

*International Journal of Reliability and Safety*  
*Volume 1, Issue 1-2*  
*DOI: 10.1504/IJRS.2006.010698*

## **A Saddlepoint Approximation Based Simulation Method for Uncertainty Analysis**

Beiqing Huang  
Graduate Research Assistant  
Department of Mechanical and Aerospace Engineering  
University of Missouri – Rolla, USA

Dr. Xiaoping Du<sup>\*</sup>, Assistant Professor  
Department of Mechanical and Aerospace Engineering  
University of Missouri – Rolla, USA  
Rolla, MO 65409  
Email: dux@umr.edu

Ramaprasad Lakshminarayana  
Graduate Research Assistant  
Department of Mechanical and Aerospace Engineering  
University of Missouri – Rolla, USA

**\*Corresponding Author**  
Department of Mechanical and Aerospace Engineering  
University of Missouri – Rolla  
1870 Miner Circle  
Rolla, Missouri 65409 – 4494  
USA  
Phone: 1-573-341-7249  
Fax: 1-573-341 - 4115  
E-mail: dux@umr.edu

January 2006

## **Abstract**

Uncertainty analysis, which assesses the impact of uncertainty of input random variables on performance functions, is an important and indispensable component in engineering design under uncertainty. In this paper, a Saddlepoint Approximation based simulation method is proposed to accurately and efficiently estimate the distribution of a response variable. The proposed method combines both simulation and analytical techniques and involves three main steps: (1) sampling on input random variables, (2) approximating the cumulant generating function of the response variable with its first four cumulants, and (3) estimating the cumulative distribution function and probability density function of the response variable using Saddlepoint Approximation. This method provides more computationally efficient solutions than the general Monte Carlo simulation while maintaining high accuracy. The effectiveness of the proposed method is illustrated with a mathematical example and two engineering analysis problems.

**Keywords:** uncertainty analysis; Saddlepoint Approximation; reliability analysis; Monte Carlo simulation; cumulant generating function

## 1 Introduction

The reliability of any engineering system is directly affected by various uncertainties, such as variations in loading, material properties, physical dimensions of parts, model structure, and operation conditions. With ever increasing demand of higher product reliability as well as the pressure of cost reduction, the effects of uncertainty on reliability must be carefully addressed. With the aid of uncertainty analysis in design stage, reliability issues can be addressed early to prevent the occurrence of failure events that may lead to significant quality losses or even catastrophic consequences. While the use of uncertainty analysis has gained significant adoption in industrial practices, to date, a major obstacle for performing all-inclusive uncertainty analysis is the unaffordable computational burden. The evaluation of the probabilistic characteristics of system performances can present severe mathematical and numerical difficulties. A full-scale uncertainty analysis may require hundreds or thousands of deterministic engineering analyses where expensive simulations, such as finite element analysis (FEA) and computational fluid dynamics, are usually involved.

Suppose a performance function is modeled as

$$Y = g(\mathbf{X}), \quad (1)$$

where  $\mathbf{X}$  is the vector of input random variables,  $\mathbf{X} = [X_1, X_2, \dots, X_d]^T$ , and  $Y$  is the response (performance) variable. In the area of structural reliability, the function  $g(\mathbf{X})$  is also referred to as a limit-state function. The main task of uncertainty analysis is to obtain the cumulative probability function (*cdf*) and probability density function (*pdf*) of the response  $Y$ , given the distribution of input variables  $\mathbf{X}$ . Theoretically, the *cdf* of  $Y$  can be calculated by a multidimensional integral,

$$F_Y(y) = P\{Y \leq y\} = \int_{g(\mathbf{X}) \leq y} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (2)$$

where  $f_{\mathbf{X}}(\mathbf{x})$  is the joint *pdf* of random variables  $\mathbf{X}$ . A close-form solution to Eq. (2) rarely exists due to the high dimensionality, the complicated integrand  $f_{\mathbf{X}}(\mathbf{X})$ , and the nonlinear integration boundary  $g(\mathbf{X}) = y$ .

Many practitioners adopt approximation methods such as the First Order Reliability Method (FORM) and the Second Order Reliability Method (SORM). These approximation methods rely on the approximation of the model at the so-called Most Probable Point (MPP) (Haldar and Mahadevan, 2000; Hohenbichler, Gollwitzer, Kruse and Rackwitz, 1987). The search of the MPP involves an optimization problem and therefore needs an iterative process where a number of deterministic analyses are required. When the number of random variables is small or moderate and when the performance function is sufficiently approximated with a low order polynomial (i.e. the first and second order in FORM and SORM, respectively), MPP-based approaches are computationally efficient and can provide satisfactorily accurate results. *However, their computational efficiency decreases significantly when the number of random variables increases.* The reason is that the typical MPP search requires numerical evaluations of derivatives of the response, and the number of such evaluations is approximately proportional to the number of random variables (Adhikari and Langley, 2002; Wu, 1998). In addition, due to the first or second order approximation, neither FORM nor SORM may be accurate enough for a highly nonlinear response. To improve accuracy, Tvedt (1988) used Saddlepoint Approximation as an alternative to SORM. After the MPP is identified and the performance function is approximated in a quadratic form in terms of standard normal variables, the cumulant generating function (*cgf*) of the performance is readily available, and then the Saddlepoint Approximation is

applied to calculate the reliability. Compared to SORM, the Saddlepoint Approximation method provides more accurate results. However, the method requires the MPP search and needs the second order derivatives at the MPP in order to obtain the quadratic approximation. When the number of random variables is large, the computation is still intensive. Furthermore, none of these MPP-based methods (Breitung, 1984; Du and Chen, 2000; Du and Chen, 2001; Du, Sudjianto and Chen 2004; Hohenbichler, Gollwitzer, Kruse and Rackwitz, 1987; Rackwitz, and Fiessler, 1978; Tvedt, 1988) is suitable for distribution estimation (*cdf* and *pdf* curve generation). Other uncertainty analysis methods (Isukapalli, Roy and Georgopoulos, 1998; Jin, Du and Chen, 2003; Zou, Mourelatos, Mahadevan and Tu, 2003) are based on Design of Experiments (DOE). Computationally cheap surrogate models are created from DOE to replace the original models during uncertainty analysis. The indicator response surface-based simulation method (Zou, Mourelatos, Mahadevan and Tu, 2003) is one representative method, which can be used to estimate reliability at both component and system levels efficiently and accurately when the dimension of the problems is not high; it can also deal with the multiple MPPs where FORM or SORM is not applicable. DOE-based methods require a large number of function evaluations to construct surrogate models when the number of random variables is large. A typical approach to improve efficiency for large scale problems is the use of variable reduction methods (Adhikari and Langley, 2002; Mohanty and Wu, 2001; Wu, 1998). According to the sensitivities of random variables to the performance response, “unimportant” variables are removed from the uncertainty analysis. Although this method improves efficiency significantly, screening out variables and neglecting the interaction between “removed variables” and “remaining variables” may result in an unknown accuracy loss (Wu, 1998).

Point-estimate method (Seo and Kwak, 2002; Zhao, Alfredo and Ang, 2003) is another uncertainty analysis method which uses approximated statistical moments to estimate the reliability. The computational demand (the number of performance function evaluations) is equal to  $s^d$ , where  $s$  is the number of points for each variable and  $d$  is the number of random variables. It is not appropriate for a large scale problem because the computational demand increases exponentially with the number of random variables.

