Implementation of the AK-MCS algorithm for structural reliability

by
Juan Camilo Lugo Rojas

Advisor
Dr. techn. Diego Andrés Alvarez Marín

Submitted in partial fulfillment of the requirements
for the degree of

Civil Engineer

Department of Civil Engineering
Faculty of Engineering and Architecture
National University of Colombia at Manizales

September, 2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The human being's desire of self-improvement leads him to try to do things better over time. Among its many creations, buildings and tools have played a fundamental role in the development of mankind. However, this development is constrained by the availability of resources (e.g. natural, economic), so their efficient use is becoming increasingly important. It is at this point where a deeper knowledge and interpretation of physical phenomena can significantly improve the design and construction of the aforementioned structures.

Deterministic models usually require a greater use of resources to deal with the uncertainty present in structural systems. This may not have as much impact when dealing with simple structures, but when it comes to the more and more common highly complex ones, the cost overruns can be quite considerable. In order to avoid this situation, a probabilistic approach is a more suitable option [Choi et al., 2006].

When using the probabilistic approach, reliability is understood as the probability that a structure will not fail, referring as failure to the structure not performing as desired. One of the most important technique to solve structural reliability problems is the Monte Carlo Simulation (MCS). The concept behind MCS is relatively simple, and it is a very powerful method, but its main drawback is that it can become computationally expensive [Nowak and Collins, 2000].

Several methods have been derived from MCS, trying to overcome the problem of its computational cost, while preserving its strengths. In this document we study one of those methods, known as AK-MCS for Active learning reliability method combining Kriging and MCS, which attempts to obtain results equivalent to those given by MCS in a more efficient manner [Echard et al., 2011].

1.1 Motivation

The fundamental reason for doing this project is the desire to be initiated in academic research, which contributes to the generation of knowledge and the development of science and the self.

Additionally, the structural reliability assesment is a very fascinating field which is constantly developing and can have a great impact. I can visualize myself working on these topics in the future.

1.2 Objective

The main objective of this final project is to implement the algorithm described in an article[1] that proposes a new approach based on MCS and Kriging metamodel to asses structural reliability. In order to satisfactorily do this, many specific objectives have to be achieved, which include:

- Learning the Monte Carlo Simulation method.
- Learning the Kriging regression algorithm.
- Reading and completely understand the article.
- Implementing the previously mentioned methods in Python.
- Reviewing articles that aim to improve the performance of the AK-MCS algorithm.
- Evaluating the performance of the algorithm, comparing it with the basic MCS.

[1] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(145-154), 2011. DOI: 10.1016/j.strusafe.2011.01.002

Chapter 2

Basic concepts of Structural Reliability

Human understanding of the laws of nature allows him to try to model the physical phenomena occurring around him. This knowledge will never be complete, but it can be attempted to be taken as in-depth as desired. How rigorous one wants to be becomes a trade-off between the benefits it brings and the difficulty and costs involved.

Regarding structures, it is of interest to study certain situations in which undesirable conditions take place. These are called limit state violations. The study of structural reliability concerns the calculation and prediction of the probability of a structure reaching such states [Melchers and Beck, 2018].

In accordance with the previously mentioned, the estimation of structural reliability can be done from different approaches, considering to a greater or lesser extent the uncertainties inherent to structural problems.

2.1 Uncertainties in Reliability Assesment

Uncertainties are what make the study of reliability meaningful. During the service life of a structure, several variables are involved that introduce uncertainties. A basic classification can be made according to their origin: [Nowak and Collins, 2000]:

- **Natural causes:** the unpredictability of loads[1] and the material properties that affects the mechanical behavior.
- **Human causes:** every task performed by a human being has a certain amount of uncertainty. Decision making is affected by many factors such as the level of knowledge, the presence of distractions,

[1] such as earthquakes, water pressure, live load, etc

communication problems, among others.

Due to uncertainties, the loads and resistances of a structure can be considered as random variables, and as such, they are a function of three factors:

- **Physical variation:** the variation that is inherent in the considered quantity.
- **Statistical variation:** parameter estimation based on a limited sample size leads to uncertainty.
- **Model variation:** simplifications made, unknown correlation between variables, correlations and unknown boundary conditions.

2.2 Deterministic approach

The deterministic approach is the traditional one. In this approach, the uncertainties in the forces and resistances are covered by assuming that the variation of these forces and resistances have limits that are not exceeded. Given this simplicity, it is usually necessary to overestimate the magnitudes in order to provide greater safety. This is usual in design codes when considering safety factors in the limit state equations. It is widely used, especially in simple and common problems. However, this required safety margin implies the need to assume higher costs.

2.3 Partial probabilistic approach

The historical record of events that expose structures to large loads, such as earthquakes, floods, tornadoes, etc., shows that they have a certain frequency according to their magnitude. For instance, during one day there may be multiple low-scale earthquakes, but every few years there is one of considerable destructive power. The partially probabilistic approach handles these return periods when considering the maximum loads to which a structure will be subjected, since it is not cost-effective to design an element for an event that it will probably never have to face. As its name indicates, this approach considers some randomness of the variables involved, but continues to make many simplifications.

