**DIGITALCOMMONS**
—@WAYNESTATE—

**Wayne State University**

Civil and Environmental Engineering Faculty Research Publications

Civil and Environmental Engineering

10-8-2010

# Reliability Estimation of Complex Numerical Problems Using Modified Conditional Expectation Method

Christopher D. Eamon
*Wayne State University, Detroit, MI*, christopher.eamon@wayne.edu

Bulakorn Charumas
*Mississippi State University, Starkville, MS*

# Reliability Estimation of Complex Numerical Problems using Modified Conditional Expectation Method

Christopher D. Eamon[1] and Bulakorn Charumas[2]

**Abstract**

A simulation-based structural reliability analysis method is presented. It is intended as an alternate approach to estimate reliability for problems for which most-probable point of failure methods fail and when computational resources are limited.   The proposed method combines conditional expectation and estimating the PDF or CDF of a selected portion of the limit state.   In the proposed approach, complex limit state functions are simplified to two random variable problems.   The success of the simplification depends on the quality of the CDF estimate.   Results indicate that the method may provide accurate and efficient solutions for some difficult reliability problems.

--------------------

1.  Associate Professor of Civil Engineering, Wayne State University, Detroit, MI.
eamon@eng.wayne.edu
2. Former Graduate Student, Civil Engineering, Mississippi State University, Starkville, MS.

**Introduction**

For structural reliability problems with well-behaved limit state functions, most probable point of failure (MPP) search or reliability index-based methods are often the first choice for reliability analysis, as they can typically achieve accurate results with many fewer calls to the response function than simulation methods such as Monte Carlo simulation (MCS) or one of the various variance reduction techniques (VRTs). The widely-used reliability-index based methods include the first- and second-order reliability methods (FORM, SORM) (Rackwitz and Fiessler 1978; Breitung 1984), with many variants presented in the literature (Chen and Lind 1983; Wu and Wirsching 1987; Fiessler et al. 1979; Hohenbichler et al. 1987; Tvedt 1990; Der Kiureghian 1987; Der Kiureghian et al. 1987; Ayyub and Haldar 1984, among many others). VRTs such as importance sampling (Rubinstein 1981; Engelund and Rackwitz 1993) and adaptive importance sampling (Wu 1992; Karamchandani et al. 1989), also make use of the MPP concept, and can similarly lead to significant reductions in computational effort over MCS. For ill-behaved or difficult to capture responses, however, such as those which may be discontinuous, highly nonlinear, or that contain multiple 'local' reliability indices on the limit state boundary, as with many complex problems requiring a numerical or finite element solution, the most probable point (MPP) search algorithms may fail or produce unstable or erroneous results. In such cases, one must rely upon a greatly reduced selection of techniques, primarily those from the simulation family that do not rely upon an MPP search such as MCS and its advanced variants (Au and Beck 2001; Au et al. 2007) or stratified sampling methods (Iman and Conover 1980).

An alternative common approach is approximating the true limit state function with a response surface (RS), of which many examples exist (Bucher et al. 1990; Gomes et al. 2004; Cheng et al. 2009, etc.) Point integration or point estimation techniques would also be

possible, although results may be highly unreliable (Eamon et al. 2005). The drawback of many sampling techniques is the effort required, particularly for high-reliability problems involving a computationally expensive, implicit limit state function. Similarly, for complex responses (highly nonlinear or discontinuous), it is may be difficult to develop a sufficiently accurate response surface for reliability analysis without expending considerable computational effort. This is particularly so when the number of random variables becomes large.

As the fast and accurate reliability analysis of complex engineering problems is a topic of great interest, there have been various recent developments in simulation methods that were developed to address these concerns. One such promising method is subset simulation (Au and Beck 2001), several versions of which have been favorably evaluated elsewhere (Au et al. 2007). This paper presents an alternative approach. Similar to existing advanced simulation methods, it describes a simulation-based method that does not rely on an MPP search, but that can generate accurate results for difficult limit states of relatively high-reliability, with a reasonable computational effort. We refer to this method as Failure Sampling (FS), as even though sample points are neither guided by nor correspond to the MPP, all samples are taken on the failure boundary.

**Concept of Failure Sampling**

The method is based on a combination of conditional expectation (CE), of which there are many versions proposed in the literature (see for example, Karamchandani and Cornell 1991; Ayyub and Chia 1992; Yasuhiro and Ellingwood 1993, Smarslok et al. 2006, Eamon 2007, among others), and estimating either the probability density function (PDF) or cumulative distribution function (CDF) of a specific portion of the limit state function by direct MCS.

Of course if the PDF (or CDF) of the complete limit state function $g$ can be estimated accurately, probability of failure, $p_f$ can be easily calculated by numerical integration over the failure region using the well-known expression:

$$p_f = \int_{-\infty}^{\infty} F_R(x) f_q(x) dx \tag{1}$$

The difficulty of estimating this curve by MCS is that most samples fall close to the mean of $g$, which may be far from the failure region, making accurate integration of this region difficult or impossible.

FS solves this problem by introducing concepts from CE. Traditional CE involves several steps: 1) From an original limit state function $g(X_j)$, a control random variable Q is chosen, which is statistically independent from the other random variables and is usually taken as the random variable (RV) with largest variability in the original limit state function $g$. 2) A new limit state function $g'$ is formed such that it has an equivalent failure boundary to that of $g(X_j)$, but in $g'$, Q is separated from the remaining RVs. That is, at failure, $g$ is expressed as $g' = 0 = R(X_i) - Q$. Here $R(X_i)$ is the portion of the limit state that is not a function of Q, and $X_i$ is the set of all RVs ($X_j$) except Q. 3) MCS is then used to simulate values for the RVs $X_i$, and $R(X_i)$ is evaluated for simulation $s$ (where $s$ is one of the total simulations taken). 4) Since at failure ($g' = 0$), $R(x_i)_s = q_s$, the cumulative distribution function (CDF) of Q, $F_Q$, evaluated at $q_s$, must also equal the value of $F_Q$ evaluated at $R(x_i)_s$: $F_Q(q_s) = F_Q(R(x_i)_s)$. Thus, $F_Q(R(x_i)_s)$ is evaluated for simulation $s$. 5) $p_f$ for simulation $s$, $p_{fs}$, can then be calculated as: $p_{fs} = P(Q > R(x_i)_s) = 1 - F_Q(R(x_i)_s)$. 6) Repeat steps 3-5 as desired, for $n$ total simulations. 7)

