Efficient sampling strategies to estimate extremely low probabilities

Cédric J. Sallaberry[a*], Robert E. Kurth[a]
[a] Engineering Mechanics Corporation of Columbus (Emc[2])
Columbus, OH, USA

Abstract: The probabilistic analysis of possible piping rupture or significant loss of coolant accident (LOCA) is a complex problem as it involves many mechanisms and generates low to extremely low probabilities of events. This topic is of particular interest in the nuclear industry and a conjoint effort between the US NRC and EPRI over the last 10 years has led to the development of the Extremely Low Probability of Rupture (xLPR) code to assess probability of rupture in nuclear piping systems. In this paper we focus on the sampling Monte Carlo methods selected to propagate uncertainty, as well as the optimization technique chosen (importance sampling) to reduce the sample size when estimating extremely rare events.

Keywords: Monte Carlo, DPD, LHS, Importance Sampling

1 INTRODUCTION

Nuclear infrastructures are required to perform in a highly reliable manner. Of particular interest is the assessment of piping rupture or a possible significant loss of coolant accident (LOCA) in a nuclear power plant. While the plant design has demonstrated a highly reliable system as the structure ages, together with the uncertainties in the process, the reliability will decrease with time. Particularly, the occurrences and evolution of damage that could occur due to several conditions (e.g. primary water stress corrosion cracking or fatigue) on welds that join two dissimilar or similar metal pipes is considered as a potential risk leading to adverse condition such as coolant leakage of pipe rupture. While the US NRC does not regulate on risk it does relies on risk-informed approach in order to understand the consequence of uncertainty and support the conclusions reached from probabilistic assessments.

To this end the US NRC, in conjunction with EPRI, has developed the Extremely Low Probability of Rupture (xLPR) code to assess probability of rupture in nuclear piping systems [1] [2]. This code models the likelihood and evolution of potential cracks in the weld. It considers several mechanisms and plant properties including crack initiation, growth, coalescence and stability, weld residual stresses and materials properties. The code also considers mechanical (MSIP, Overlay and Inlay) and/or chemical mitigation (Hydrogen concentration and Zinc addition), as well as the impact of in service inspections and leak detection. Due to the defense in depth approach, and the regular inspection schedule implemented coupled with constant measurement to detect potential leakage, the probability of rupture or other adverse events such as LOCA is expected to be extremely low (in the order of 10-5 or lower) that may render the standard, direct Monte Carlo method impractical.

This paper will focus on the probabilistic framework into which the xLPR v2.0 code has been integrated, in order to estimate low and extremely low probability of events with the constraints and features of the problem at hand: a large number of potentially uncertain inputs (more than 300), a large number of outputs of interest (crack occurrences, crack growth in both axial and circumferential directions, through wall crack occurrences, leak rate, LOCA,

* Csallaberry@emc-sq.com

rupture, etc.) during nominal operating conditions. Methods must also be able to address alternate conditions (e.g. mitigation, inspection, leak rate detection). Finally, upset conditions such as seismic events or sever transient events such as loss of station power, must also be able to be modeled.

The sampling strategies included in the current version of the code (xLPR v2.0) will be presented in light of their strengths and weaknesses: Simple Random Sampling (SRS), Latin Hypercube Sampling (LHS) and Discrete Probability Distribution (DPD). Importance Sampling implemented by the variance reduction technique, whose purpose is to reduce the sample to a practical size when estimating these extremely low probabilities of adverse events. The methodologies implemented in xLPR as well as the pre-processing code PROMETHEUS will be described and presented with practical examples.

2 CONTEXT OF THE ANALYSIS

The problems under consideration are usually complex, involving several different mechanisms. The number of uncertain inputs is fairly large (several hundreds) due to inherent randomness, spatial variability and lack of knowledge. In a risk-informed decision making process, it is necessary to look at several indicators, meaning that, often, 10 or more outputs have to be considered. Due to the multidimensionality of the problem, both in the input and output space, Monte Carlo methods are considered the most appropriate techniques to propagate the uncertainty through the system.

The adverse events are expected to be extremely rare (in the range of 10^{-5} or less). Such assumption is supported by the fact that such events have not occurred over any power plant in service. As a result, it means that a classical Monte Carlo approach would require a large number of runs (that can reach a million or more) to estimate with enough confidence their probability. The complexity of the code, as well as the amount of information required underlined the need of optimized techniques in order to reduce the sample size that would in theory be required.

A joint paper describes how the results can be analyzed to optimize subsequent runs of the code. The current paper focus on the techniques used upfront to generate the required information as well as those that can be used once the main drivers of the output uncertainty are identified. We present the sampling strategies considered in xLPR v2.0 and compare them via illustrative examples. We describe then the result of a generic analysis in which the strategy has been efficiently applied.