2.4 Probabilistic approach

As discussed above, there is some randomness associated with the variables that affect the loads and strengths of structures. Taking this into account, allows for a deeper analysis, and therefore, to know

in greater detail the behavior of the structure, which is always desirable.

By making use of statistics, it is possible to infer an approximate distribution that the values of the variables mentioned above follow. By considering this, it is possible to estimate the probability that the loads are greater than the resistances, being this the probability of failure of the structure. However, in most cases this calculation is very complex and it is not possible to solve it analytically.For this reason, it is necessary to use methods that provide an approximation of the probability of failure.

2.5 Monte Carlo Simulation

With knowledge of the distributions followed by the variables involved in the reliability problem, by means of numerical techniques it is possible to generate random samples of numbers following these distributions. For a given sample it is possible to estimate the average number of points in this space that cause the structure to fail. By the law of large numbers, this average gets closer and closer to the expected value of the distribution. The Monte Carlo simulation is based on this principle. Despite its simplicity, this method is very robust when dealing with structural reliability problems.

Chapter 3

The AK-MCS algorithm

Crude Monte Carlo Simluations are easy to implement and robust in dealing with reliability problems. However, they require the evaluation of the expensive performance function a considerable amount of times. AK-MCS[1], that and must be seen as a modification of the traditional MCS, aims to overcome this downside making use Gaussian Processes Regressions and a learning function to optimize the number of calls of the performance function on a Monte Carlo simulation.

[1] stands for Active learning method combining Kriging and Monte Carlo Simulation and is described in [Echard et al., 2011]

3.1 *Step-by-step of the method*

1. **Generation of a Monte Carlo population in the design space.** According to the involved random variables, this population named S of n_{MC} points is generated.

2. **Definition of the initial design of experiments (DoE).** The initial DoE consists of a random selection of N_1 points of S. It is preferred to be small, adding at each iteration only the point that improves the metamodel the most. We will be using a dozen points as suggested in [Echard et al., 2011].

3. **Computation of the Kriging model.** The Kriging regressor is trained with the performance function G evaluated on the initial DoE. A model with an squared-exponential kernel is used.

4. **Prediction by Kriging and estimation of the probability of failure.** Predictions of the performance function over the Monte Carlo population, $\widehat{G}(x_i)$ for $i = 1, ..., n_{MC}$ are made with the previously trained Kriging model. Then, the estimated probability of failure $\widehat{p_f}$ is obtained by calculating the ratio of the points $x_i \in S$ such that $\widehat{G}(x_i) \leq 0$, i.e.:

$$\widehat{p_f} = \frac{n_{\widehat{G}(x_i) \leq 0}}{n_{MC}} \tag{3.1}$$

5. **Identification of the best next point in S to be evaluated on the performance function.** At this stage, the learning function is computed for each point of S to determine the next point that should be added to the DoE to improve the most the metamodel.

6. **Evaluation of the stopping condition.** The stopping condition associated to the learning function is evaluated for the point selected in the previous stage. If the criterion is met, we skip to Stage 8. Otherwise, we continue with Stage 7.

7. **Update of the design of experiments.** The point identified at stage 5 is added to the current DoE, such that $N_{i+1} = N_i + 1$. Then, the method goes back to Stage 3.

8. **Computation of the coefficient of variation of the probability of failure** If the stopping condition is satisfied, the metamodel is said to be accurate enough on the sign of the performance function on S. Subsequently, it is checked if S is large enough to obtain a low coefficient of variation on the estimation of p_f. If the coefficient of variation is lesser than 5%, AK-MCS stops and the last estimation of the probability of failure is considered as the result. In other case, it continue to Stage 9.

9. **Update of the population.** S is increased with other n_{MC} points from the design space generated in the same way that in Stage 1.

3.2 *Learning Functions*

The learning functions are used to decide the next point of the population to be included in the DoE. In order to do that, they give a score to each x_i of S according to the value $\widehat{G}(x_i)$ and by its variance $\sigma^2_{\widehat{G}}(x_i)$ given by the Kriging regressor. Every learning function comes along with a learning criterion and with a stopping condition on the obtained scores.

3.2.1 *Expected feasibility function (EFF)*

It is defined as

$$\begin{aligned} EFF(x) = &\left(\widehat{G}(x) - a\right)\left(2\Phi(C) - \Phi(C^+) - \Phi(C^-)\right) \\ &-\sigma_{\widehat{G}}(x)\left(2\phi(C) - \phi(C^+) - \phi(C^-)\right) \\ &+\epsilon\left(\Phi(C^+) + \Phi(C^)-\right) \end{aligned} \tag{3.2}$$

where

$$C = \frac{a - \widehat{G}(x)}{\sigma_{\widehat{G}}(x)}$$
$$C^{+} = \frac{(a+\epsilon) - \widehat{G}(x)}{\sigma_{\widehat{G}}(x)}$$
$$C^{-} = \frac{(a-\epsilon) - \widehat{G}(x)}{\sigma_{\widehat{G}}(x)}$$

and Φ is the standard normal cumulative distribution and ϕ the standard normal density distribution.