4

The final $p_f$ estimate is calculated as the mean of the $n$ failure probabilities $p_{fs}$:

$$p_f = \left(\sum_{s=1}^{n} p_{fs}\right)/n.$$

To clarify the above process with an example, consider a simple limit state function $g = X_1 X_2 - X_3/X_4$. Assume $X_3$ is chosen as the control variable Q. An equivalent limit state boundary is formed by rewriting $g$ as $g'$ such that $X_3$ is separated from the remaining RVs. The new limit state boundary is written as $g' = 0 = X_1 X_2 X_4 - X_3$. Here, $X_1 X_2 X_4 = R(X_i)$ and $X_3 = Q$. $p_f$ of $g'$ can be calculated as: $p_f = P(X_3 > X_1 X_2 X_4)$. To calculate $p_f$, MCS is used to simulate values for $X_1$, $X_2$, and $X_4$. Using these simulated values, $p_f$ for simulation $s$ can be written as: $p_{fs} = P(X_3 > x_1 x_2 x_4) = 1 - F_Q(x_{3s})$, where $x_{3s}$ is the value of $X_3$ that would cause a failure ($g' = 0$) for simulation $s$. Since at failure, $x_3 = x_1 x_2 x_4$, $p_f$ can also be written as: $p_f = 1 - F_Q(x_1 x_2 x_4)$. Multiple simulations are conducted, and the average $p_f$ result from all simulations is taken as the final $p_f$ estimate.

FS shares steps 1-3 above with CE, but replaces steps 4-7 with a different procedure. In CE, neither the PDF nor CDF of $R(X_i)$ is determined, but only the response of R for a simulated set of RV values $(x_i)$. In contrast, with the FS approach, estimating the PDF or CDF (as needed) of $R(X_i)$ is the primary concern. Consider the reformulated limit state per step 2 of CE as a starting point (and renaming $g'$ to $g_{fs}$):

$$g_{fs} = R(X_i) - Q \qquad\qquad (2)$$

As with CE, the limit state boundary of $g_{fs}$ is simply the boundary of $g$ rewritten, separating Q from the remaining RVs. There is no theoretical limitation to the selection of the control variable Q, other than statistical independence from the remaining RVs $X_i$, and thus $X_i$ and Q

are not required to be the actual resistance and load RVs of the problem. Note that $g_{fs}$ can be completely implicit, and there is no need for an analytical or closed-form formulation.

The advantage of estimating the PDF (or CDF) of $R(X_i)$, rather than the entire PDF of $g$, is twofold. First, as with CE, the variance of the control variable is removed from the Monte Carlo simulation. More importantly, the value of $R(X_i)$ at the MPP of $g_{fs}$ is generally closer to the central region of the PDF of $R(X_i)$, $f_R$, than $g(X_j) = 0$ is to the central region of the PDF of $g$, $f_g$. Therefore, if $f_R$ is estimated, such as by MCS, the calculation of $p_f$ of $g_{fs}$ becomes less sensitive to the tail region of $f_R$, than is the $p_f$ of $g$ to the tail of $f_g$. This allows a more accurate computation of $p_f$ of $g_{fs}$ than the $p_f$ of $g$, with the same number of simulations. This is illustrated in Figure 1, which represents the results of example problem 1, discussed below. Clearly, the failure region of $g_{fs}$ that is associated with $f_R$ is larger (lower graph in the figure), and thus easier to capture for the same number of simulations, than the failure region of $f_g$ (upper graph in the figure).

When compared to CE, the advantage of FS is that (an estimate of) the PDF or CDF is used to calculate $p_f$ rather than the independent data points $x_i$. For reasonable reliability problems, this typically results in a greater number of data available in the area of interest, that representing the lowest values of $R(X_i)$. That is (assuming the number of simulations is kept constant between CE and FS), for FS, more data are available in the tail region of the PDF or CDF estimate of $R(X_i)$, for use in eq (2), then are available in the tail region of the direct $R(x_i)$ samples used in the CE expression $p_f = \left( \sum_{s=1}^{n} (1 - F_Q(R(x_i)_s)) \right) / n$. This higher data density in the area of interest for FS leads to a more accurate estimate of $p_f$. For example, consider the CDFs shown in Figure 2, which are from the results of example problems described in detail below. The CDF to the left is calculated directly from the cumulative

probability of 1000 MCS samples: $F_R(x_i) = s/(1+n)$. These $R(x_i)$ values would be used in the CE $p_f$ calculation above. The CDF to the right represents an (50 point) estimated $F_R$ from the same 1000 samples, which is used in FS. Although there are 20 times as many samples that define the CDF to the left in the Figures, they are clearly less dense in the lower tail of the distribution, the region which is most critical to the $p_f$ estimate, than is the CDF estimate to the right. These results are typical.

As with CE, complex reliability problems solved with FS are simplified to a algebraically linear, two random variable problem. By doing so, the complexity of the original model is clearly lost in the solution of the surrogate problem. Using this approach, how well the solution of the surrogate problem represents that of the true problem critically depends on how accurately $f_R$ (or $F_R$, as needed) is estimated. Details on obtaining a suitable estimate of $f_R$ or $F_R$ from the $R(x_i)$ samples are discussed below

For FS, the estimation of $f_R$ or $F_R$ is conducted by imposing the condition $g_{fs} = 0$. For many practical (i.e. nonlinear and implicit) problems, imposing $g_{fs} = 0$ generally requires a higher computational cost per FS sample than a MCS sample, as each FS sample requires solution of $g_{fs} = 0$ for q. For an implicit, nonlinear problem, this may require multiple iterations. However, as will be shown, this cost is greatly outweighed by the savings in overall number of samples required, especially as $p_f$ decreases.

**FS Algorithm**

The FS algorithm is simple to implement and works as follows.

1. Choose a control RV and reformulate $g$ to $g_{fs}$ as per eq. (2). Although it is not required to explicitly re-write the limit state in terms of the control variable, doing so would avoid the need for a nonlinear solution and decrease overall computational effort. The control variable is best taken as the RV with highest variability, but is often most easily taken as a load magnitude RV for implicit problems. The only restrictions are that the control RV must be statistically independent from the remaining RVs and be able to satisfy $g_{fs} = R(x_i) - q = 0$, for the expected range of values of $X_i$ that will be obtained by the sampling method chosen. The effect of control variable choice is discussed further in a later section.