Monte Carlo simulation (MCS) is easy and flexible to use and does not exhibit the limitations shared by analytical and surrogate methods. However, directly applying MCS is usually not a practical alternative because running computer simulation is an expensive exercise and the probability of failure is typically very small for a high reliability product. Some variants of MCS (Dey and Mahadevan, 1998; Moarefzadeh and Melchers, 1999; Rubinstein, 1981), such as importance sampling and adaptive importance sampling, have been developed to address rare events with small probability. Several other methods have also been proposed to improve the efficiency of the basic MCS. McKay et al. (1979) introduced the use of Latin Hypercube Sampling (LHS) for computer experiments. Owen (1997) showed that, for finite samples, LHS is not worse than MCS.

Recently, the First Order Saddlepoint Approximation (FOSPA) method has been reported (Du and Sudjianto, 2004). It is more accurate and efficient than FORM and in some cases more accurate than the general SORM and the saddlepoint-based SORM (Tvedt, 1988). However, FOSPA is still expensive when the problem dimension is high, because the MPP search is required.

Considering the above challenges, there is a need to develop general, accurate and efficient uncertainty analysis methods which are especially suitable for large scale problems. In

this paper, taking the advantage of the accuracy of Saddlepoint Approximation (SPA) and the dimension-independent feature of MCS-type methods, we propose a Saddlepoint Approximation Based Simulation (SABS) method. This method provides a solution to the tradeoff between accuracy and efficiency. It inherits the good features from both MCS and SPA and gives an accurate estimate for the cumulative distribution function (*cdf*) and probability density function (*pdf*) of a response while only needing a moderate number of simulations. Consequently, it is especially suitable and beneficial to uncertainty analysis for large scale problems that involve a large number of random variables.

The remaining of the paper is organized as follows. Section 2 gives an overview of Saddlepoint Approximation, which serves as a theoretical foundation of the SABS method. In Section 3, the details of the SABS method are presented. In Section 4, a mathematical example and two engineering problems are used to illustrate the SABS method in comparison with FORM, SORM, and MCS. Section 5 is the closure of this paper.

## 2 Saddlepoint Approximation

Saddlepoint Approximation has become a powerful tool to estimate the *pdf* and *cdf* since it was first introduced by Daniels (1954). During the last two decades, research in this area has vastly increased (Goutis and Casella, 1999). Next, we will give a brief introduction to Saddlepoint Approximation.

Let  $Y$  (a response variable) be a random variable with a *pdf*  $f_Y(y)$ , and then the moment generating function of  $Y$  is given by

$$\phi(t) = \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy. \quad (3)$$

The cumulant generating function (*cgf*)  $K(t)$  of  $Y$  is defined as

$$K(t) = \log [\phi(t)], \quad (4)$$

where  $\log$  is the natural logarithm.

The *pdf* of  $Y$  can be restored from  $K(t)$  by the inverse Fourier transformation

$$f_Y(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \phi(it) dt = \frac{1}{2\pi} \int_{-i\infty}^{i\infty} e^{[K(t)-ty]} dt, \quad (5)$$

where  $i = \sqrt{-1}$  (the pure imaginary quantity).

Differentiating the integrand in Eq. (5) and letting the result equal zero yields the following equation

$$K'(t) = y, \quad (6)$$

where  $K'(\cdot)$  is the first order derivate of the *cgf*.

The solution to Eq. (6) is called the *saddlepoint* and is denoted by  $t_s$ . Then the integrand is approximated at the saddlepoint and an integration path passing through the saddlepoint of the integrand is selected. Since the saddlepoint is an extreme point as shown in Eq. (6), the function value of the integrand falls away rapidly as we move from this point. Thus, the influence of neighboring points on the integral is diminished (Goutis and Casella, 1999; Huzurbazar, 1999).

Based on the above idea, Daniels (1954) used the exponential power series expansion to estimate the integral in Eq. (5) and derived the following equation for *pdf* estimation,

$$f_Y(y) = \left[ \frac{1}{2\pi K''(t_s)} \right]^{\frac{1}{2}} e^{[K(t_s)-t_s y]}, \quad (7)$$

where  $K''(\cdot)$  is the second order derivative of the *cgf*. Lugannani and Rice (1980) provided a concise formula for calculating *cdf*,

$$F_Y(y) = P\{Y \leq y\} = \Phi(w) + \phi(w) \left( \frac{1}{w} - \frac{1}{v} \right), \quad (8)$$



where  $\Phi(\cdot)$  and  $\phi(\cdot)$  are the *cdf* and *pdf* of a standard normal distribution, respectively,

$$w = \text{sgn}(t_s) \left\{ 2 [t_s y - K(t_s)] \right\}^{1/2} \quad (9)$$

and

$$v = t_s \left[ K''(t_s) \right]^{1/2}, \quad (10)$$

where  $\text{sgn}(t_s) = +1, -1, \text{ or } 0$ , depending on whether  $t_s$  is positive, negative or zero.

Briefly speaking, the central idea of SPA is that the probability integration is approximated through the saddlepoint where the integrand has the highest contribution to the integration. For the complete methodology, interested readers can refer to Huzurbazar (1999), Jensen (1995), and Lugannani and Rice (1980).

Saddlepoint Approximation has several excellent features. (1) It yields extremely accurate probability estimation, especially in the tail area of a distribution where probability calculation for high reliability system is needed (Daniel, 1987; Jensen, 1995). (2) It requires only a process of finding one saddlepoint without any integration. 3) It provides the estimations of both *cdf* and *pdf* simultaneously. Hence, there is no need of taking numerical derivative of *cdf* to obtain *pdf* or performing numerical integration of *pdf* to obtain *cdf*.

Even though Saddlepoint Approximation has widespread applications in statistics (Daniel, 1987; Kolassa, 1991; Kuonen, 2001), its applications in engineering design and simulation has not been well studied (Du and Sudjianto, 2004). The integration of Saddlepoint Approximation with SORM (Tvedt, 1988) has demonstrated its potential use in engineering fields. A recent attempt is the First Order SaddlePoint Approximation (FOSPA) for reliability analysis (Du and Sudjianto, 2004). With FOSPA, the performance function is linearized at the MPP in the *original space* without any nonlinear transformation from non-normal random variables to standard normal random variables. Therefore, the chance of increasing the nonlinearity of the

performance function is eliminated. However, the search of the MPP is an iterative process and needs a number of performance function evaluations. Because of the computational cost, FOSPA may not be applicable for a large scale problem. To alleviate the computational demand, this research work develops an efficient Saddlepoint Approximation based simulation method by taking advantages of the desirable features of Saddlepoint Approximation and Monte Carlo simulation. Details of the proposed method are given in the following section.