Proposed in [Bichon et al., 2008]. It provides an indication of how well the true value of $\widehat{G}(x)$ is expected to satisfy the equality $G(x) = a$. In AK-MCS we have $a = 0$ and $\epsilon = 2\sigma_{\widehat{G}}(x)$. The learning criterion of $EFF(x)$ is the maximum value, so the best next point to add to DoE is $x^* \in S$ such that $EFF(x^*) = \max(EFF(x))$. The stopping condition is $EFF(x^*) \leq 0.001$.

3.2.2 *Learning function U*

$$U(x) = \frac{\left|\widehat{G}(x)\right|}{\sigma_{\widehat{G}}(x)} \tag{3.3}$$

Proposed in [Echard et al., 2011]. Considering that in MCS only the sign of the performance function is important, this function selects as the next best point of S to be added to the DoE the one that has the higher potential risk of crossing the separator $\widehat{G}(x) = 0$. Then, the best next point to add to the DoE is $x^* \in S$ such that $U(x^*) = \min(U(x))$. The stopping condition is $U(x^*) \geq 2$.

3.2.3 *Learning function H*

$$H(x) = \left| \ln\left(\sqrt{2\pi}\sigma_{\widehat{G}}(x) + \frac{1}{2}\right)\left[\Phi\left(\frac{D^-}{\sigma_{\widehat{G}}(x)}\right) - \Phi\left(\frac{-D^+}{\sigma_{\widehat{G}}(x)}\right)\right] - \left[\frac{D^-}{2}\phi\left(\frac{D^-}{\sigma_{\widehat{G}}(x)}\right) + \frac{D^+}{2}\phi\left(\frac{-D^+}{\sigma_{\widehat{G}}(x)}\right)\right]\right| \tag{3.4}$$

where

$$D^{+} = 2\sigma_{\widehat{G}}(x) + \widehat{G}(x)$$
$$D^{-} = 2\sigma_{\widehat{G}}(x) - \widehat{G}(x)$$

Proposed in [Lv et al., 2015]. This function is based on the information entrophy theory and it measures the uncertainty of $\widehat{G}(x)$. The best next point to add to the DoE is $x^* \in S$ such that $H(x^*) = \max(H(x))$. It will be used the same stopping condition that is chosen in [Lv et al., 2015], it is $H(x^*) \leq 0.5$.

3.3 *Examples*

In this section some examples are presented. They consist of the statement of the design space and the performance function. Results of each problem are presented in tables that contains the number of calls to the performance function N_{call}, the estimated probability of failure $\widehat{p_f}$, the coefficient of variation for the Monte Carlo method, and the relative error of each method compared to the probability obtained with the Monte Carlo method.

3.3.1 *Example 1*

The first example is a simple problem with just one variable, that follows a normal distribution with mean 0 and standard deviation 2, and the performance function considered is 3.5.

$$G(\boldsymbol{x}) = \sin x \tag{3.5}$$

The problem is solved only with the learning function U, and the metamodel is initialized with just 5 points in the initial DoE.

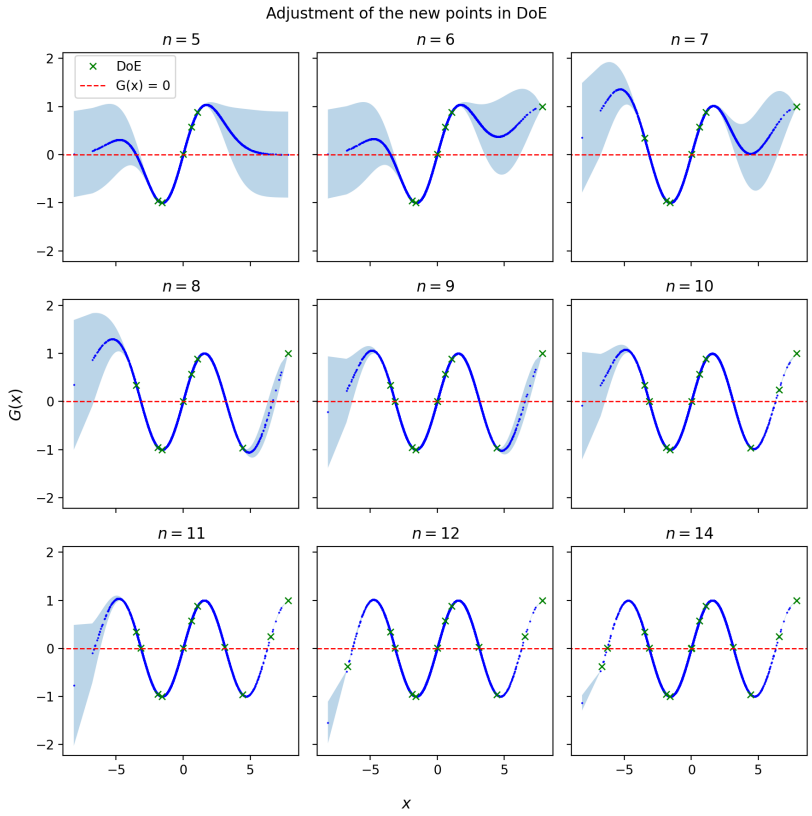

Figure 3.1: Example 1. Prediction and standard deviation obtained with n points in the DoE.