3 SAMPLING STRATEGIES

The multidimensional aspect of both the input space and output space, as well as the potential interaction between inputs (dependencies or correlation) have led to choose the Monte Carlo approach [3] to represent and propagate the uncertainty in the system. Probabilistic distributions are used to represent inputs uncertainty as well as model uncertainty via their parameters. Values are sampled from these distribution to create a set of values that is use to run the code deterministically. The operation is repeated several time (from a few hundreds to several hundred thousands) to generate a set of potential values for the output of interest which then can be analysed with classical statistical methods.

In the xLPR code, three sampling methods are considered: Simple Random Sampling, Latin Hypercube Sampling and Discrete Probability Distribution.

3.1. Simple Random Sampling (SRS)

SRS corresponds to the original Monte Carlo technique [4] . Each value is sampled randomly according to its probability distribution using a pseudo random number generator. The method is good when a large number of simulations can be performed with respect to the probability of event occurring. The law of great numbers indicates that the solution will converge in $\frac{1}{\sqrt{n}}$ where n represents the sample size (see [5] and [6] for instance). Experience suggest that the sample size has to be at least 10 to 20 times larger than $\frac{1}{P}$ where P represents the probability of interest, to generate stable results.

With simple random sampling, n quantiles values $(q_{i,k})_{i=1,n}$ are sampled uniformly from the interval $(0,1)$ for a selected input X_k , associated with a probability distribution. The inverse CDF is then used to estimate the sampled values for this input, with $(x_{i,k})_{i=1,n} = F^{-1}(q_{i,k})_{i=1,n}$ where F^{-1} represents the inverse of the cumulative distribution function.

3.2. Latin Hypercube Sampling (LHS)

LHS stratifies the input space in each of its dimension with (usually equally probable) denser stratification for all uncertain input individually. LHS usually reduces variance in the estimates as demonstrated in [5]. LHS leads to a better coverage of the CDF of each sampled input and is more efficient when some inputs have a strong influence on the output of interest by themselves.

Latin Hypercube requires actually two samples. The quantile space is decomposed into n equal strata. Each strata is selected randomly one and only one time. Within the strata, a value is selected randomly. The stratification forces the sampling to have at least one value in any interval of size $\frac{1}{n}$ allowing an even coverage of each distribution.

3.3. Discrete Probability Distribution (DPD)

As for LHS, the DPD [6] sampling technique discretizes each input's distribution into n strata as defined by the user. However, the number of strata is taken smaller than the actual sample size (usually the sample size is taken as a multiple of the number of strata n). Furthermore in each strata the *conditional expected* value is always used instead of a sampled value. Thus this method could be assimilated to some LHS implementation where the center of the strata is used as reference (except that expected value is used instead of median value). The discretization is not as dense as with LHS. The reduction in stratification coverage decreases the denser coverage in each direction while increasing the multidimensional coverage. This technique is more suited when conjoint influence of multiple inputs is expected to be a major contributor to the uncertainty on the output of interest.

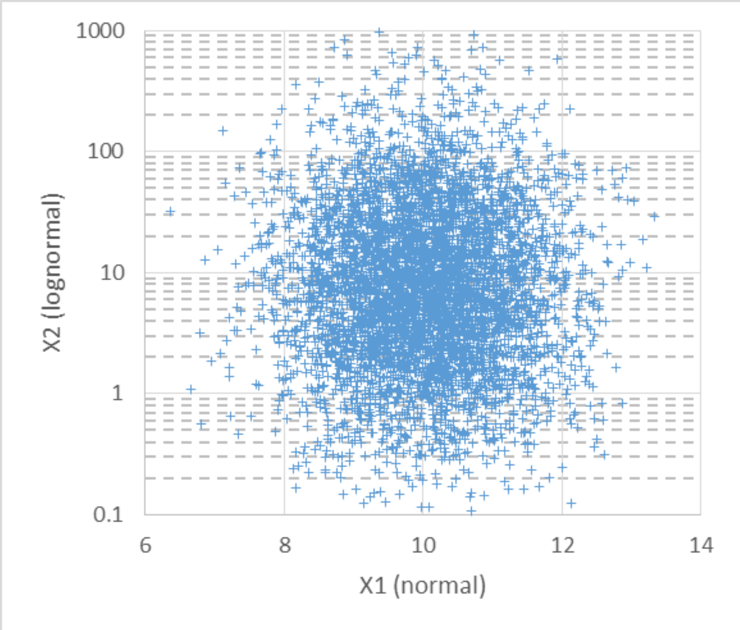
3.4. Comparison of the three methods

Both LHS and DPD tend to improve over simple random sampling, the former with allowing a better coverage of each input independently, while the later focusses more on the multidimensional aspects. The differences between the two techniques can be summarized as follow:

- In LHS the number of strata matches the sample size while it is smaller in DPD
- The value in each strata is sampled randomly LHS while the conditional expected value is used in DPD
- LHS sample without replacement (each strata is sampled once in each direction) while DPD sample with replacement.

