estimate $\hat{P}(I)$
- Associated standard deviation $\hat{\sigma}\left(\hat{P}(I)\right)$
- Coefficient of variation $\delta = \hat{\sigma}/\hat{P}(I)$
- Figure of Merit $FOM = 1/(\hat{\sigma}^2 \cdot N_P)$

Notes:

- The computational time was not considered due to the nature of the numerical integration and of the machine, and the number of simulations was preferred instead as reference parameter
- The FOM parameter puts more emphasis on the accuracy of the results than on the computational burden [1]

Application of the proposed methods

Case 1: 2010RF12 – Definition

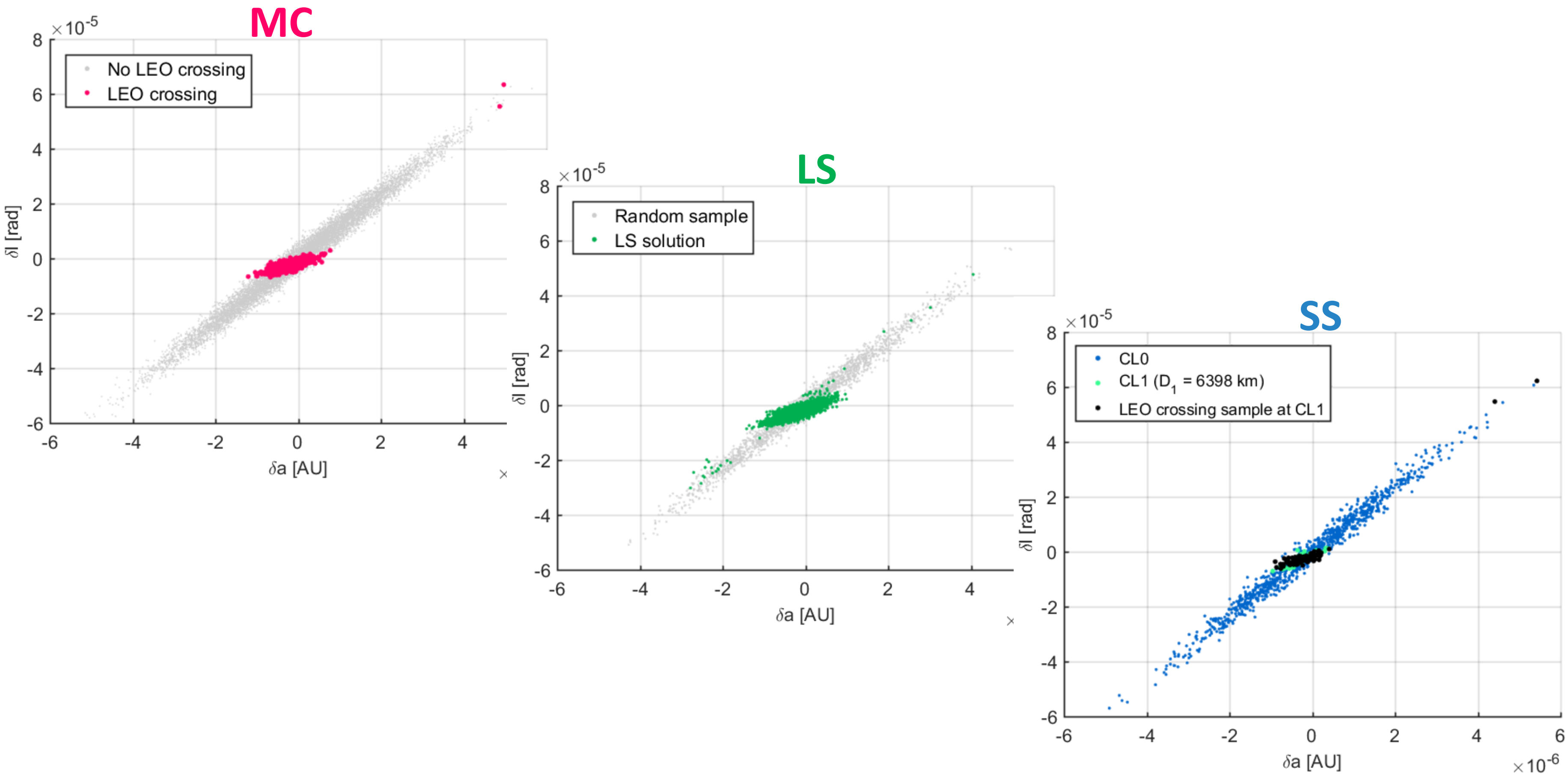


- 2010 RF_{12} is a small Near Earth Asteroid (NEA) with currently **the highest probability of hitting the Earth** (around 6%)
- 2095 fly-by was chosen as a test case due to the high expected impact probability
- **LEO crossing** (with LEO under 2000 km from Earth's surface) was considered as main event for the application of the methods