Method	N_{call}	$\widehat{p_f}$ (C.O.V$_{\widehat{p_f}}$)	$\epsilon_{\widehat{p_f}}(\%)$
Monte Carlo	1×10^4	$4.944 \times 10^{-1}(1.01\%)$	-
AK-MCS+U	14	4.944×10^{-1}	0

Table 3.1: Results of example 1

In figure 3.1 it can be seen how the algorithm selects the next point to be added to the DoE such that it reduces the most the variance near to the limit $G(\boldsymbol{x}) = 0$. Even though there is still some considerable variance in some regions, it stops at $N = 14$ because this uncertainty is away from the limit, so it is unlikely to affect the estimation of the probability of failure. In fact, the results in table 3.1 reveal that a very accurate estimation is obtained.

3.3.2 *Example 2*

In the second place there is a series system with four branches that is worked in [Echard et al., 2011]. Both random variables are standard normal distributed. The performance function is:

$$G(x_1, x_2) = \min \left\{ \begin{array}{c} 3 + 0.1(x_1 - x_2)^2 - \frac{(x_1+x_2)}{\sqrt{2}}; \\ 3 + 0.1(x_1 - x_2)^2 + \frac{(x_1+x_2)}{\sqrt{2}}; \\ (x_1 - x_2) + \frac{4}{\sqrt{2}}; \\ (x_2 - x_1) + \frac{4}{\sqrt{2}} \end{array} \right\} \tag{3.6}$$

with k taking two different values: 6 and 7.

From the results on tables 3.2 and 3.3 we can see that the learning function U obtained the most accurate estimations of $\widehat{p_f}$, being exactly the same of the MCS. In the case $k = 6$, the other two functions performed well, they only failed at clasifying one point. In the other one, with $k = 7$, EFF made some misclassifications, but the learning function H, although calling the performance function less times, had an error of 30%.

Method	N_{call}	$\widehat{p_f}$ (C.O.V$_{\widehat{p_f}}$)	$\epsilon_{\widehat{p_f}}(\%)$
Monte Carlo	1×10^6	$4.433 \times 10^{-3}(1.5\%)$	-
AK-MCS+U	126	4.433×10^{-3}	0
AK-MCS+EFF	123	4.432×10^{-3}	0.02
AK-MCS+H	113	4.434×10^{-3}	0.02

Table 3.2: Results of example 2 with $k = 6$

In the figure 3.2 the actual distribution of the two classes in the MC population is displayed, while in the figure 3.3 there is the distribution predicted by AK-MCS+U at several stages. Additionally, figure 3.3 give some insights about the update of the DoE, showing how the selected

Method	N_{call}	$\widehat{p_f}$ (C.O.V$_{\widehat{p_f}}$)	$\epsilon_{\widehat{p_f}}(\%)$
Monte Carlo	1×10^6	2.161×10^{-3}(2.15%)	-
AK-MCS+U	103	2.161×10^{-3}	0
AK-MCS+EFF	107	2.156×10^{-3}	0.23
AK-MCS+H	65	1.5×10^{-3}	30.59

Table 3.3: Results of example 2 with $k = 7$

points tend to come from near the limit $G(\boldsymbol{x}) = 0$.

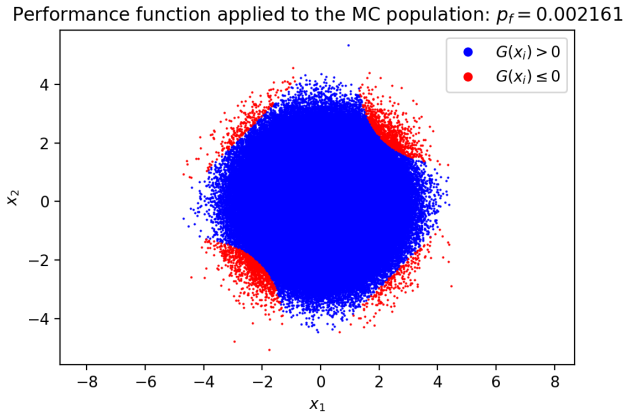

Figure 3.2: Example 2 with $k = 7$. Evaluation of the performance function on the MC population.

Figures 3.4 and 3.5 display how the probability of failure converges with each learning function, and how the selected points make the learning criteria tend to the stopping conditions. In both cases there are similar behaviors. The EFF is proned to converge more consistently to the stopping condition. The function U usually gets a good estimation well before finishing. The function H presents the most irregular results. A stricter stopping condition would be more appropiate for this last learning function.

3.3.3 *Example 3*

The next example is a non-linear undamped single degree of freedom system, as the one shown in figure 3.6. The involved variables are listed in table 3.4.

$$G(\boldsymbol{x}) = 3r - \left| \frac{2F_i}{m\omega_0^2} \sin\left(\frac{\omega_0 t_1}{2}\right) \right| \tag{3.7}$$

where $\omega_0 = \sqrt{\frac{c_1+c_2}{m}}$

Variable	P.D.F	Mean	Standard deviation
m	Normal	1	0.05
c_1	Normal	1	0.1
c_2	Normal	0.1	0.01
r	Normal	0.5	0.05
F_1	Normal	1	0.2
t_1	Normal	1	0.2

Table 3.4: Random variables of example 3

Table 3.5 contains the obtained results. In contrast with the other examples, in this case the function H performs quite well, obtaining the exact result, just as the function U, but with some more calls to the performance function.