To illustrate the differences between the three sampling techniques, a set of samples of size 50 for two variables ($X_1 \backsim N(\mu = 10; \sigma = 1)$ and $X_2 \backsim LN(\mu = 2; \sigma = 1.5)$) will be displayed on a semi log scale. The same 50 random numbers will be used to generate the values for SRS and the LHS and DPD corresponding strata in order to take our any bias due to a difference in random numbers for such a small sample size. Furthermore, DPD will use 10 strata. With a larger sample size the values displayed on a semi-log scale will presents the classical profile of a bi-normal distribution with a circular pattern whose density decreases from the center of the shape.

Figure 1: example of sampled values for a sample of size 5000

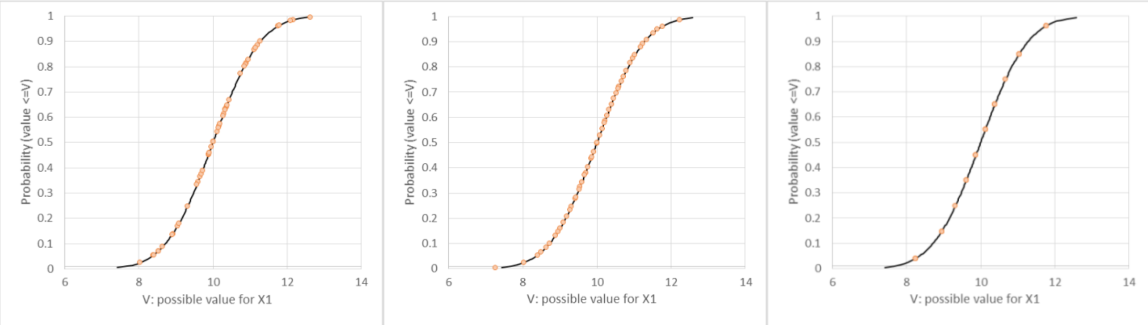


The hope is that a smaller sample size will be able to represent this behaviour as well as cover each distribution correctly.

These methods are compared for the sampled values for each uncertain input (represented by the theoretical CDF as well as the repartition of values on the (X_1, X_2) space.

Figure 2 displays the sampled values for X_1 using Simple Random Sampling (left frame), Latin Hypercube Sampling (middle frame) and Discrete Probability Distribution (right frame). SRS has a coverage of the response space that is uneven, with many areas left uncovered (the parts of the CDF not sampled appear as a dark line). As expected, the use of equiprobable strata gives a more even coverage of the CDF. DPD also has an even coverage but with larger distances between the points (note that in practice, DPD will be used with 100 to 200 strata for larger sample sizes, leading to a better coverage).

Figure 2: Comparison of sampled values for X_1using SRS (left), LHS (middle) and DPD (right)



Similar conclusions can be drawn from Figure 3 with the sampled values of X_2.

Figure 3: Comparison of sampled values for X_2using SRS (left), LHS (middle) and DPD (right)

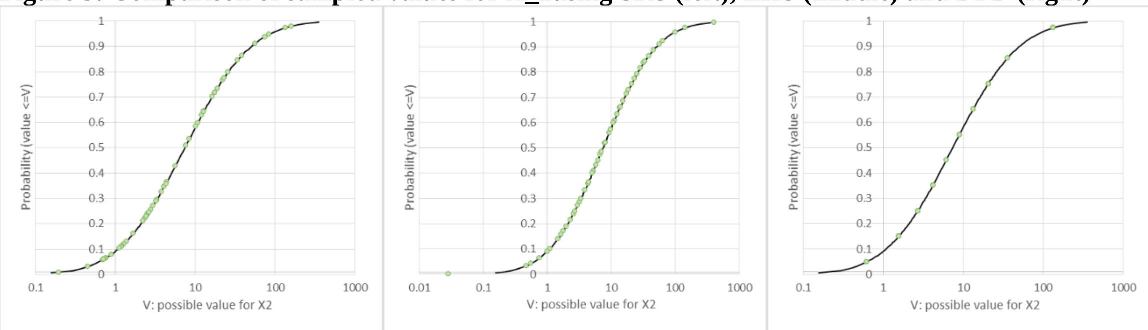
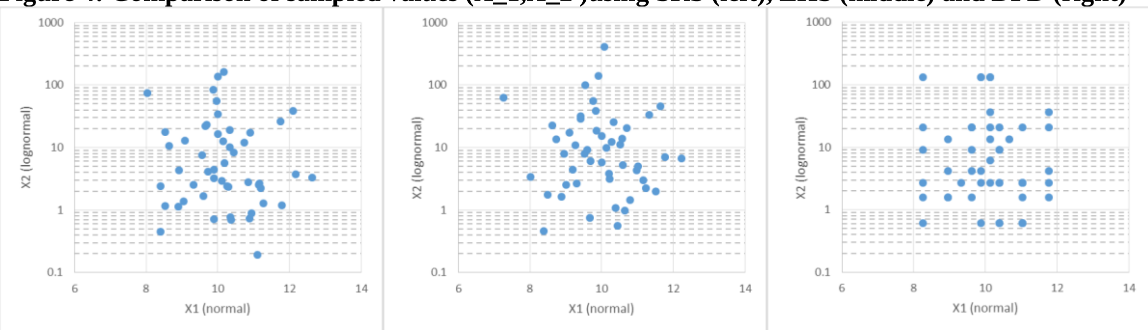