Application of the proposed methods

Case 1: 2010RF12 – Comparison of results

Visualisation of initial uncertainty set and impact region in $(\delta a, \delta l)$ plane



Application of the proposed methods

Case 1: 2010RF12 – Comparison of results

	N_S	N_P	$\hat{P}(I)$	$\hat{\sigma}$	δ	FOM	$\hat{\sigma}_{MC}/\hat{\sigma}$	FOM/FOM_{MC}
Std MCS	1e4	1e4	7.07e-2	2.56e-3	2.76e1	15.22	1.00e0	1.00e0
LS	1e3	~6e3	7.29e-2	1.13e-3	1.56e-2	122.17	2.26e0	8.03e0
SS	1e3	1.9e3	8.01e-2	6.00e-3	7.50e-2	14.62	4.27e-1	9.61e-1

- Both methods correctly identify the impact region
- For high values of the expected probability, the two methods are comparable with the standard Monte Carlo in terms of accuracy

Application of the proposed methods

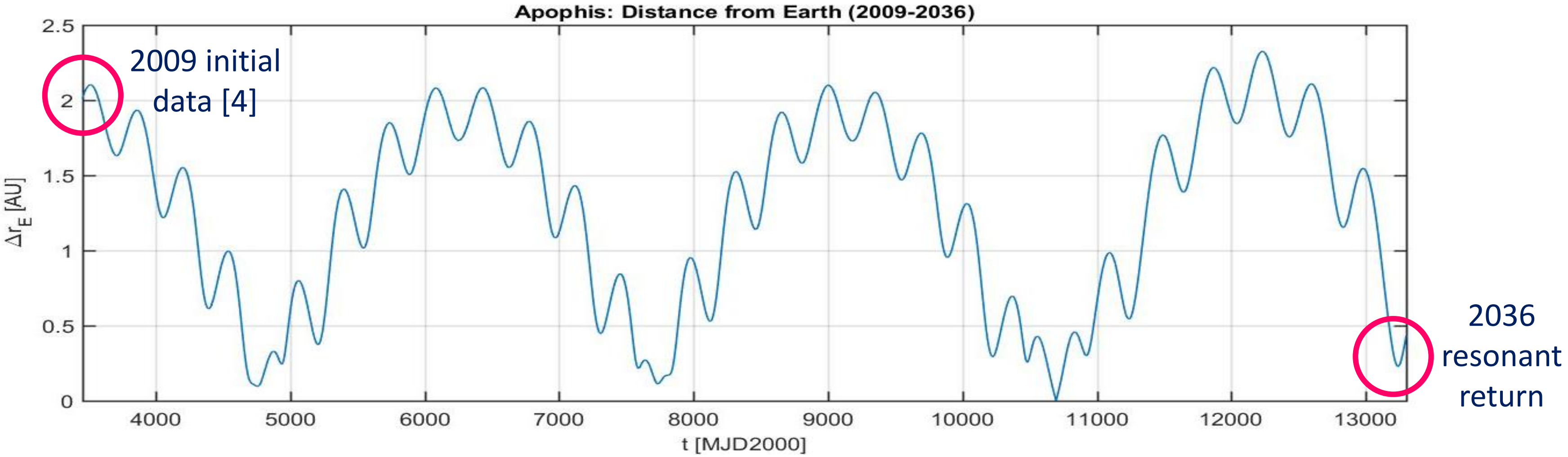
Case 1: 2010RF12 – Comparison of results

	N_S	N_P	$\widehat{P}(I)$	$\widehat{\sigma}$	δ	FOM	$\widehat{\sigma}_{MC}/\widehat{\sigma}$	FOM/FOM_{MC}
Std MCS	1e4	1e4	7.07e-2	2.56e-3	2.76e1	15.22	1.00e0	1.00e0
LS	1e3	~6e3	7.29e-2	1.13e-3	1.56e-2	122.17	2.26e0	8.03e0
SS	1e3	1.9e3	8.01e-2	6.00e-3	7.50e-2	14.62	4.27e-1	9.61e-1