In figure 3.7 it is evident that the function EFF met its stopping condition before converging to the solution, unlike the other two functions.

Method	N_{call}	$\widehat{p_f}$ (C.O.V$_{\widehat{p_f}}$)	$\epsilon_{\widehat{p_f}}(\%)$
Monte Carlo	7×10^4	2.7814×10^{-2}(2.23%)	-
AK-MCS+U	60	2.7814×10^{-2}	0
AK-MCS+EFF	43	2.7671×10^{-2}	0.51
AK-MCS+H	66	2.7814×10^{-2}	0

Table 3.5: Results of example 3

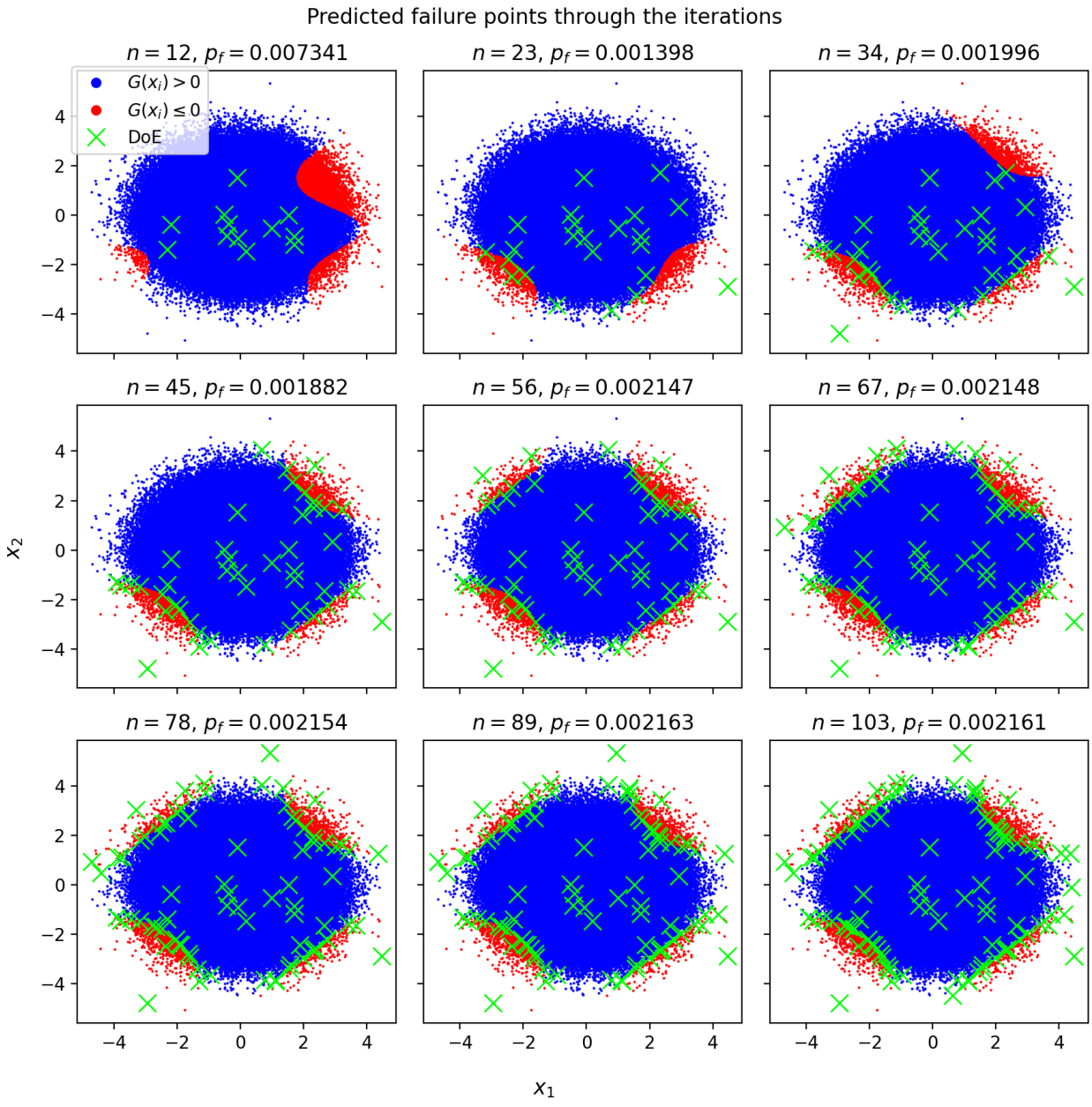

Figure 3.3: Example 2 with $k = 7$. Prediction made by AK-MCS+U at several stages.