Figure 4 displays the pairs $(X_{1,i}, X_{2,i})$ for each element i of the sample on a 2D scatterplot. There are few variations between SRS and LHS with a slight change in the values for a better coverage of each input independently. On the other hand, DPD graph looks different, but it provides a better systematic coverage of the space.

Figure 4: Comparison of sampled values (X_1,X_2)using SRS (left), LHS (middle) and DPD (right)



This simple example illustrates the differences among sampling techniques and the importance of understanding their strengths and weaknesses. With large enough sample sizes, *all* techniques will converge to the correct solution. In conclusion, any of these techniques is appropriate for a first exploratory analysis of the PFM problem. Then the most appropriate sampling technique can be selected, once the important mechanisms and drivers have been identified.

3.5. Importance Sampling

Importance sampling is considered., for SRS and LHS, as a variance reduction technique because its purpose is to reduce the variance of the estimate of the quantity of interest. The idea behind the reduction is to under-sample the regions which have very little, or no, impact on the estimate (since less samples are required) and over sample the regions of importance. This way, the estimate of the event of interest can be more accurate with fewer samples.

The main difficulty with importance sampling is to know which regions of the input space are important to the response being studied. (These regions can change depending on the response). Since the regions of importance are defined in the input space, it means that for an importance sampling to be efficient, one needs to know which input variables affect the output of interest and how they affect it. This part is often not straightforward in complex systems usually considered in PFM studies, due to many interactions in the system, and require a prior sensitivity analyses in order to support the selected important inputs, often confirmed with expert elicitation. An incorrect choice of inputs may lead to an under-sampling of the region of interest and a worse estimate than with standard sampling. As a result, a careful and thorough analysis prior to input selection is required and needs to be documented to support any use of importance sampling. One sampling strategy that does not require tis added step, adaptive sampling, is discussed later.

The implementation of the methodology is usually fairly straightforward and involves classical integration calculations on the ratio of density between the initial distribution and the importance distribution. In Monte Carlo methods, when the traditional approach of sampling from a uniform distribution and the inverse CDF is used to generate samples from a large number of distributions, importance sampling is often applied to the initial uniform quantile sampling.

As an example, a Latin Hypercube sample of size 50 has been generated for a standard normal distribution (of mean $\mu = 0$ and standard deviation $\sigma = 1$). The resulting sample is displayed in Figure 5. The coverage of the normal distribution is good for 80% of the range ($q = 0.1$ to $q = 0.9$) but outside of this range the coverage is less (which is consistent with the density associated with a normal distribution). If the event is only associated with low to extremely low (or high) to extremely low (or high) values, a LH sample of this size may not be sufficient.