2. Assign random values to the RVs within R, $(x_i)$, using MCS or an alternative non-MPP based simulation method. In this paper, MCS is used and the potential advantages of using other methods with FS have not been explored.

3. Determine the value of the control variable needed to satisfy $g_{fs} = 0$: $q = R(x_i)$. For implicit, nonlinear problems, a variety of numerical nonlinear solvers are available, but in general q is incrementally increased (or decreased) until the limit state $g_{fs}$ equals zero.

4. Repeat steps 2-3 until a sufficient number of samples have been taken. As with MCS, the most reliable way to determine appropriate sample size is to increase the number of samples until the solution converges. Some guidance can be obtained from the example problems below.

5. Using the data from step 4, the PDF of $R(X_i)$, $f_R$, or its CDF, $F_R$, can be estimated, as needed, depending on the method that will be used to calculate $p_f$. If a sufficient number of samples are taken, good estimations of the statistical moments of $R(X_i)$ as well as its distribution can be obtained, if desired. The methods used to estimate $f_R$ and $F_R$ are discussed in the next section.

6. With the statistical parameters of $R(X_i)$ estimated, the problem is effectively reduced to a 2-RV linear problem as shown by eq. (2), where $R(X_i)$ is now represented as a single random variable R, and Q as the control variable. Although algebraically linear, this equation will likely be nonlinear in standard normal space, as R is typically non-normal. Either reliability index or $p_f$ can now be readily computed using any desired method. For the validation problems considered in this paper, two separate approaches were considered for comparison. The first is to calculate $p_f$ by numerically integrating eq. (1).

There are various ways to estimate $F_R$. The most direct way is take $F_R$ as the CDF of the $R(x_i)$ samples; i.e. the cumulative probability of the $n$ samples ($F_R(x_i) = s/(1+n)$). As noted above, however, using this method offers little advantage over CE. Alternatively, if a PDF of the data samples is constructed, $F_R$ can be estimated by numerical integration of the PDF. The advantage of this approach was discussed above (see Fig. 1). This is the method used for FS, with specific implementation details given in the next section.

Error in failure probability prediction caused by error in the FS-estimated CDF as compared to the of the original resistance data $R(X_i)$ can be directly calculated as the

differences in the estimated contributions to failure probability from the two cases at the considered sample points, which can be numerically estimated as:

$$pf_{err} \approx \left[ \sum_{s=1}^{n} \left( F_R(x_s) - F_{R\,FS}(x_s) \right) \right] f_q(x_s) w(x_s) \qquad (3)$$

where $F_R(x_s)$ represents the CDF of the original data and is calculated based on the direct cumulative probability of the n samples: $F_R(x_i) = s/(1+n)$); $F_{R\,FS}(x_s)$ is calculated based on the FS-estimated CDF; i.e. formed by numerical integration of the PDF; and $w(x_s)$ is the interval width associated with resistance sample $x_s$; $w(x_s) = (x_s - x_{s-1})/2 + (x_{s+1} - x_s)/2$. As the number of sample points taken increases, differences between $F_R(x_s)$ and $F_{R\,FS}(x_s)$, and thus the failure probability error estimate, decrease as the solutions converge.

For the second approach used to evaluate the reliability of eq (2), a PDF of the data is constructed, then an analytical distribution is fit to the PDF. This analytical distribution thus becomes an estimate of $f_R$. With R represented by a known distribution type, $p_f$ or $\beta$ can be easily calculated from eq. (2) using any method. In this paper, reliability index $\beta$ was calculated using FORM with this approach.

Although various functions are available for distribution fitting, the authors have obtained consistently good results with the Generalized Lambda Distribution (GLD), which is highly flexible and can accurately describe many distribution shapes. The PDF of GLD has the form

10

$$f_g = \frac{\lambda_2}{[\lambda_3 u^{(\lambda_3 - 1)} + \lambda_4 (1-u)^{(\lambda_4 - 1)}]} \qquad (4)$$

where $\lambda_i$ are parameters determined from the first four statistical moments of the simulated data.  Various references describe the GLD and how to obtain its parameters (Karian et al. 2000; Ozaturk and Dale 1982; Asif and Helmut 2000).  With this procedure, differences between the skew and kurtosis coefficients of the original data and the fitted curve can be used as a measure of how accurately the data are represented.  However, eq. (3) could also be used.  Note that once the FS samples are taken in step 4, no more calls are made to the original limit state $g_{fs}$ to evaluate eq. (1) or (4),  and the additional computational effort needed to conduct steps 5 and 6 to obtain $p_f$ or $\beta$ is negligible (several seconds on a desktop computer).

Although the authors expected the first approach (direct integration of the data samples using eq. (1)) to give superior results than the curve-fitting method of approach 2 using the GLD, in most cases, little difference in accuracy was found between these two approaches.

**Example Problems**

Three representative problems are chosen to illustrate the results of FS: 1) a material-only nonlinear structural system problem; 2) a dynamic problem with a discontinuous limit state; and 3) a large strain, large displacement problem with nonlinear geometric and material response.

Note that for each of the numerical problems, the MPP could not be located, and thus FORM, SORM, Importance Sampling, and other popular methods that rely on MPP cannot be accurately applied here.

As with any simulation method, the accuracy of FS tends to improve as the number of simulations increases. In this paper, 1000 simulations were taken for most problems as the baseline for consistent comparison. For all problems, considering the first FS approach discussed in step 6 above, for use in eq. (1), $f_R$ was estimated by dividing the 1000 $R(x_i)$ results into 50 intervals. A 50-point estimation of $F_R$ was then obtained by numerically integrating $f_R$. Finally, $p_f$ was calculated from eq. (1) by integrating with the trapezoidal rule using 100 intervals, then the standard normal transformation $\beta = -\Phi^{-1}(p_f)$ was used to report generalized $\beta$ (given as the result "FS (NI)" in the tables).

When the second approach discussed in step 6 was considered, the GLD was fit to the $f_R$ estimate as described above. Reliability index was then computed using FORM from the 2 RV problem $g = R - Q$ based on eq. (1), with results given as "FS (GLD)" in the tables.