### **3 Saddlepoint Approximation Based Simulation**

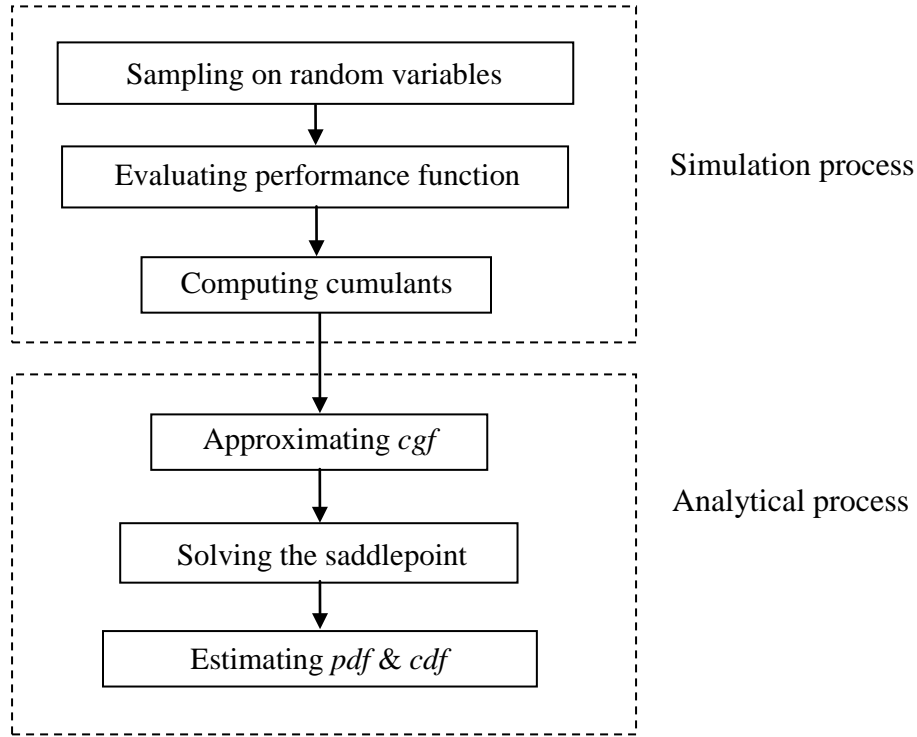
As discussed previously, the first order Saddlepoint Approximation method is accurate but inefficient when the number of random variables is high. On the other hand, the accuracy of Monte Carlo simulation (MCS) is independent of the problem dimension, but MCS is inefficient for rare events. The proposed Saddlepoint Approximation based simulation (SABS) method combines both methods and provides accurate solutions with moderate computational cost.

#### **3.1 General procedure**

The SABS method consists of simulation and analytical processes (simulation and Saddlepoint Approximation). As discussed in the preceding section, the use of Saddlepoint Approximation rests on the *cgf* of the performance function  $Y = g(\mathbf{X})$ , which is usually a nonlinear function of the random variables  $\mathbf{X}$ . Except for special cases, a close-form *cgf* of  $Y = g(\mathbf{X})$  cannot be obtained analytically. In this work, a sampling based method is used to estimate the first four statistical cumulants. The *cgf* of the performance function is then approximated with the empirical cumulants. Finally, Saddlepoint Approximation is applied to estimate the *pdf* and *cdf* of the performance function. Figure 1 outlines the procedure involved in

the SABS method, and the details of each step of the procedure are presented in the following subsections.

**Figure 1** Procedure of the SABS method



### 3.2 Sampling on random variables and evaluating performance function

Random variable sampling requires the generation of  $n$  samples of a random vector consisting of all of the  $d$  random variables,  $\mathbf{X} = [X_1, X_2, \dots, X_d]^T$ . If variable  $X_j$  follows a distribution  $F_{X_j}$ , the samples are generated using Quantile-Quantile transformation from *i.i.d* samples,  $\mathbf{v}$ , uniformly distributed on  $u[0, 1)^d$ ,

$$x_j = F_{X_j}^{-1}(v_j). \quad (11)$$

McKay et al. (1979) introduced the use of Latin Hypercube Sampling (LHS) for computer experiments where the samples are drawn by

$$v_j^i = \frac{\pi_j(i-1) + w_j^i}{n} \quad (i=1, \dots, n), \quad (12)$$

where  $\pi_j$  is a uniform permutation of  $0, \dots, n-1$ , and  $w_j^i$  is a random observation from  $u[0,1)$ .

$v_j^i$  can also be obtained by the widely used median version

$$v_j^i = \frac{\pi_j(i-1) + 0.5}{n}. \quad (13)$$

Because LHS exhaustively stratifies across the whole range of each sampled variable, it provides an efficient way for sampling the entire range of each variable in accordance with the assumed probability distribution. Thus, it mitigates the problem of random sampling for which important intervals with low probability but high consequences may be missing from the samples. LHS requires fewer samples than the Monte Carlo sampling with the same accuracy and produces more stable results (Ding, Zhou and Liu, 1998; Helton and Davis, 2003; Owen, 1997).

After  $n$  samples of  $\mathbf{X}$  are drawn, the corresponding  $n$  samples of the performance function  $Y$  are then obtained through substituting the samples of  $\mathbf{X}$  into the performance function

$$y^i = g(x_1^i, x_2^i, \dots, x_d^i) \quad (i=1, \dots, n), \quad (14)$$

where  $(x_1^i, x_2^i, \dots, x_d^i)$  is the  $i$ th sample ( $\mathbf{x}^i$ ) of  $\mathbf{X}$ .

### 3.3 Estimating cumulants

Cumulants,  $\kappa_1, \kappa_2, \dots, \kappa_r$ , of a random variable are theoretically derived from the power expansion of its *cgf* (Kendall and Stuart, 1958)

$$K(t) = \kappa_1 t + \frac{\kappa_2 t^2}{2!} + \dots + \frac{\kappa_r t^r}{r!} + \dots \quad (15)$$

The  $r$ th cumulant  $\kappa_r$  ( $r = 1, 2, \dots$ ) is generated by

$$\kappa_r = \left. \frac{d^r}{dt^r} [K(t)] \right|_{t=0}. \quad (16)$$

Cumulants have been used to estimate *pdf*. The representative method is the cumulant-neglect closure for dynamics problems which involve stochastic differential equations (Wojtkiewicz, Spencer and Bergman, 1996; Wu and Lin, 1984). In this work, we use cumulants of the response variable to estimate its *cgf*. We first estimate the cumulants and then use the cumulants to approximate the *cgf* of the performance function as shown in Eq. (15).

Once a set of samples of the performance  $y^i$  ( $i = 1, 2, \dots, n$ ) are generated, the cumulants of the performance function are computed using the following equations (Fisher, 1928):

$$\left\{ \begin{array}{l} \kappa_1 = \frac{s_1}{n} \\ \kappa_2 = \frac{ns_2 - s_1^2}{n(n-1)} \\ \kappa_3 = \frac{2s_1^3 - 3ns_1s_2 + n^2s_3}{n(n-1)(n-2)} \\ \kappa_4 = \frac{-6s_1^4 + 12ns_1^2s_2 - 3n(n-1)s_2^2 - 4n(n+1)s_1s_3 + n^2(n+1)s_4}{n(n-1)(n-2)(n-3)}, \end{array} \right. \quad (17)$$

where  $n$  is the sample size;  $s_r$  ( $r = 1, 2, 3$  and  $4$ ) are the sums of the  $r$ th powers of the sample values of the performance function, namely,

$$s_r = \sum_{i=1}^n (y^i)^r. \quad (18)$$

For higher order of cumulants, one can refer to Fisher (1928).

### 3.4 Approximating cumulant generating function

As suggested in Eq. (15), the *cgf* can be expressed with an infinite series of powers of  $t$  (in an interval of  $[-h, h]$  for some small  $h > 0$ ):

$$K(t) = \sum_{j=1}^{\infty} \kappa_j \frac{t^j}{j!}. \quad (19)$$

An approximation formula could be obtained by constructing a series up to a certain order. In practice, only the first four cumulants in the above power expansion are used, and then the Saddlepoint Approximation technique is applied. This approximation can yield remarkable results (Canty and Davison, 1996; Gatto and Ronchetti, 1996; Wang, 1992). In this work, up to the fourth cumulant items,  $\kappa_1, \kappa_2, \kappa_3$  and  $\kappa_4$ , are used to approximate the *cgf*.