Figure 5: regular LHS sample of size 50 on a standard normal distribution

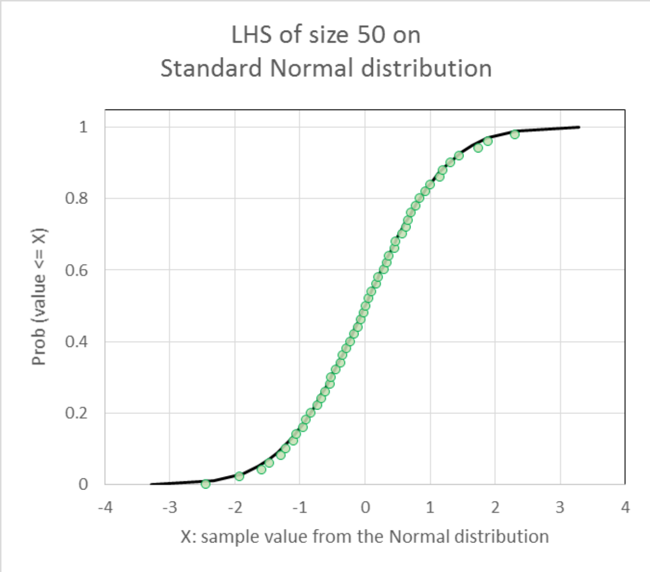
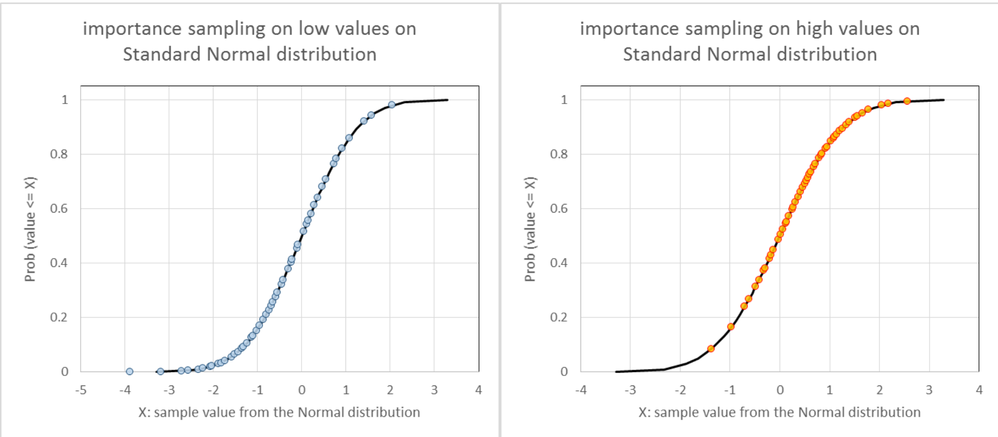


Figure 6 displays two importance samples of same size (50), one emphasizing the low values (left frame – blue dots) and the other the high values (right frame, orange dots). Importance sampling compresses the sampling in the area of interest while expanding it elsewhere. These two samples illustrate one of the disadvantages of the method, the extreme region not focused on by the importance sampling is only covered by a few points. An error in the region of interest may thus lead to dramatically incorrect results.

Figure 6: importance sampling focusing on low values (left) and high values (right) on a standard normal distribution

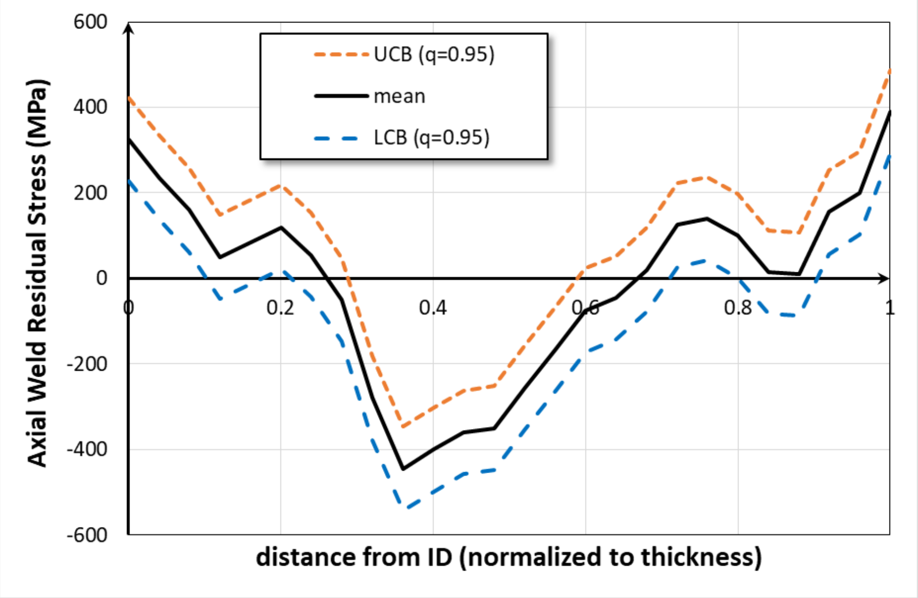


The expansion and contraction of the coverage means that the sampling is no longer even in the quantile sense (Y-axis on the CDF). Each realization is thus associated with a weight, representing the ratio between the original density and the importance density, in order to preserve the initial distribution.