For comparison, the problems were also solved with traditional conditional expectation (CE), Monte Carlo simulation (MCS), and a common quadratic response surface approach (RS). For MCS, often no solution could be obtained if the computational effort is limited to that of FS; this is reported as 'no failures', or 'n.f.' in the results tables (in this case, failure probability would be estimated as 0 with MCS). For RS, a design of experiments was used to approximate the limit state function with a full quadratic response equation (full factorial composite design) using the code VisualDOC (Vanderplaats 2008), then MCS was used to evaluate failure probability of the response surface using a large enough number of samples

to obtain an 'exact' $p_f$ estimation of the response surface. For complex problems, CPU time is proportional to the number of function calls to the true response (i.e. FEA code), which is reported in the solutions. For consistent comparison, CPU time was kept constant for the comparison methods (FS, MCS, and CE), since a typical FS or CE sample takes more computational time than a MCS or RS sample, as discussed above.

For all methods, multiple trials were conducted and the average of absolute value of error, or 'typical' error, is presented, as well as the coefficient of variation of reliability index, reported as COV in the tables. Also indicated is whether a method usually produces a reliability index higher or lower than the true solution, which is indicated with a positive or negative value on the reported error, respectively. Most of the example problems have a range of $\beta$ from 2-4, which we consider a realistic range for most structural reliability problems, though problems with $\beta$ as high as 8 were also accurately solved with this approach.

**Problem 1. Nonlinear Static Truss**

The truss shown in Figure 3 is subjected to a random load P. The material is modeled as steel with a bi-linear stress-strain curve. Each member $i$ has 4 independent RVs for cross-sectional area A (mean 2 sq. in, COV 0.05), yield strength $\sigma_y$ (mean 50 ksi, COV 0.10), and post-yield modulus $E_2$ (mean 1,200 ksi, COV 0.25). Mean load P is taken as 40 or 55 kips, with COV of 0.10, for a total of 31 RVs. All RVs are normal. Failure occurs when the maximum stress in member no. 1 (only) $\sigma_1$ reaches its yield strength $\sigma_{y1}$. The limit state function is

$$g = \sigma_{y1} - \sigma_1(P, \sigma_{yj}, E_{2i}, A_i) \quad \text{for i = 1 to 10, j = 2 to 10} \qquad (5)$$

The limit state was evaluated with the commercial FEA code MSC Nastran (2005) using a Newton-Raphson solution. For FS and CE, P was taken as the control variable and the bisection method was used to find the value for which $g_{fs} = 0$ (error tolerance taken as 0.01). When using the bisection method, an algorithm was written that bases the $g_{fs} = 0$ starting point for future sample iterations on extrapolations from past root values, which considerably reduces iteration effort for future samples.

Results are given in Table 1, where the exact solution is taken from the result of $1 \times 10^6$ MCS samples, which required approximately 430 total CPU hours shared among multiple processors. Here, direct simulation could conduct approximately twice the number of calls as FS during the same CPU time, achieving similar results as FS for the low $\beta$ (1.24) case. For the high $\beta$ (3.31) case, MCS could not provide a solution for the CPU time given, and CE resulted in 20% error for the low-$\beta$ case, whereas FS gave a reasonable error between 2-3%. A full quadratic response surface would require much more computational effort than allocated for this problem (i.e. $2^n + 2n + 1$, or $2.14 \times 10^9$ samples with n=31 RVs), so a fractional design was used with 2111 samples, which resulted in just slightly more than the computational effort allocated to the other methods. Here, RS performed reasonably well but with higher error than FS.

**Example Problem 2. Dynamic Problem with Discontinuous Limit State**

This problem is taken from Eamon (2007). A 4 ft x 8 ft plywood panel on the corner of a hip roof is secured to five roof trusses with 33 nails as shown in Figure 4. The panel is subject to 4 independent dynamic wind loads, each described by a 3-minute pressure time-history. These time histories apply uplift and downwards pressure loads normal to the panel surface, with peak uplift pressures of approximately 25 psf. RVs are the withdrawal strengths of

three critical nails $n_1$-$n_3$ (it was found that these nails governed the panel failure), the pressure magnitude of the dynamic wind loads $W_1$-$W_4$, and panel dead load, for 8 RVs. RVs $N_i$ are independent and normally distributed with mean of 169 lb and COV of 0.40, while dead load is also normal with mean of 3.5 psf with COV of 0.10. Wind pressure RVs $W_j$ are lognormally distributed with COV of 0.41. The mean value of each $W_j$ is taken as 1.0, which represents the factor applied to the entire pressure time-history of the load area corresponding to $W_j$. The panel is modeled with plate elements while the nails are modeled with bar elements using MSC Nastran. For evaluation of the limit state, a transient dynamic analysis is conducted. If the withdrawal strength (axial force) of a nail is exceeded (the non-RV nails have a deterministic withdrawal strength of 169 lb), the analysis is stopped and the 'withdrawn' nail element is removed from the panel. The analysis is then restarted from the time in the load history that it was stopped and continues for the remaining duration of time or until all nails are withdrawn. Failure is defined as the event when all 33 nails are removed. The limit state is given by:

$$g = R(N_i)\ - Q(W_j) \tag{6}$$

Where R is the function of panel resistance, as a function of the nail withdrawal capacity RVs, while Q is the function of load effect on the panel, as a function of the wind load RVs. This is a difficult problem since the limit state response may be discontinuous, as it is a function of the sequence of discrete nail removals. Results are given in Table 2. The exact solution is taken from the result of $1 \times 10^5$ MCS samples, which required approximately 7200 total CPU hours shared among multiple processors. For FS the control variable was taken as $W_1$. For this problem, only the numerical integration approach for FS (NI) was considered, as the GLD could not be fit to the simulation data. Similarly, CE could not be applied here

due to computational difficulties with this method. Here FS produced an excellent result at 1% error, while MCS and RS produced significantly higher errors. The effect of partial correlation on the nail strength RVs was also considered ($\rho$=0.50). In this case, reliability index increased slightly but had no significant effect on the accuracy of the FS solution.


**Example Problem 3. Crush Tube**

An aluminum hollow tube 400 mm long with an 80 mm square cross-section representing a idealized vehicular side-rail (Rais-Rohani et al. 2006) is subjected to an initial velocity and impacts a rigid wall. The tube is modeled with shell elements as shown in Figure 5 in its failure state. A concentric point-mass applied at the end of the tube represents the relevant portion of the vehicular mass. Random variables are the density of the tube ($\rho_t$), point mass (M); wall thickness (T); initial velocity (V); elastic modulus (E); yield stress ($\sigma_y$); and post-yield modulus ($E_t$). Statistical parameters of these RVs are given in Table 3. Failure is defined as the event when the length-wise deformation of the tube (*D*) exceeds more than 50% (200 mm) of its original length at a time of 20 ms into the analysis. The limit state is given as


$$g = 200 - D(\rho_t, M, T, V, E, \sigma_y, E_t) \qquad (7)$$


The limit state is evaluated with the finite element code LS-DYNA (2003). Due to the computational cost of this problem, only 100 FS samples were considered. Initial velocity V was taken as the control variable. As each FS sample took approximately three times as long as an MCS sample, MCS and RS could achieve approximately 300 samples for the same CPU time. Results are shown in Table 4. The exact solution is taken from the result of $1 \times 10^5$ MCS samples, requiring approximately 4000 total CPU hours shared among multiple

processors. The MCS samples were plotted on probability paper to verify results. Here only FS was able to provide an accurate solution. As with the previous problem, only the numerical integration approach for FS (NI) was considered, as the GLD could not be fit to the simulation data.