### 3.5 Evaluating *pdf* and *cdf*

Once an analytical formulation of the *cgf* for the performance function is obtained, the use of Saddlepoint Approximation becomes straightforward. The analytical formulation of the *cgf* is in a polynomial form in terms of variable  $t$ , and it is easy to obtain an analytical solution of the saddlepoint and the derivatives of the *cgf*. According to Eqs. (6) and (15), the equation for the saddlepoint is derived as

$$K'(t) = \kappa_1 + \sum_{j=2}^r \kappa_j \frac{t^{j-1}}{(j-1)!} = y. \quad (20)$$

Solving the above equation yields the saddlepoint  $t_s$ .

The *cgf* and its second order derivative at the saddlepoint  $t_s$  are then given by

$$K(t_s) = \sum_{i=1}^r \frac{\kappa_i t_s^i}{i!} \quad (21)$$

and

$$K^n(t_s) = \kappa_2 + \sum_{j=3}^r \kappa_j \frac{t_s^{j-2}}{(j-2)!}, \quad (22)$$

respectively.

Thereafter, *pdf* and *cdf* are then calculated by Eqs. (7) and (8), respectively.

## 4 Examples

In this section, a mathematical example and two engineering problems are used to illustrate our proposed method (SABS). Up to the fourth cumulant items are used to solve the saddlepoint. The performance of SABS is compared with that of MCS, FORM, and SORM, whenever necessary. For FORM and SORM, the sequential quadratic programming is employed to search the MPP.

### 4.1 A mathematical problem

Consider a performance function with  $d$  independent random variables (Du and Sudjianto, 2004) given by

$$Y = g(\mathbf{X}) = \frac{\sum_{j=1}^d X_j - d}{\sqrt{d}}, \quad (23)$$

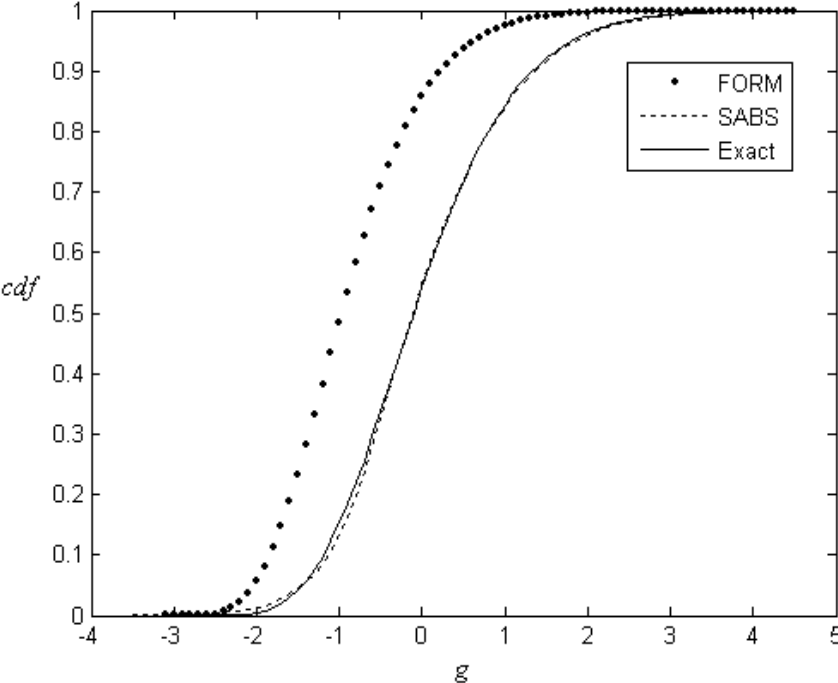
where  $d$  is the number of random variables, and  $X_j$  are random variables following the standard exponential distributions with *pdf*

$$f_{X_j}(x_j) = e^{-x_j}. \quad (24)$$

This example is selected because a theoretical solution exists. FORM, SORM, and SABS are used to evaluate the *cdf* and *pdf* of the performance function in a range of  $[-3.5, 4.5]$ . The results from FORM, SABS and exact solutions are depicted in Figures 2 and 3 for  $d = 10$ . It is

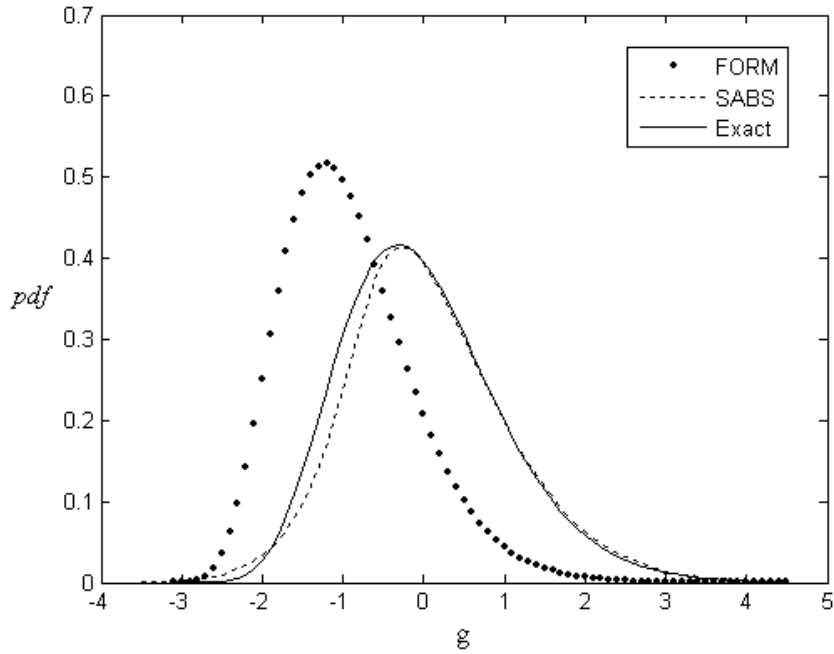
noted that the proposed method is much accurate than FORM while only 250 samples are used. SORM gives better solutions than FORM in the most of the range, but they are still much less accurate than SABS. Singularity occurs in the left tail of the distribution when SORM is used, and therefore the results from SORM are not included in the figures.

**Figure 2** *cdf* when  $d = 10$ , the sample size of SABS  $n = 250$



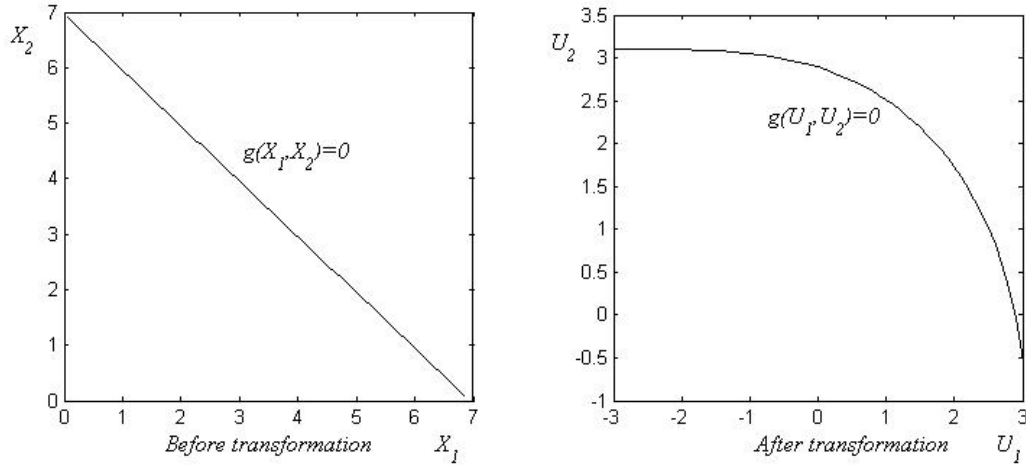


**Figure 3** *pdf* when  $d = 10$ , the sample size of SABS  $n = 250$



As indicated in Du and Sudjianto (2004), the reason that FORM and SORM are not accurate is due to the nonlinear transformation from a non-normal distribution to a normal distribution. Note that the performance function in Eq. (23) is linear in the original random variable space. However, after the random variable transformation from an exponential distribution into a standard normal distribution, the performance function becomes highly nonlinear. Both FORM and SORM fail to accommodate such nonlinearity. Figure 4 demonstrates this situation when  $d = 2$ .