- Both methods correctly identify the impact region
- For high values of the expected probability, the two methods are comparable with the standard Monte Carlo in terms of accuracy

Application of the proposed methods

Case 2: Apophis – Definition

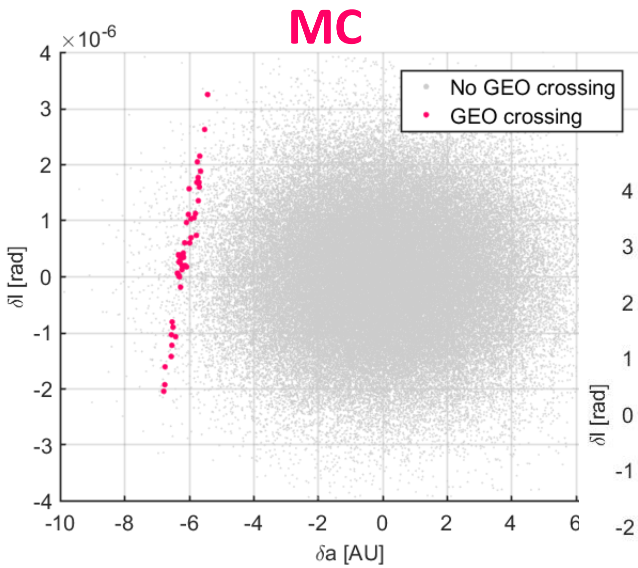


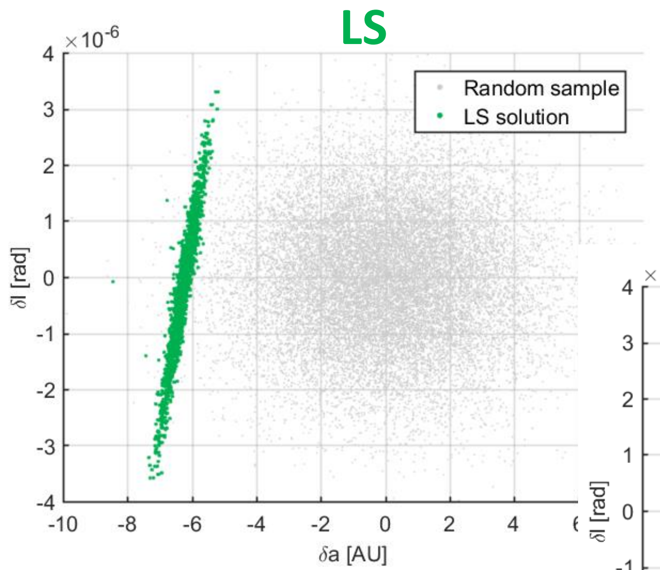
- 2036 close approach may present impacts with Earth depending on the uncertainty given by 2009 observations, and was chosen as a test case as more challenging for the application of the methods
- **GEO crossing** was considered as main event for the application of the methods for computational reasons

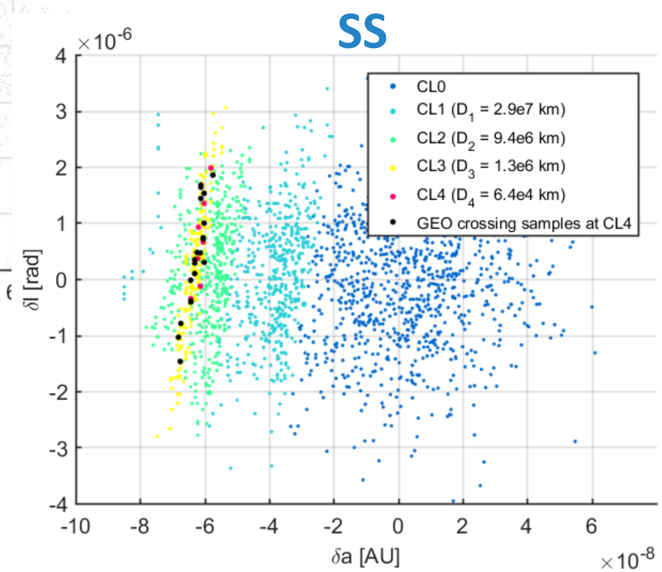
Application of the proposed methods

Case 2: Apophis – Comparison of results

Visualisation of initial uncertainty set and impact region in $(\delta a, \delta l)$ plane