## Effect of Reducing Sample Size

As noted earlier, the computational time for 1000 samples was somewhat arbitrarily chosen to compare solutions. The effect of reducing sample size on FS results is investigated by considering a previously studied problem. Table 5 gives results for example problem 1 for mean P = 40. As expected, sample size greatly affects the quality of the $p_f$ estimate, although significant reductions in sample size below 1000 can often be maintained with acceptable results. As shown in the table, less than 5% error was obtained between 50 and 250 samples this problem, for $\beta = 3.31$.

## Effect of control variable selection

Changing the control variable will create an equivalent but different limit state function. The FS solution procedure applied to these different limit states will create different distributions for R in $g_{fs}$, as the control variable is changed. As with CE, selecting an FS control variable with the most variability typically results in the best solutions, as that variance is removed from the formation of $f_R$ and $F_R$. To explore the effect of these differences on the FS solution, problem 3 was reconsidered with different selections of control variable. The results given in Table 6 show that there is less than a 2% difference in the final solution regardless of control variable choice, with error ranging from 1.4-3.1%  As expected, selecting a control variable with highest COV (RV T) results in the lowest error. For all

control variable choices, however, error tends to decrease toward the exact solution as the number of samples is increased. These results are typical.

Note that, for some oddly-formulated problems, it may not be possible to satisfy $g_{fs} = 0$, depending on the control variable selection and the sampled values of the remaining RVs. For this situation, FS (and CE) cannot be applied. It would be incorrect to disregard the non-conforming sample, as this would amount to eliminating the corresponding MCS samples, biasing the results, and likely producing a poor solution.

Another special case involves the situation where the chosen control variable has multiple roots in $g_{fs} = 0$. For this problem, a separate distribution for R would have to be generated for each root, and a separate failure probability calculated for each case. The solution would then be taken as the mean failure probability of the cases.

**Conclusion**

A reliability analysis method was presented that is a modification of conditional expectation, and the results of three representative example problems were shown for comparison. A variety of additional problems were considered for validation, with variations in failure probability, level of variance, distribution types, degree of non-linearity, and resistance RV correlation, with similar results (Charumas 2008). For the problems considered, FS produced reasonable, relatively efficient, and in most cases, very accurate solutions.

As FS represents complex limit state functions with an algebraically linear, two random variable problem, the complexity of the original model is discarded and the success of the

solution depends on the accuracy of the estimation of $R(X_i)$. Other considerations when using FS were given in the section above.

As the available advanced simulation methods may offer different advantages for particular types of problems, the authors would recommend that the various alternate solution possibilities be considered when approaching complex and computationally demanding reliability problems.

**Nomenclature**

| | |
|---|---|
| CDF | cumulative distribution function |
| CE | Conditional Expectation |
| $f_g$ | PDF of g; also, PDF of the GLD (eq 4) |
| $F_Q$ | CDF of  Q |
| FORM | First Order Reliability Method |
| $f_Q$ | PDF of Q |
| $F_R$ | CDF of $R(X_i)$ |
| $f_R$ | PDF of $R(X_i)$ |
| FS | Failure Sampling |
| g | limit state function |
| g' | CE limit state function |
| $g_{fs}$ | FS limit state function |
| GLD | Generalized Lambda Distribution |
| MCS | Monte Carlo simulation |
| MPP | most probable point of failure |
| NI | numerical integration |
| PDF | probability density function |
| $p_f$ | failure probability |
| Q | control RV |
| q | a specific value of Q |
| $R(X_i)$ | resistance function |
| R | $R(X_i)$ represented as a single equivalent resistance RV |
| RS | quadratic response surface |
| RV | random variable |
| s | simulation number |
| VRT | variance reduction technique |
| $X_i$ | set of RVs |
| $x_i$ | a set of specific values for RVs $X_i$ |
| $\beta$ | reliability index |
| $\lambda_i$ | GLD parameter |
| $\Phi$ | standard normal CDF |

## References

Au, S.K., Ching, J., and Beck, J.L., "Application of subset simulation methods to reliability benchmark problems." Structural Safety, Vol. 29, No. 3, pp. 183-193, 2007.

Au, S.K. and Beck, J.L. "Estimation of small failure probabilities in high dimensions by subset simulation." Probabilistic Engineering Mechanics, Vol. 16, No. 4, pp. 263-277, 2001.

Ayyub, B.M. and Chia, C-Y. "Generalized conditional expectation for structural reliability assessment." Structural Safety, Vol. 11, No. 2, pp.131-146, 1992.

Ayyub, B.M., and Haldar, A. "Practical Structural Reliability Techniques." Journal of Structural Engineering, ASCE, Vol. 110, No. 8, pp. 1707-1724, 1984.

Benjamin, J.R. and Cornell, C.A. *Probability, Statistics and Decisions for Civil Engineers*, McGraw-Hill, New York, 1970.

Breitung, K. "Asymptotic Approximations for Multinormal Integrals." Journal of Engineering Mechanics, ASCE, Vol. 110, No. 3, pp 357-366, 1984.

Bucher, C.G. and Bourgund, U. "Fast and efficient response surface approach for structural reliability problems." Structural Safety, v 7, n 1, pp. 57-66, 1990.

Charumas, B. "A New Technique for Structural Reliability Analysis: The Full Failure Sampling." Report Submitted to The Center for Advanced Vehicular Studies, Mississippi State University, August 2008.

Charumas, B. "A New Technique for Structural Reliability Analysis." Masters Degree Thesis, Department of Civil and Environmental Engineering, Mississippi State University, May 2008.

Chen, X. and Lind, N.C. "Fast Probability Integration by Three-Parameter Normal Tail Approximation." Structural Safety, Vol. 1, pp. 269-276, 1983.