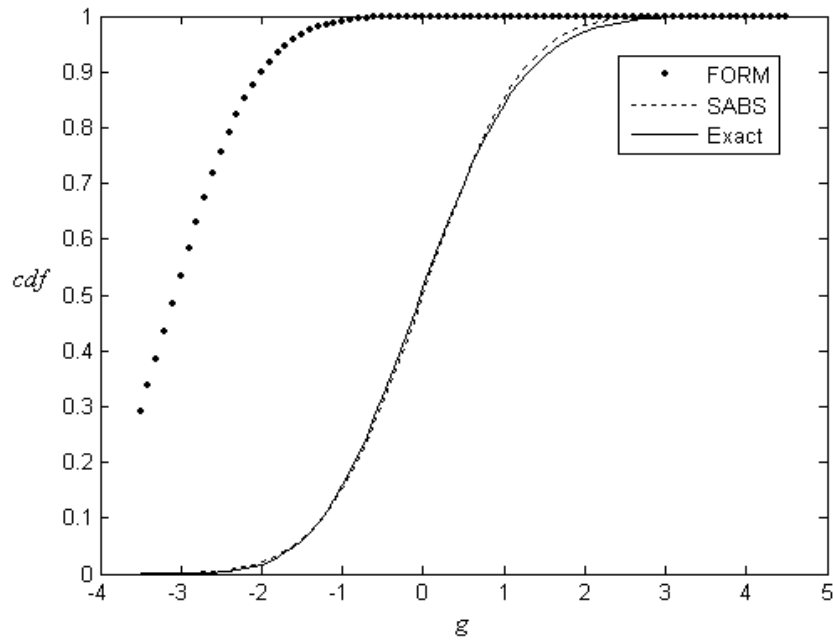
**Figure 4** Performance function before and after the transformation



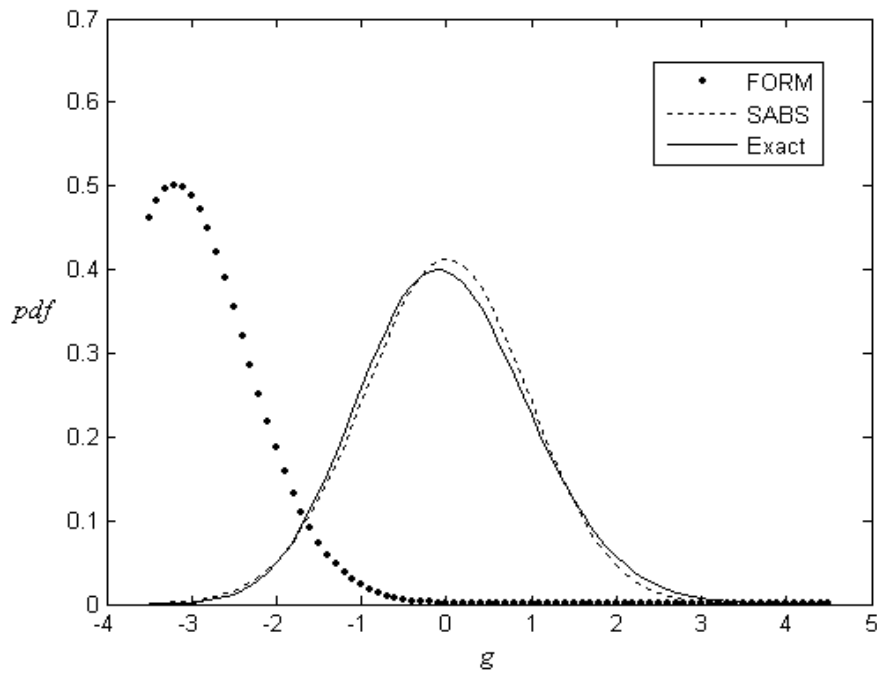
We further investigate a higher dimension with  $d = 100$ . The *cdf* and *pdf* of the performance function for this large scale case are plotted in Figures 5 and 6, respectively. It is noted that the results from SABS and the exact solution are very close. The error of FORM is unacceptably large. At  $g = -1.0$ , the exact *cdf* is less than 20%, but FORM gives a *cdf* more than 95%. Only 250 samples are used in SABS for this large scale problem. Figures 2, 3, 5 and 6 demonstrate that SABS is evenly accurate over the whole range  $[-3.5, 4.5]$ .

Note that this problem involves a summation of a large number of random variables. According to the central limit theorem, when the number of random variable is large, the distribution of the performance variable approaches a normal distribution. The *pdf* curves from the exact solution and SABS are in agreement with the central limit theorem while this is not the case with FORM due to its poor approximation of the performance function in the transformed **U**-space.

**Figure 5**  $cdf$  when  $d = 100$ , the sample size of SABS  $n = 250$



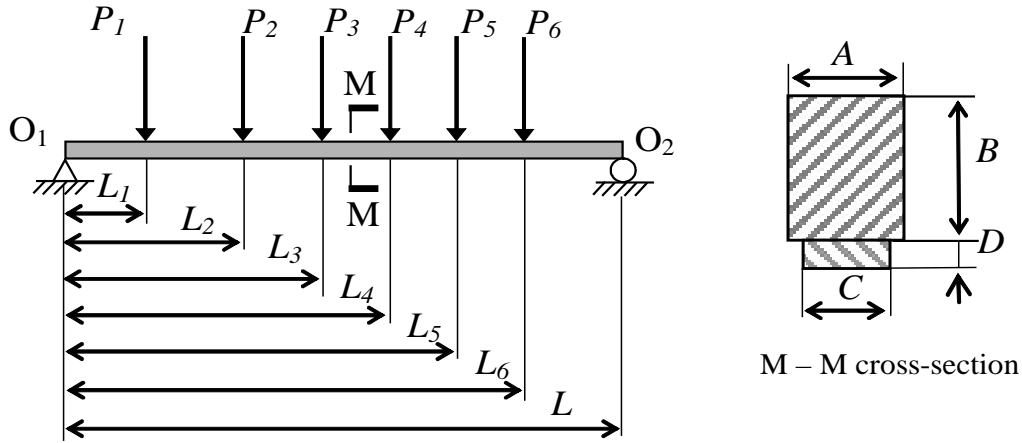
**Figure 6**  $pdf$  when  $d = 100$ , the sample size of SABS  $n = 250$



## 4.2 A composite beam problem

In what follows, we present an example where the engineering model has an analytical form. Consider a composite beam with 20 independent random variables. The beam with Young's modulus  $E_w$  and  $A$  mm wide by  $B$  mm deep by  $L$  mm long, has an aluminum plate with Young's modulus  $E_a$  and a net section  $C$  mm wide by  $D$  mm high securely fastened to its bottom face, as shown in Figure 7. Six external vertical forces,  $P_1, P_2, P_3, P_4, P_5$  and  $P_6$  are applied at six different locations along the beam,  $L_1, L_2, L_3, L_4, L_5$ , and  $L_6$  from the left end. The allowable tensile strength is  $S$ .