4 EXAMPLE

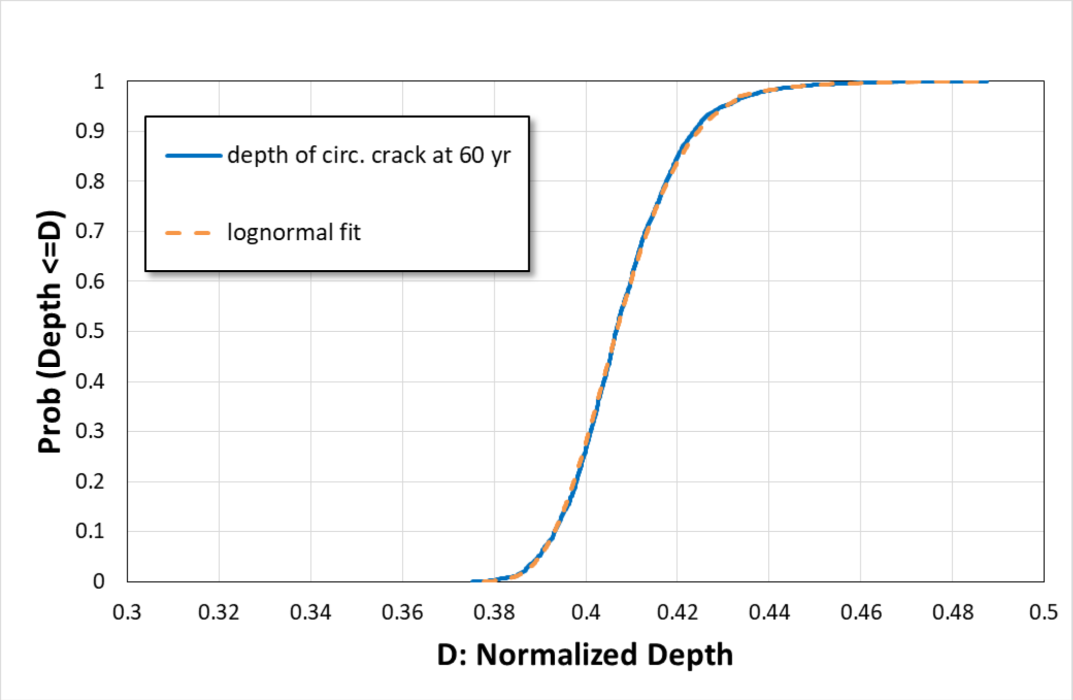
In this generic example, we consider a crack growing circumferentially through the depth of a weld. The Axial Weld Residual Stress Profile is considered uncertain and is displayed in Figure 7. Due to the double V-groove shape of the weld, there is a strong compressive zone about half through the thickness, stopping the growth of the vast majority of circumferential cracks. If some cracks can grow beyond the bottleneck, then the stress are going tensile again, leading potentially to through wall crack and even pipe rupture. Due to axisymmetric conditions, the WRS profile is equilibrated to integrate to zero. As a result high positive values are associated with low negative value and reciprocally.

Figure 7: axial WRS mean value and q=0.95 confidence interval



A sample of size 2,500 was used to generate the distribution on the maximum circumferential depth for the duration of the simulation of 60 yr. (Figure 8).

Figure 8: Cumulative Distribution Function (CDF) of maximum circumferential crack depth at 60 yr.



The lognormal fit applied on the distribution allowed an estimate of the potential likelihood of a crack reaching a certain depth. The threshold depth that could lead to the crack too begin to grow again is not known. As a result, we list in Table 1 the resulting likelihood for different potential threshold depths ranging from 0.5 to 0.55 of the thickness

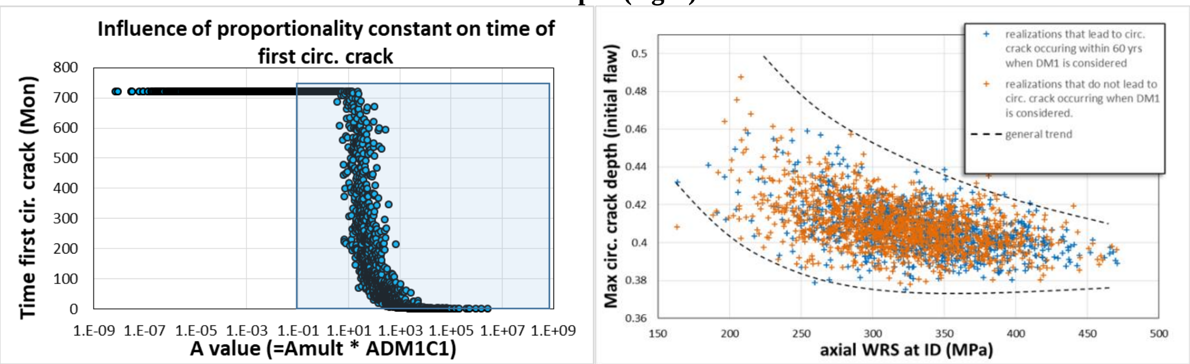
Table 1: theoretical probabilities using the fitting distribution

a	$P(X > a)$
0.5	1.28×10^{-4}
0.51	5.88×10^{-5}
0.52	2.73×10^{-5}
0.53	1.29×10^{-5}
0.54	6.14×10^{-6}
0.55	2.97×10^{-6}

Such intermediate analysis helps to define the expected number of realization that would be required to generate some realizations through wall crack or rupture (between 20,000 and 1,000,000 realizations).