Cheng, J. and Li, Q.S. "Application of the response surface methods to solve inverse reliability problems with implicit response functions." Computational Mechanics, v 43, n 4, pp. 451-459, 2009.

Der Kiureghian, A., Lin, H.Z., and Hwang, S.J. "Second-Order Reliability Approximations," Journal of Engineering Mechanics, ASCE, Vol. 113, No. 8, pp. 1208-1225, 1987.

Eamon, C. "Reliability of Concrete Masonry Unit Walls Subjected to Blast Threats," ASCE Journal of Structural Engineering, July 2007, No. 7, pp. 935-944 .

Eamon, C. "Reliability of Wood Roofs Exposed to Hurricane Winds," Durability of Wood-Framed Housing: Lessons Learned from Natural Disasters. Forest Products Society Conference, Biloxi, MS, November 2007.

Eamon, C., Thompson, M., and Liu, Z. "Evaluation of Accuracy and Efficiency of some Simulation and Sampling Methods in Structural Reliability Analysis," Journal of Structural Safety, Vol. 27 No. 4, pp. 356-392, 2005.

Engelund, S. and Rackwitz, R. "A benchmark study on importance sampling techniques in structural reliability, Structural Safety, Vol. 12, No. 4, pp. 255-276, 1993.

Fiessler, B., Neumann, H.J., and Rackwitz, R. "Quadratic Limit States in Structural Reliability." Journal of Engineering Mechanics, ASCE, Vol. 1095, No. 4, pp. 661-676, 1979.

Gomes, H. M. and Awruch, A. M. "Comparison of response surface and neural network with other methods for structural reliability analysis." Structural Safety, v 26, n 1, pp. 49-67, 2004.

Helmut, M. and Asif, L. "Estimating the Parameters of the Generalized Lambda Distribution", ALGO Research Quarterly, Vol. 3, No.3, December, 2000.

Hasofer, A.M. and Lind, N.C., "Exact and Invariant Second Moment Code Format." Journal of the Engineering Mechanics Division, ASCE, Vol. 100, No. EM1, pp. 111-121, 1974.

Hohenbichler, M., Gollwitzer, S., Kruse, W., and Rackwitz, R. "New Light on First- and Second-Order Reliability Methods." Structural Safety, Vol. 4, pp. 267-284, 1987.

Huang, B., and Du, X. "Uncertainty Analysis by Dimension Reduction Integration and Saddlepoint Approximations." Journal of Mechanical Design, ASME, Vol. 126, No. 1, pp. 26-33, 2006.

Iman, R.L. and Conover, W.J. "A distribution-free approach to inducing rank correlation among input variables." Communications in Statistics Vol. 11, No. 3, pp 311-334, 1982.

Iman, R.L., and Conover, W.J. "Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment." Communications in Statistics: Theory and Methods. Vol. A9 No. 17, pp. 1749-1842, 1980.

Karamchandani, A., Bjerager, P., and Cornell, A.C. "Adaptive Importance Sampling." Proceedings, International Conference on Structural Safety and Reliability (ICOSSAR), San Francisco, CA., pp. 855-862, 1989.

Karamchandani, A. and Cornell, C. "Adaptive hybrid conditional expectation approaches for reliability estimation." Structural Safety, Vol. 11, No.1, pp. 59-74, 1991.

Karian, A. Zaven and Dudewicz, J. Edward. *Fitting Statistical Distributions: The generalized lambda distribution and generalized bootstrap methods*, CRC Press, 2000.

LS-DYNA Keyword User's Manual, Livermore Software Technology Corporation, April 2003.

Mori, Y. and Ellingwood, B.R. "Time-dependent system reliability analysis by adaptive importance sampling," Structural Safety, Vol. 12, No. 1, pp. 59-73, 1993.

MSC/NASTRAN Users Guide. MacNeal-Schwendler Corporation, Los Angeles, CA. 2005

Ozaturk, A. and R. F. Dale. "A study of Fitting the Generalized Lambda Distribution to Solar Radiation Data." American Meteorological Society, July, 1982.

Rackwitz, R. and Fiessler, B. "Structural Reliability Under Combined Random Load Sequences." Computers and Structures, Vol. 9, No. 5, pp. 484-494, 1978.

Rais-Rohani, M, Solanki, K, and Eamon, C. "Reliability-Based Optimization of Lightweight Automotive Structures for Crashworthiness," 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Portsmouth, VA, September 2006.

Rubinstein, R.Y. *Simulation and the Monte Carlo Method*, John Wiley & Sons, New York, 1981.

Smarslok, B., Haftka, R.T., and Kim, N.H. "Taking Advantage of Separable Limit States in Sampling Procedures," 8th AIAA Non-deterministic App. Conference, 2006, pp. 366-378

Tvedt, L. "Distribution of Quadratic Forms in Normal Space--Application to Structural Reliability." Journal of Engineering Mechanics, ASCE, Vol. 116, No. 6, pp. 1183-1197, 1990.

Vanderplaats Research and Development. *VisualDOC Design Optimization Software Version 6.1 Theoretical Manual.* Colorado Springs, 2008.

Yasuhiro, M. and Ellingwood, B.R. "Time-dependent System Reliability Analysis by Adaptive Importance Sampling." Structural Safety, Vol. 12, No. 1, pp. 59-73, 1993.

Wu, Y.T. "An Adaptive Importance Sampling Method for Structural Systems Analysis, Reliability Technology 1992." Edited by T.A. Cruse, ASME Winter Annual Meeting, Vol. AD 28. Anaheim, CA, pp. 217-231, 1992.

Wu, Y.T., and Wirsching, P.H. "New Algorithm for Structural Reliability Estimation." Journal of Engineering Mechanics, ASCE, Vol. 113, No. 9, pp 1319-1336, 1987.