**Figure 7** A composite beam



In this problem, the twenty random variables are

$$\mathbf{X} = [X_1, \dots, X_{20}]^T = [A, B, C, D, L_1, L_2, L_3, L_4, L_5, L_6, L, P_1, P_2, P_3, P_4, P_5, P_6, E_a, E_w, S]^T.$$

Details of these random variables are given in Table 1.

The maximum stress occurs in the middle cross-section M-M and is then given by

$$\sigma_{\max} = \frac{\left[ \frac{\sum_{i=1}^6 P_i(L-L_i)}{L} L_3 - P_1(L_2-L_1) - P_2(L_3-L_2) \right] \left[ \frac{0.5AB^2 + \frac{E_a}{E_w} DC(B+D)}{AB + \frac{E_a}{E_w} DC} \right]}{\frac{1}{12} AB^3 + AB \left\{ \left[ \frac{0.5AB^2 + \frac{E_a}{E_w} DC(B+D)}{AB + \frac{E_a}{E_w} DC} \right]^2 - 0.5B \right\} + \frac{1}{12} \frac{E_a}{E_w} CD^3 + \frac{E_a}{E_w} CD \left\{ 0.5D + B - \left[ \frac{0.5AB^2 + \frac{E_a}{E_w} DC(B+D)}{AB + \frac{E_a}{E_w} DC} \right]^2 \right\}}. \quad (25)$$

The maximum stress  $\sigma_{\max}$  should be less than the allowable strength  $S$ . The performance function of the beam is defined by

$$Y = g(\mathbf{X}) = S - \sigma_{\max}. \quad (26)$$

**Table 1** Random variables of the beam reliability problem

Variable No.	Variable	Parameter 1*	Parameter 2 <sup>+</sup>	Distribution type
1	$A$	100 mm	0.2 mm	Normal
2	$B$	200 mm	0.2 mm	Normal
3	$C$	80 mm	0.2 mm	Normal
4	$D$	20 mm	0.2 mm	Normal
5	$L_1$	200 mm	1 mm	Normal
6	$L_2$	400 mm	1 mm	Normal
7	$L_3$	600 mm	1 mm	Normal
8	$L_4$	800 mm	1 mm	Normal
9	$L_5$	1000 mm	1 mm	Normal
10	$L_6$	1200 mm	1 mm	Normal
11	$L$	1400 mm	2 mm	Normal
12	$P_1$	12 kN	18 kN	Uniform
13	$P_2$	12 kN	18 kN	Uniform
14	$P_3$	12 kN	18 kN	Uniform
15	$P_4$	12 kN	18 kN	Uniform
16	$P_5$	12 kN	18 kN	Uniform
17	$P_6$	12 kN	18 kN	Uniform
18	$E_a$	70 GPa	7 GPa	Normal
19	$E_w$	8.75 GPa	0.875 GPa	Normal
20	$S$	20 MPa	2.0 MPa	Normal

\*Parameter 1 is the mean for normal distribution and is the lower bound for uniform distribution

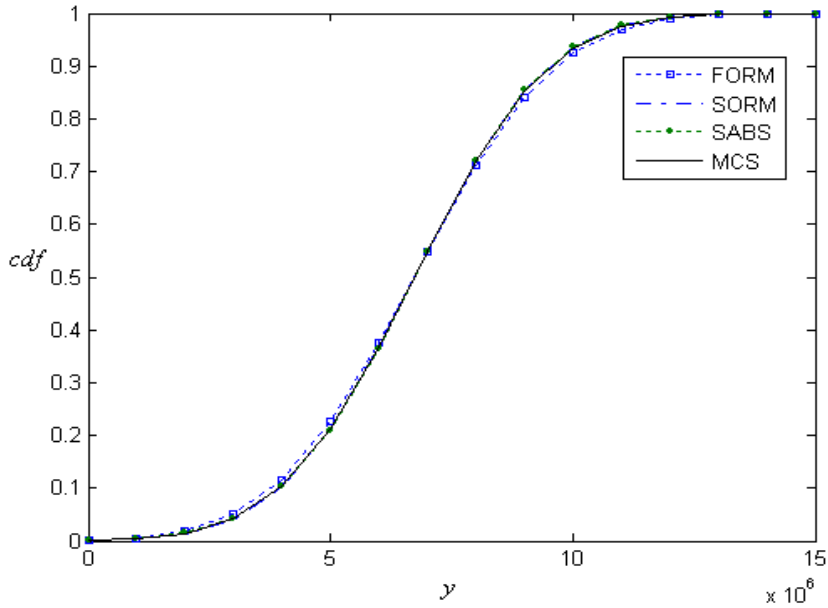
<sup>+</sup>Parameter 2 is the standard deviation for normal distribution and is the upper bound for uniform distribution.

The estimated *cdf* of  $Y$  and computational cost from each method are given in Table 2. The sample sizes of SABS and MCS are 500 and 1,000,000, respectively. FORM and SORM require the MPP search at each realization value of  $Y$  listed in the first column of Table 2. The numbers of function evaluations used by FORM and SORM are provided in the parentheses. The result of MCS is considered accurate because of the large number of runs. It is noted that both SORM and SABS provide accurate results compared to that of MCS. SORM is more accurate than FORM, and SABS is the most efficient method since it consumes only 500 function evaluations for the entire distribution generation. FORM and SORM are *extremely* inefficient since they need to search the MPP at each point of the performance function. The total numbers of function evaluations of FORM and SORM are 16,890 and 20,670, respectively. The probability results are plotted in Figure 8.

**Table 2** The estimated probability of  $Y \leq y$  (MPa) and computational cost

$y$ (MPa)	FORM		SORM		SABS	MCS
0.0	0.0014	(882)	0.0008	(1134)	0.0010	0.0009
1.0	0.0058	(1220)	0.0036	(1472)	0.0042	0.0039
2.0	0.0188	(1102)	0.0131	(1354)	0.0147	0.0142
3.0	0.0511	(885)	0.0392	(1137)	0.0424	0.0424
4.0	0.1164	(1672)	0.1008	(1924)	0.1026	0.1033
5.0	0.2253	(768)	0.2113	(1020)	0.2097	0.2120
6.0	0.3755	(1302)	0.3711	(1554)	0.3650	0.3678
7.0	0.5473	(1216)	0.5474	(1468)	0.5477	0.5493
8.0	0.7106	(1342)	0.7197	(1594)	0.7217	0.7215
9.0	0.8397	(1210)	0.8555	(1462)	0.8553	0.8534
10.0	0.9241	(1185)	0.9374	(1437)	0.9375	0.9352
11.0	0.9697	(1114)	0.9780	(1366)	0.9779	0.9761
12.0	0.9899	(999)	0.9936	(1251)	0.9937	0.9928
13.0	0.9972	(853)	0.9985	(1105)	0.9986	0.9982
14.0	0.9994	(1140)	0.9997	(1392)	0.9997	0.9996