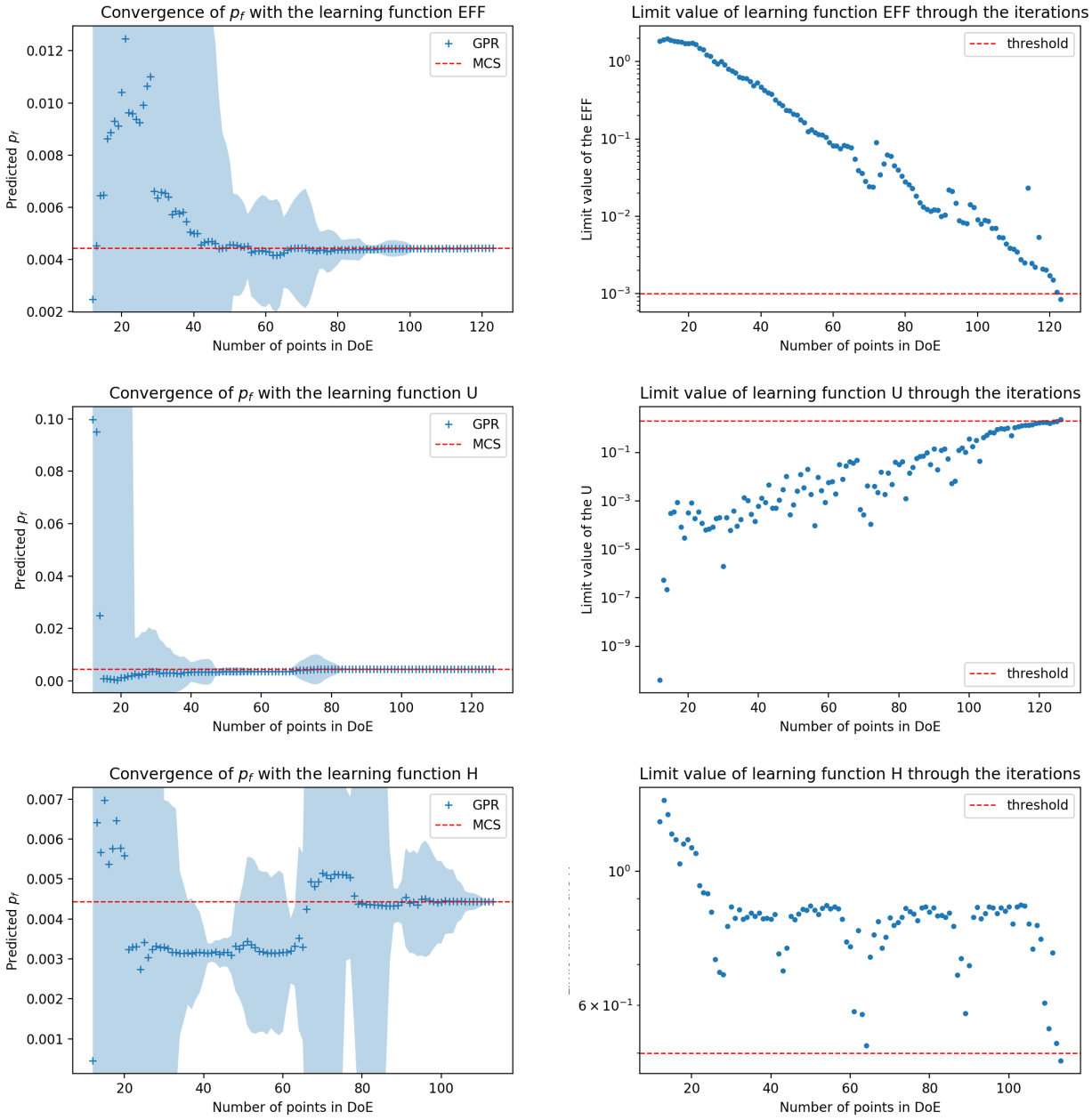

Figure 3.4: Results of example 2 with $k = 6$

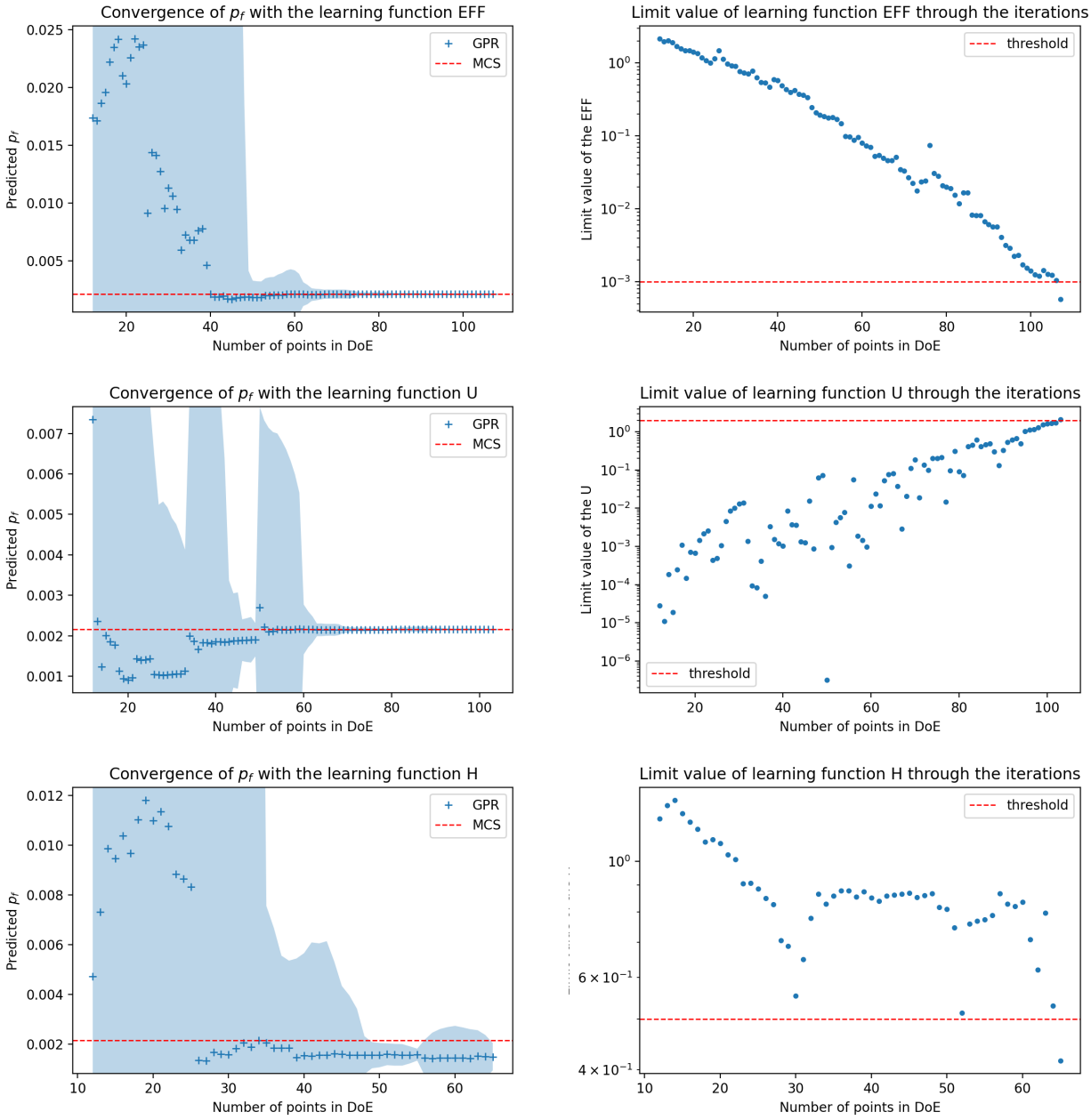

Figure 3.5: Results of example 2 with $k = 7$

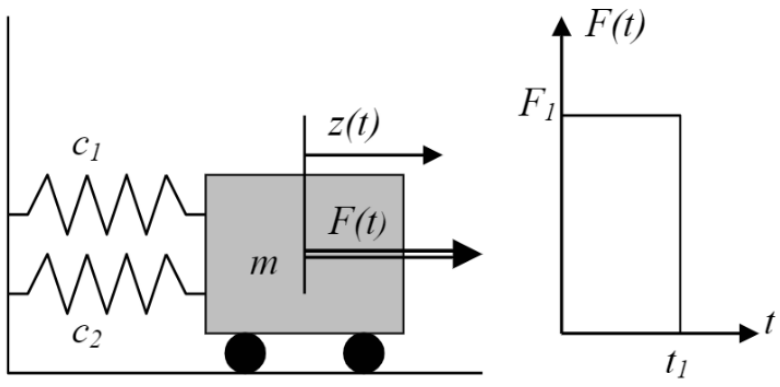

Figure 3.6: Example 3. Nonlinear oscilator. Taken from [Schueremans and Gemert, 2005]