**Table 1. Problem 1 (Truss) Results**

| method | nominal no. of calls | mean P = 55 | | | mean P = 40 | | |
|---|---|---|---|---|---|---|---|
| | | β | %err | COV | β | %err | COV |
| Exact Solution | | 1.24 | -- | -- | 3.31 | -- | -- |
| FS (NI) | 1000 | 1.19 | -2.7 | 0.014 | 3.39 | 2.3 | 0.0068 |
| FS (GLD) | 1000 | 1.21 | -2.5 | 0.025 | 3.42 | 3.4 | 0.002 |
| MCS | 2000 | 1.20 | -3.7 | 0.062 | n.f. | -- | -- |
| CE | 1000 | 0.99 | 20 | 0.023 | 3.38 | 2.2 | 0.010 |
| RS | 2111* | 1.29 | 4.2 | -- | 3.47 | 4.8 | -- |

*Fractional factorial design used


**Table 2. Problem 2 (Dynamic Roof Panel) Results**

| method | nominal no. of calls | β | %err | COV |
|---|---|---|---|---|
| Exact Solution | | 2.85 | -- | -- |
| FS (NI) | 1000 | 2.83 | 1.0 | 0.051 |
| MCS | 5000 | 3.09 | 5.1 | 0.45 |
| RS | 273* | 3.98 | 39 | -- |

*maximum, as controlled by number of RVs.


**Table 3. Problem 3 RV Statistical Parameters**

| RV | Mean | COV | Distribution |
|---|---|---|---|
| $\rho$ | $2.70 \times 10^{-6}$ kg/mm$^3$ | 0.05 | normal |
| $M$ | 800 kg | 0.15 | normal |
| $T$ | 4 mm | 0.01 | normal |
| $V$ | 9.5 m/s | 0.10 | normal |
| $E$ | 69 GPa | 0.10 | lognormal |
| $\sigma_y$ | 0.175 GPa | 0.15 | normal |
| $E_t$ | 0.265 GPa | 0.15 | normal |


**Table 4. Problem 3 (Crush Tube) Results**

| method | nominal no. of calls | mean V = 11 | | | mean V = 9.5 | | |
|---|---|---|---|---|---|---|---|
| | | β | %err | COV | β | %err | COV |
| Exact Solution | | 2.28 | -- | -- | 3.89 | -- | -- |
| FS (NI) | 100 | 2.23 | -2.3 | 0.036 | 3.89 | 0.04 | 0.027 |
| MCS | 300 | 2.11 | 7.1 | 0.43 | n.f. | -- | -- |
| CE | 100 | 3.22 | 41 | 0.092 | 4.66 | 20 | 0.040 |
| RS | 143* | 2.02 | -11 | -- | 3.42 | -12 | -- |

*maximum, as controlled by number of RVs.

**Table 5. Effect of Reduced Sample Size on FS Results for Example Problem 1.**

|  | no. of calls | β | %err | COV |
|---|---|---|---|---|
| Exact solution |  | 3.31 | -- | -- |
| FS | 25 | fail | **--** | **--** |
|  | 50 | 4.39 | 33 | 0.050 |
|  | 250 | 3.42 | 3.4 | 0.028 |
|  | 500 | 3.40 | 2.5 | 0.011 |
|  | 1000 | 3.39 | 2.3 | 0.007 |

**Table 6. Comparison of the Effect of Selected Control Variable on FS Results**

| control variable | β | % err | COV |
|---|---|---|---|
| T | 2.31 | 1.4 | 0.027 |
| V | 2.23 | -2.3 | 0.036 |
| $\sigma_y$ | 2.35 | 3.1 | 0.041 |
| M | 2.34 | 2.6 | 0.035 |
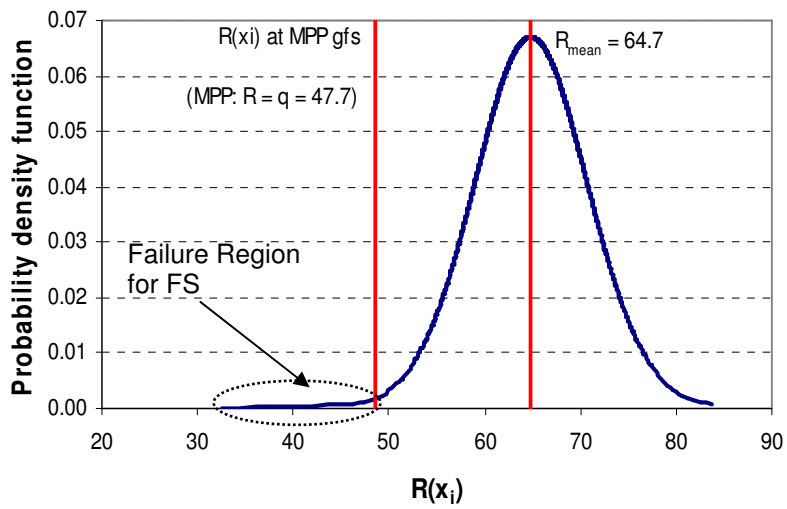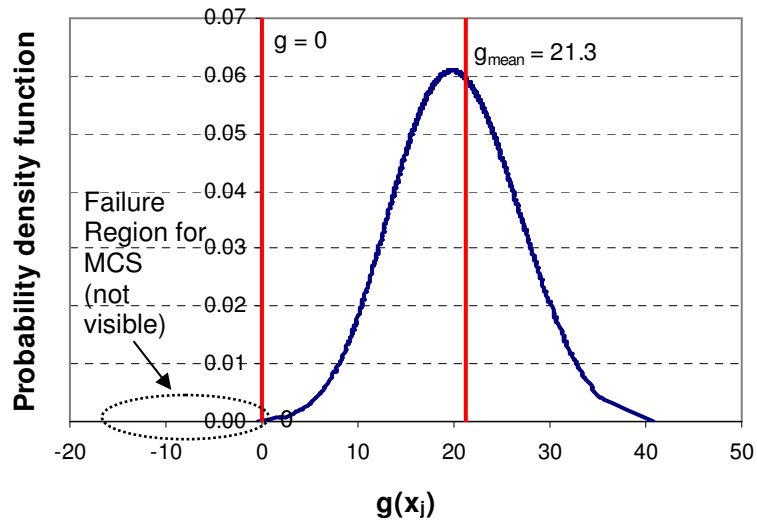
**Figure 1. PDFs of g and R($x_i$) for Example Problem 1, stress limit ($\beta$ = 3.31, mean P = 40).**
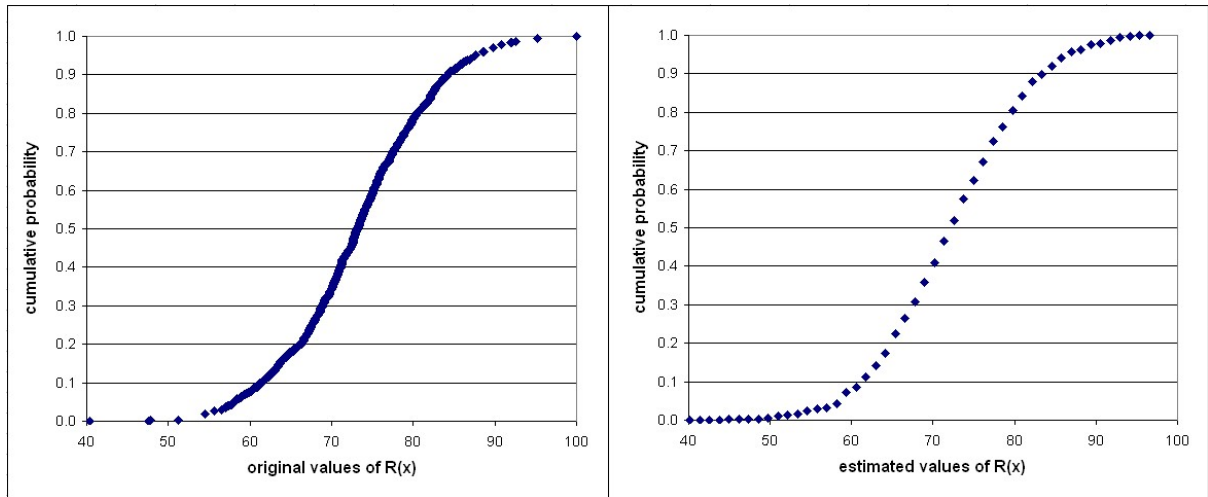
**CDF of original samples (left) and CDF estimate used by FS (right) for Example Problem 1.**
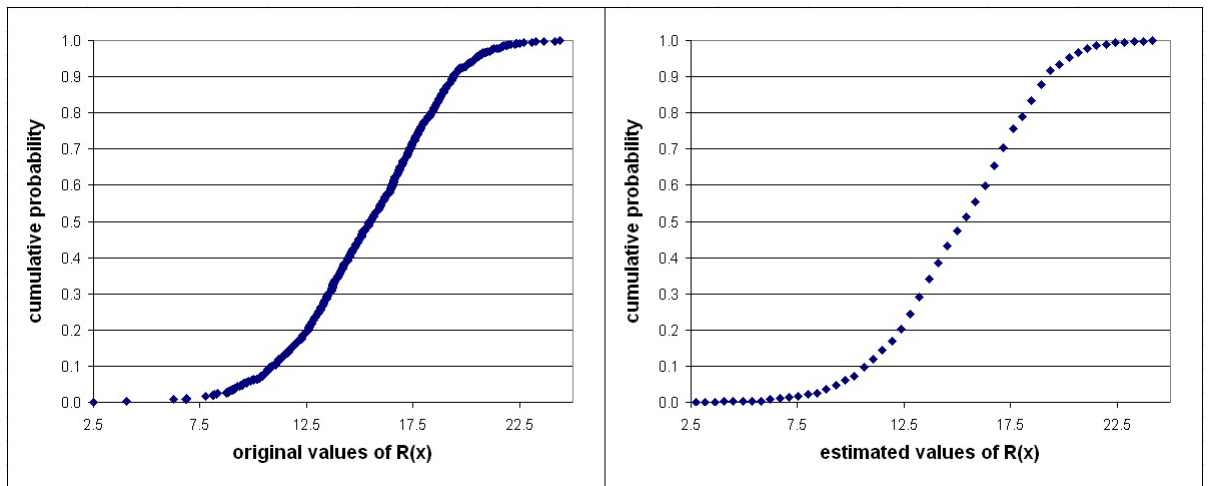