Using sensitivity analysis, the main contributors to the crack maximum depth have been identified as the proportionality constant parameter on the selected crack initiation model and the Axial Weld residual Stress. Scatter plots of their influence is displayed below. High values of the proportionality constant (x-axis left frame) leads to earlier cracks occurrences, leaving time for the crack to grow. WRS at the ID has a negative influence due to the axisymmetric condition. Low values at the Inner Diameter (ID) lead to higher values around the bottleneck, which leads to deeper circumferential cracks.

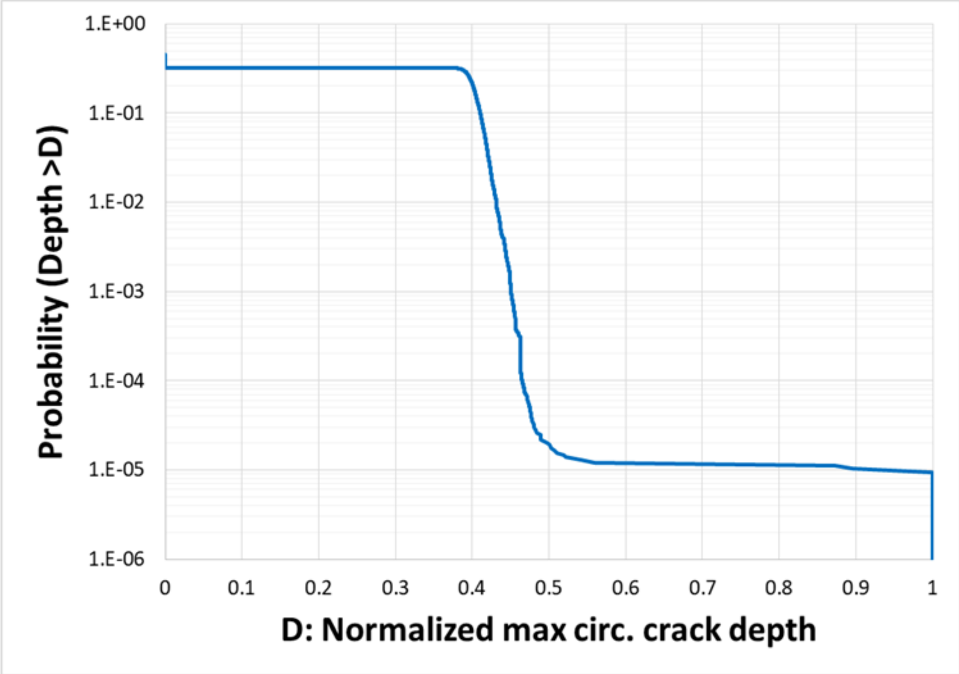
Figure 9: scatterplot of the influence of the A parameter on circ. crack initiation (left) and of the WRS value at the Inner Diameter on Maximum crack depth (right)



Those two values have been considered as candidates for importance sampling, with emphasis on high values for the proportionality constant and low values for WRS at the ID. An importance sampling of size 10,000 was then run.