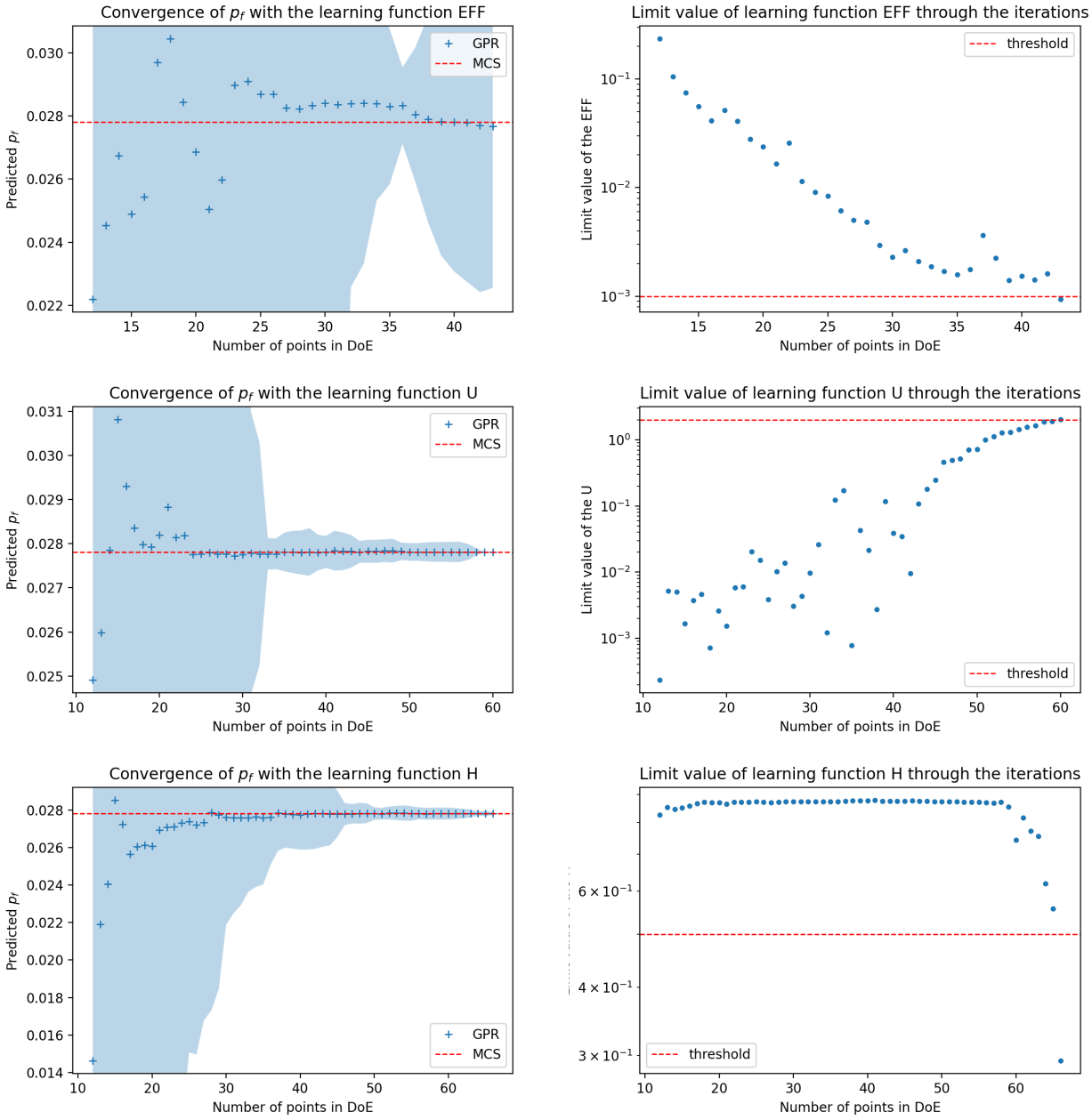

Figure 3.7: Results of example 3

Chapter 4

Conclusion

The AK-MCS algorithm proved to have several advantages, being a substantial improvement of the basic MCS method. It retains the advantages of the latter, improving its main disadvantage which is having to call the expensive performance function so many times.

The choice of one learning function over another should be influenced by the nature and form of the problem. Although they are all based on the same parameters, the manner in which they determine the evolution of the prediction process varies by giving greater importance to different aspects of the design space and the performance function. Likewise, the choice of thresholds for the stopping conditions is of great importance, as they represent the usual trade-off between cost and benefit. It could be seen that in the worked examples, the suggested threshold of the learning function H was too high, yet for some particular problems it was low enough.

Given the simplicity of the nature of the method, its potential for improvement is very high. In principle, the formulation of new learning functions is a boundless task. Moreover, different ways of approaching the problem following the same principles have been studied. For example, in [Peijuan et al., 2017] geometrical considerations are presented that allow to improve the efficiency of the method, although limiting its applicability. In [Balesdent et al., 2013] work is done on a variant of the crude MCS. And as well as these, there are many modifications that are pending to be studied, and even proposed, that could further improve the qualities of the method.

Bibliography

M. Balesdent, J. Morio, and J. Marzat. Krigin-based adaptive importance sampling algorithms for rare event estimation. *Structural Safety*, 44(1-10), 2013. DOI: 10.1016/j.strusafe.2013.04.001.

B. Bichon, M. Eldred, L. Swiler, S. Mahadevan, and J. McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal*, 46(10), 2008. DOI: 10.2514/1.34321.

S. Choi, R. Grandhi, and R. Canfield. *Reliability-based Structural Design*. Springer, first edition, 2006. ISBN 9781846284441.

B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(145-154), 2011. DOI: 10.1016/j.strusafe.2011.01.002.

Z. Lv, Z. Lu, and P. Wang. A new learning function for Kriging and its applications to solve reliability problems in engineering. *Computers and Mathematics with Applications*, 70(1182-1197), 2015. DOI: 10.1016/j.camwa.2015.07.004.

R. Melchers and A. Beck. *Structural Reliability Analysis and Prediction*. John Wiley & Sons Ltd, third edition, 2018. ISBN 9781119266075.

A. Nowak and K. Collins. *Reliability of Structures*. McGraw-Hill, first edition, 2000. ISBN 0-07-048163-6.

Z. Peijuan, W. Ming, Z. Zhouhong, and W. Liqi. A new active learning method based on the learning function U of the AK-MCS reliability analysis method. *Engineering Structures*, 148(185-194), 2017. DOI: 10.1016/j.engstruct.2017.06.038.

L. Schueremans and D. Van Gemert. Benefit of splines and neural networks in simulation based structural reliability analysis. *Structural Safety*, 27(246-261), 2005. DOI: 10.1016/j.strusafe.2004.11.001.