**CDF of original samples (left) and CDF estimate used by FS (right) for Example Problem 3.**

**Figure 2. Example CDFs of original samples (left) and CDF estimates used by FS (right).**
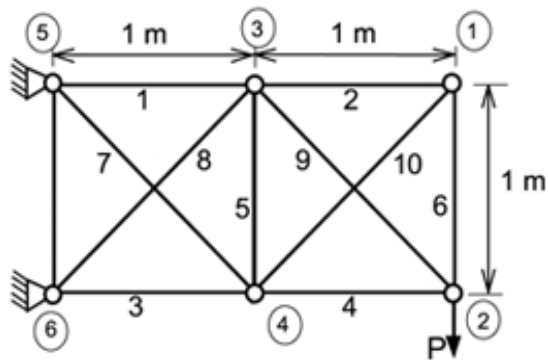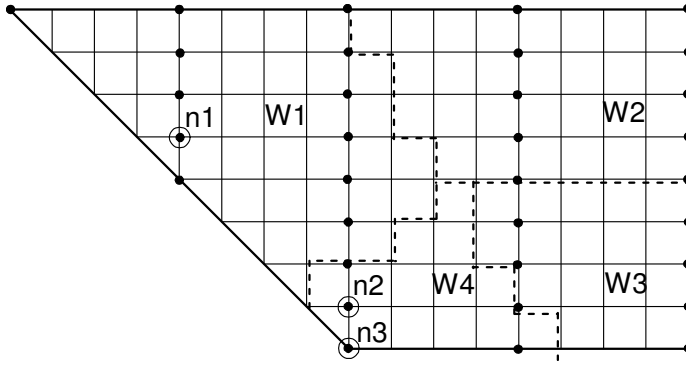


**Figure 3. 10-bar Cantilever Truss**

**Figure 4. Plan view of Dynamically-Loaded Panel.** The grid is the FEA mesh of plate elements, while dots represent the location of nails. Dotted lines separate wind load areas.
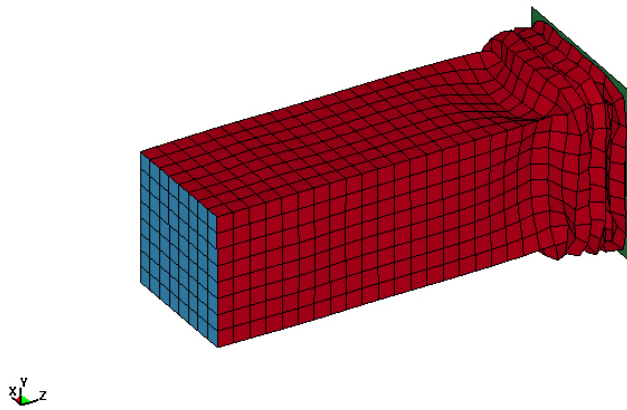


**Figure 5. Crush Tube.** Figure of the typical deformed shape at t=20 ms. The end of the hollow tube is covered with a plate.