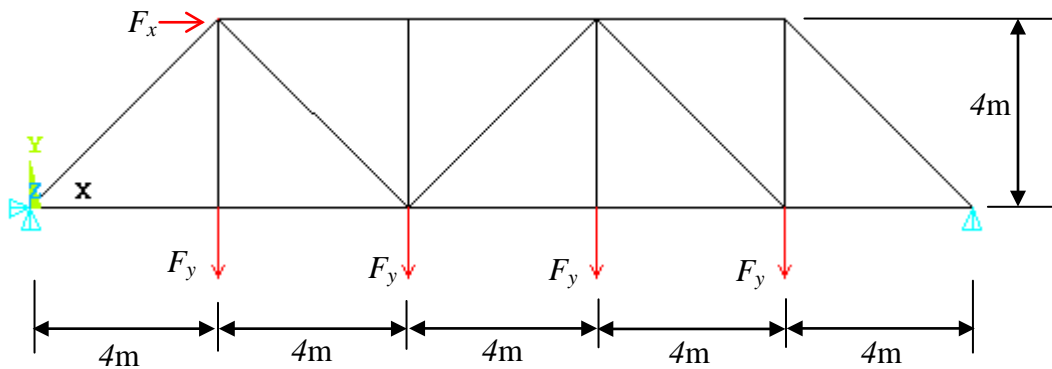
**Figure 8** The *cdf* of  $Y$  for the beam problem



### 4.3 A steel truss bridge

In the previous example, the performance function is in an analytical form. Next, we will demonstrate the use of SABS for a performance function that involves a back-box model, which is obtained from finite element analysis (FEA). A truss bridge with applied constraints and loads is shown in Figure 9.

**Figure 9** A steel truss bridge with constraints and loads

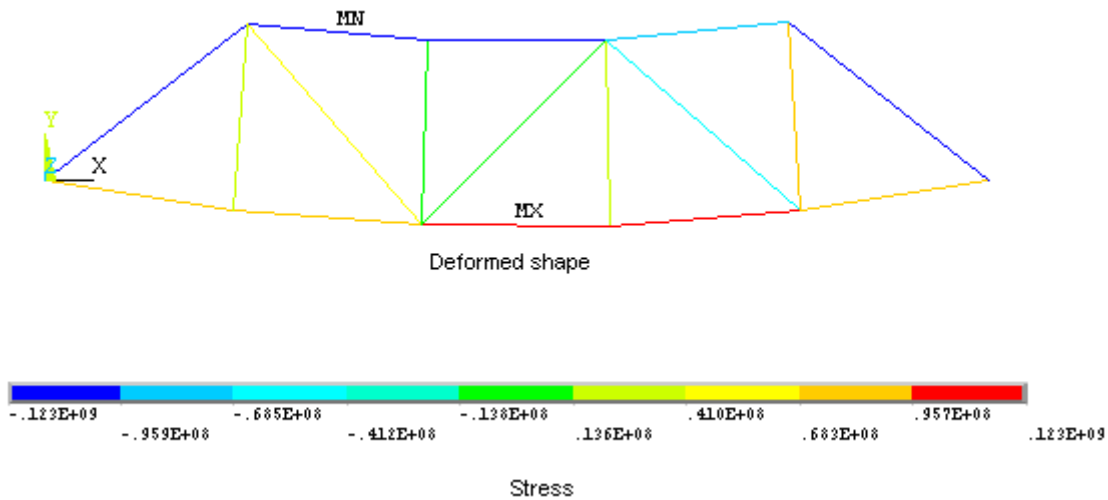


There is one fixed support in the horizontal and vertical directions at the bottom left end, and one fixed support in the vertical direction at the bottom right end. The total span of the bridge is 20 m. Four external forces  $F_y$  are applied at the four locations (with 4 m distance in between) to represent the dead loads throughout the deck. A horizontal force  $F_x$  is applied at the top left corner to represent the wind load acting on the bridge. The height of the bridge is 4 m. Five random variables are  $F_x$ ,  $F_y$ , the truss' cross-sectional area  $A$ , Young's modulus  $E$ , and the material's strength  $S$ . Their distributions are given in Table 3. Figure 10 illustrates the deformed shape and stress plot from one ANSYS FEA simulation.

**Table 3** Random variables of the steel truss bridge problem

Variable	Mean	Standard deviation	Distribution type
$F_x$	1000 N	100 N	Normal
$F_y$	45000 N	4500 N	Normal
$A$	0.0011 m <sup>2</sup>	0.00011 m <sup>2</sup>	Normal
$E$	200 GPa	20 GPa	Normal
$S$	200 MPa	20 MPa	Normal

**Figure 10** The deformed shape and stress plot of one FEA simulation





The performance function of the truss bridge is given by

$$Y = S - \sigma_{\max}, \quad (26)$$

where  $S$  is the strength of the material and  $\sigma_{\max}$  is the maximum tensile stress in the trusses that is calculated from FEA simulation.

The probability of failure is defined as the probability that the strength is less than the maximum stress, i.e.  $p_f = P\{Y \leq 0\}$ . Table 4 shows the estimated probability of failure  $p_f$  and the computational cost by SABS and MCS. It is seen that the result of SABS is close to that of MCS. SABS is much more efficient than MCS since it only uses 500 simulations while MCS uses 10,000 simulations. The finite element analyses are performed with ANSYS 9.0 on an HP workstation x4000 with Intel(R) Xeon(TM) CPU 1.70GHz and Microsoft Windows XP professional operating system. SABS takes 14 minutes and 39 seconds for the 500 simulations while MCS takes 4 hours 50 minutes and 57 seconds for the 10,000 simulations.

**Table 4** Probability of  $Y \leq 0$  (MPa) and computational costs by SABS and MCS

Method	$P\{Y \leq 0\}$	$n$ (No. of simulations)
SABS	0.0039	500
MCS	0.004	10,000

## 5 Conclusion

The proposed Saddlepoint Approximation based simulation method is an attractive alternative approach to uncertainty analysis, especially for large scale problems. It incorporates both simulation and analytical techniques. The simulation is performed for generating the *cgf* of a performance function, and the analytical technique is used for approximating the *cdf* and *pdf*. With a moderate sample size, SABS can produce accurate results for which a large number of

Monte Carlo simulations are needed. This feature especially benefits the large scale problems with high reliability. If FORM or SORM is used for the same large scale problem, much more computational effort is required because the computational demand is proportional to the dimension. Since SABS estimates both *cdf* and *pdf* directly, there is no need of taking derivative of *cdf* to get *pdf* or taking integration of *pdf* to acquire *cdf*. In addition, SABS does not have the limitations shared by FORM and SORM, such as the requirement of the differentiability of the performance function and the need of a unique MPP. It should be pointed out that, due to the use of finite cumulants (first four cumulants), SABS may not be applicable for multimodal distributions.

The error of SABS comes from the errors of moment estimation, *cgf* approximation, and Saddlepoint Approximation. The statistical error of moment estimation can be established from the general MCS. The error of the estimation of Saddlepoint Approximation can be obtained analytically (Goutis and Casella, 1999). The error of the approximation of the cumulant generating function is similar to that of Taylor series expansion. The overall error estimation needs a further investigation and will be our future work.

### **Acknowledgements**

The support from the U.S. National Science Foundation grant DMI – 040081 and the University of Missouri – Rolla Intelligent Systems Center is gratefully acknowledged.