Application of the proposed methods

Case 2: Apophis – Comparison of results

	N_S	N_P	$\widehat{P}(I)$	$\widehat{\sigma}$	δ	FOM	$\widehat{\sigma}_{MC}/\widehat{\sigma}$	FOM/FOM_{MC}
Std MCS	1e6	1e6	4.70e-5	6.86e-6	6.85e0	2.13e4	1.00e0	1.00e0
LS	1e5	~1e6	5.32e-5	3.45e-7	6.48e-3	8.46e6	1.99e1	3.97e2
SS	1e3	4.6e3	6.10e-5	1.30e-5	2.10e-1	1.28e6	5.27e-1	6.05e1

Both methods become more convenient as the expected probability decreases, outperforming standard MC either in computational time (SS) or accuracy (LS)

- SS provides an estimate of the probability using a much **lower number of propagations**, with significant reductions in the computational burden, but also a larger variance
- LS is computationally heavier in its current implementation, but offers the **highest accuracy**, especially when the estimated probability is low

Application of the proposed methods

Case 2: Apophis – Comparison of results

	N_S	N_P	$\hat{P}(I)$	$\hat{\sigma}$	δ	FOM	$\hat{\sigma}_{MC}/\hat{\sigma}$	FOM/FOM_{MC}
Std MCS	1e6	1e6	4.70e-5	6.86e-6	6.85e0	2.13e4	1.00e0	1.00e0
LS	1e5	~1e6	5.32e-5	3.45e-7	6.48e-3	8.46e6	1.99e1	3.97e2
SS	1e3	4.6e3	6.10e-5	1.30e-5	2.10e-1	1.28e6	5.27e-1	6.05e1

Both methods become more convenient as the expected probability decreases, outperforming standard MC either in computational time (SS) or accuracy (LS)

- SS provides an estimate of the probability using a much **lower number of propagations**, with significant reductions in the computational burden, but also a larger variance
- LS is computationally heavier in its current implementation, but offers the **highest accuracy**, especially when the estimated probability is low

1. Final considerations
2. Future developments

CONCLUSIONS

Conclusions

Final considerations

- Line Sampling
  - Advantages
    - Using the same number of random samples, LS can achieve a lower variance of the solution (**higher accuracy**) thanks to the use of exact integrals on each line
    - Larger efficiency with respect to standard techniques as the impact probability gets lower
  - Current limitations
    - Current implementation supposes a unique impact region with a regular shape
    - Current implementation uses **extra evaluations** to probe each line, thus decreasing the efficiency of LS

Conclusions

Final considerations

- Subset Simulation
  - Advantages
    - It drives the generation of samples towards the impact region of the initial uncertainty set
    - Larger efficiency with respect to standard techniques as the impact probability gets lower
    - **Reduced computational burden** with respect to standard MC and LS
  - Current limitations
    - Correlation among conditional samples: possible bias in the estimates
    - **Larger variance** with respect to LS method

Conclusions

Future developments

- Line Sampling
  - Analysis of the influence of the main implementation degrees of freedom on the accuracy of the results
  - Obtain an analytical expression for the confidence interval in order to determine a minimum number of lines required for a given confidence level
  - Improvement of the zeros computation
  - Application to **planetary protection** studies as part of a Ph.D. research

- Subset Simulation
  - Analysis of the influence of the main implementation degrees of freedom on the accuracy of the results
  - Optimization of the Monte Carlo Markov Chain for samples generation

References

1. Enrico Zio, Nicola Pedroni, *Subset Simulation and Line Sampling for Advanced Monte Carlo Reliability Analysis*, Proceedings of the European Safety and RELiability (ESREL) 2009 Conference, 2009, pp.687-694. <hal-00721047>
2. Murray Rosenblatt, *Remarks on a Multivariate Transformation*, The Annals of Mathematical Statistics, Vol. 23, No. 3 (Sep., 1952), pp. 470-472
3. Zuev K. et al., *Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions*, Computers and Structures,92, 283-296, 2011
4. Initial conditions and uncertainty values taken from Near Earth Object Dynamic Site NEODyS (http://newton.dm.unipi.it/neodys)

THANK YOU FOR YOUR ATTENTION

This project 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)

Source: "http://www.redorbit.com/media/uploads/2013/01/shutterstock_105231248.jpg"

POLITECNICO MILANO 1863

Estimation of impact probability of asteroids and space debris through Monte Carlo Line Sampling and Subset Simulation

KePASSA2017 – ESTEC/ESA

Matteo Romano, Matteo Losacco, Camilla Colombo, Pierluigi Di Lizia