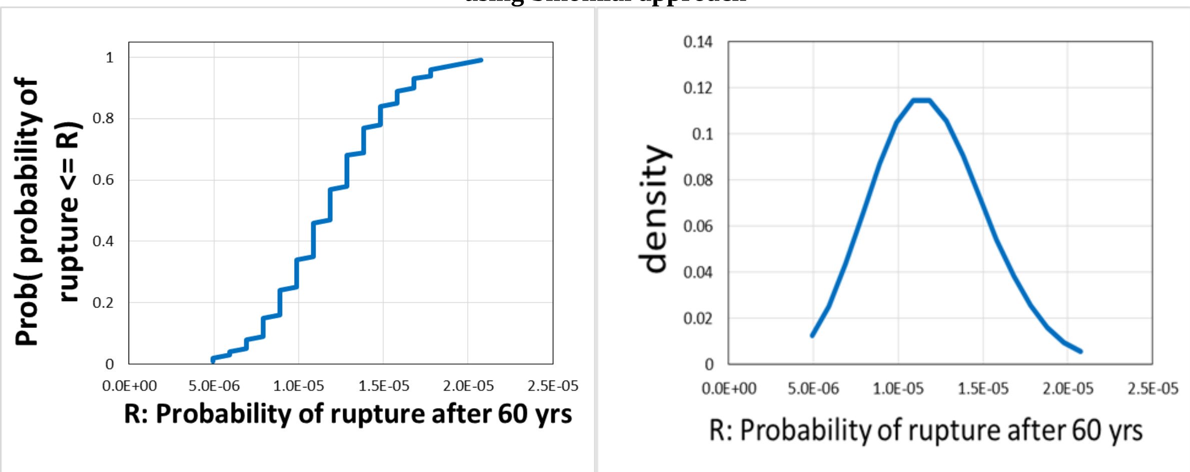
The resulting Complementary Cumulative Distribution of circumferential crack depth, displayed below (Figure 10), estimate the probability of TWC and potential rupture occurring to be around 10^{-5} which is consistent with the log-normal fit.

Figure 10: CCDF of circumferential crack depth for the 10,000 importance sampling run



A binomial distribution has been used to represent the potential error toward the potential probability of rupture estimate. The resulting CDF and PDF are displayed below (Figure 11). It is satisfying to discover a normal pattern shape in the error representation as it is a good indicator of convergence of the estimate.

Figure 11: CDF and PMF on distribution of mean probability of rupture from a circumferential crack using binomial approach



A run of the xLPR code on a single processor with a sample size of 10,000 takes about 10 hours. The use of importance sampling allowed to estimate with enough confidence a probability in the range of 10^{-5} within reasonable computational time compared to what would be needed with a traditional Monte Carlo method. In our case, a single realization runs in about 4 seconds which would require about 46 days on a single processor to run 1,000,000 and reach similar accuracy in the estimate.

5 CONCLUSION

Direct Monte Carlo method cannot sometimes be used to estimate the likelihood of extremely rare events in complex systems. Even with optimized method, the code may be too slow (or the amount of information too large) for the analyst to afford millions of runs or more. The use of optimized methods using more efficient strategies are a way to overcome this pitfall and generate stable answer that can support risk-informed decision making.

These methods require often a better understanding of the system under consideration and to identify the most important parameters influencing the risk. Consequently they have to be supplemented with sensitivity analysis (presented in a conjoint paper) as well as other sensitivity studies to increase confidence in the results.

In order to improve the approach, we are now working on adaptive sampling that find automatically the area of interest in the input space so the analyst is guided toward this process. However, we fully acknowledge that there is no single universal method that will work on any situation and that at the end, it is the responsibility of the analyst and the experts to check any results.

Acknowledgements
The development of the Version 2.0 xLPR code is a group effort that spans a variety of experts across many fields of expertise from the U.S. NRC, EPRI, and their contractors. The ongoing success of this program emphasizes the dedication of the xLPR team, the strength of its leadership, and the support from both the NRC and EPRI.

6 REFERENCES

[1] D. Rudland and C. Harrington, "xLPR Pilot Study Report," U.S NRC-RES Washington, DC and EPRI, Palo Alto, CA, NUREG-2110 and EPRI 1022860, May 2012.

[2] D. Rudland, C. Harrington and R. Dingreville, "Development of the extremely low probability of rupture (xLPR) Version 2.0 code," in *Proceedings of the ASME 2015 Pressure Vessels & Piping Division*, Boston, MA, USA, 2015.

[3] C. Lemieux, Monte Carlo and Quasi-Monte Carlo Sampling, Springer, 2009.

[4] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association*, vol. 44, p. 335, 1949.

[5] J. C. Helton and F. J. Davis, "Latin Hypercube Sampling and the propagation of uncertainty in analyses of complex systems," *Reliability Engineering and System Safety,* vol. 81, no. 1, pp. 23-69, 2003.

[6] C. B. Storlie and J. C. Helton, "Multiple predictor smoothing methods for sensitivity analysis: Description of Techniques," *Reliability Engineering and System Safety,* vol. 93, no. 1, pp. 28-54, 2008.

[7] Wikipedia contributers, "Central Limit theorem," Wikipedia, The Free Encyclopedia., 26 september 2017. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Central_limit_theorem. [Accessed 28 september 2017].

[8] "Wikipedia page on law of large numbers," [Online]. Available: https://en.wikipedia.org/wiki/Law_of_large_numbers. [Accessed 2017].

[9] R. E. Kurth and D. C. Cox, "Discrete Probability Distributions for Probabilistic Fracture Mechanics," *Risk Analysis,* vol. 5, no. 3, pp. 235-240, 1985.