## References

- Adhikari, S. and Langley, R.S. (2002) 'Reduction of random variables in structural reliability analysis', *3rd international conference on mathematical methods in reliability methodology and practice*, Trondheim, Norway.
- Breitung, K. (1984) 'Asymptotic approximations for multinomial integrals', *Journal of engineering Mechanics*, Vol. 110, No. 3, pp.357-366.
- Canty, A.J. and Davison, A.C. (1996) 'Implementation of saddlepoint approximations to bootstrap distribution', *Proceedings of the 28th Symposium on the Interface, Computer Science and Statistics*, pp.248-253.
- Daniels, H.E. (1954) 'Saddlepoint approximations in statistics', *Annals of Mathematical Statistics*, Vol. 25, pp.631-650.
- Daniel, H.E. (1987) 'Tail probability approximations', *International Statistical Review*, Vol. 55, No. 1, pp.37-48.
- Dey, A. and Mahadevan, S. (1998) 'Ductile structural system reliability analysis using adaptive importance sampling', *Structural Safety*, Vol. 20, No. 2, pp.137-154.
- Ding, K., Zhou, Z. and Liu, C. (1998) 'Latin hypercube sampling used in the calculation of the fracture probability', *Reliability Engineering & system safety*, Vol. 59, No. 2, pp.239-242.
- Du, X. and Chen, W. (2000) 'Towards a better understanding of modeling feasibility robustness in engineering design', *ASME Journal of Mechanical Design*, Vol. 122, No. 4, pp.385-394.
- Du, X. and Chen, W. (2001) 'A most probable point based method for uncertainty analysis', *Journal of Design and Manufacturing Automation*, Vol. 4, No. 1, pp.47-66.

- Du, X., Sudjianto, A. and Chen, W. (2004) 'An integrated framework for optimization under uncertainty using inverse reliability strategy', *ASME Journal of Mechanical Design*, Vol. 126, No. 4, pp.562-570.
- Du, X. and Sudjianto, A. (2004) 'The first order saddlepoint approximation for reliability analysis', *AIAA Journal*, Vol. 42, No. 6, pp.1199-1207.
- Fisher, R.A. (1928) 'Moments and product moments of sampling distribution', *Proceeding of London Mathematical Society*, Vol. 30, No. 2, pp.199-238.
- Gatto, R. and Ronchetti, E. (1996) 'General saddlepoint approximations of marginal densities and tail probabilities', *Journal of the American Statistical Association*, Vol. 91, No. 433, pp.666-673.
- Goutis, C. and Casella, G. (1999) 'Explaining the saddlepoint approximation', *The American Statistician*, Vol. 53, No. 3, pp.216-224.
- Haldar, A. and Mahadevan, S. (2000) *Probability, reliability, and statistical methods in engineering design*, John Wiley and Sons.
- Helton, J.L. and Davis, F.J. (2003) 'Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems', *Reliability Engineering & system safety*, Vol. 81, pp.23-69.
- Hohenbichler, M., Gollwitzer, S., Kruse, W. and Rackwitz, R. (1987) 'New light on first- and second-order reliability methods', *Structural Safety*, Vol. 4, pp.267-284.
- Huzurbazar, S. (1999) 'Practical saddlepoint approximations', *The American Statistician*, Vol. 53, No. 3, pp.225-232.

- Isukapalli, S.S., Roy, A. and Georgopoulos, P.G. (1998) 'Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological system', *Risk Analysis*, Vol. 18, No. 3, pp.351-363.
- Jensen, J.L. (1995) *Saddlepoint approximations*. Clarendon Press, Oxford.
- Jin, R., Du, X. and Chen, W. (2003) 'The use of metamodeling techniques for design under uncertainty', *Structural and Multidisciplinary Optimization*, Vol. 25, pp.99-116.
- Kendall, M.G. and Stuart, A. (1958) *The Advanced Theory of Statistics Volume 1: Distribution Theory*, London: Charles Griffin & Company Ltd.
- Kolassa, J.E. (1991) 'Saddlepoint approximations in the case of intractable cumulant generating functions', *Selected Proceedings of the Sheffield Symposium on Applied Probability*, Institute of Mathematical Statistics, Hayward, CA, USA, pp.236-255.
- Kuonen, D. (2001) *Computer-intensive statistical methods: saddlepoint approximations with applications in bootstrap and robust inference*, PhD thesis, Swiss Federal Institute of Technology.
- Lugannani, R. and Rice, S.O. (1980) 'Saddlepoint approximation for the distribution of the sum of independent random variables', *Advances in Applied Probability*, Vol. 12, pp.475-490.
- Mckay, M.D., Conover, W.J. and Beckman, R.J. (1979) 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code', *Technometrics*, Vol. 21, No. 2, pp.239-245.
- Moarefzadeh, M.R. and Melchers, R.E. (1999) 'Directional importance sampling for ill-proportioned spaces', *Structural Safety*, Vol. 21, No. 1, pp.1-22.
- Mohanty, S. and Wu, Y.-T. (2001) 'CDF sensitivity analysis technique for ranking influential parameters in the performance assessment of the proposed high-level waste repository at

- Yucca Mountain, Nevada, USA', *Reliability Engineering & System Safety*, Vol. 73, No. 2, pp.167-176.
- Owen, A. (1997) 'Monte Carlo variance of scrambled equidistribution quadrature', *SIAM Journal of Numerical Analysis*, Vol. 34, No. 5, pp.1884-1910.
- Rackwitz, R. and Fiessler, B. (1978) 'Structural reliability under combined random load sequences', *Computer and Structures*, Vol. 9, No. 5, pp.489-494.
- Rubinstein, R.Y. (1981) *Simulation and the Monte Carlo method*, New York: Wiley.
- Seo, H.S. and Kwak, B.M. (2002) 'Efficient statistical tolerance analysis for general distributions using three-point information', *International Journal of Production Research*, Vol. 40, No. 4, pp.931-944.
- Tvedt, L. (1988) 'Second order reliability by an exact integral', *Reliability and Optimization of Structural System'88; Proceedings of the 2<sup>nd</sup> IFIP WG7.5 Conference*, London, UK, pp.377-384.
- Wang, S. (1992) 'General saddlepoint approximations in the bootstrap', *Statistics & Probability Letters*, Vol. 13, pp.61-66.
- Wojtkiewicz, S.F., Spencer Jr, B.F. and Bergman, L.A. (1996) 'On the Cumulant-Neglect closure method in stochastic dynamics', *International Journal of Non-linear Mechanics*, Vol. 31, No. 5, pp. 657-684.
- Wu, Y-T. (1998) 'Methods for efficient probabilistic analysis of systems with large number of random variables', *7th AIAA/USAF/NASA/ISSMO symposium on multidisciplinary analysis and optimization*, St. Louis, MO, USA.

- Wu, W.F. and Lin, Y.K. (1984) 'Cumulant-Neglect closure for non-linear oscillators under random parametric and external excitations', *International Journal of Non-linear Mechanics*, Vol. 19, No. 4, pp.349-362.
- Zhao, Y.G., Alfredo, H.S. and Ang, H.M. (2003) 'System reliability assessment by method of moments', *ASCE Journal of structural engineering*, Vol. 129, No. 10, pp.1341-1349.
- Zou, Z., Mourelatos, Z., Mahadevan, S. and Tu, J. (2003) 'An indicator response surface-based Monte Carlo method for efficient component and system reliability analysis', *Proceedings of ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois, USA.