Failure Surface Frontier for Reliability Assessment on Expensive Performance Function

Songqing Shan
Dept. of Mech. & Manuf. Engineering
The University of Manitoba
Winnipeg, MB
Canada, R3T 5V6

G. Gary Wang*
Dept. of Mech. & Manuf. Engineering
The University of Manitoba
Winnipeg, MB, Canada, R3T 5V6
Tel: 204-474-9463 Fax: 204-275-7507
Email: gary_wang@umanitoba.ca

Abstract

This work proposes a novel concept of failure surface frontier (FSF), which is a hyper-surface consisting of the set of non-dominated failure points on the limit states of a failure region. Assumptions, definitions, and benefits of FSF are described first in detail. It is believed that FSF better represents the limit states for reliability assessment (RA) than conventional linear or quadratic approximations on the most probable point (MPP). Then, a discriminative sampling based algorithm is proposed to identify FSF, based on which the reliability can be directly assessed for expensive performance functions. Though an approximation model is employed to approximate the limit states, it is only used as a guide for sampling and a supplementary tool for RA. Test results on well-known problems show that FSF-based RA on expensive performance functions achieves high accuracy and efficiency, when compared with the state-of-the-art results archived in literature. Moreover, the concept of FSF and proposed RA algorithm are proved to be applicable to problems of multiple failure regions, multiple most probable points, or failure regions of extremely small probability.

Keywords: Failure surface frontier, reliability assessment, discriminative sampling, reliability analysis

1. Introduction

Conventionally, variation and uncertainties in materials, manufacturing, assembly, field usage, and so on are taken into consideration through safety factors, which are often determined through empirical data. In modern approaches, the statistical distribution of input variables and design performance are used to quantify the probability of achieving a specific performance target. This is the task of reliability assessment (RA). RA is the foundation for reliability engineering and reliability-based design optimization (RBDO) [1-5].

As a general formulation, consider an *n*-component system with design (variable) space $S_x \subseteq \Re^n$, where $x = (x_1, x_2, ..., x_n)$. Let the uncertain state of the system be denoted by a random vector $X = (X_1, X_2, ..., X_n)$ with a joint probability density function (JPDF) $j_x(x)$. The performance function, or state function, is defined as

$$f(x) = f(x_1, x_2, ..., x_n) \tag{1}$$

* Corresponding author

where $x = (x_1, x_2, ..., x_n) \in S_x \subseteq \Re^n$.

In reliability analysis, a limit state function, $g(x)$, is defined such that for a particular limit state value f_0 of the performance function,

$$g(x) = f(x) - f_0 = 0 \tag{2}$$

The limit state value, f_0, is equivalent to an engineering performance specification. This limit state function divides the variable space into the safe region for which $g(x) > 0$ and the failure region for which $g(x) \le 0$. Geometrically, the limit state Eq. (2) is an n-dimensional surface that is called the failure surface. One side of the failure surface is the safe region, whereas the other side is the failure region. Fig. 1 illustrates two limit state surfaces $g_1(x)=0$ and $g_2(x)=0$; the dark region is the failure region and the rest is the safe region, or success region. The reliability of a system or product, R, is defined as the rate of compliance to the performance specification and is given by

$$R = P(g(X) > 0) = \int_{g(x)>0} ... \int j_x(x) dx \tag{3}$$

Eq. (3) is simply the volume integral of $j_x(x)$ over the safe region. Conversely the failure probability of a system or product, P_f, is defined as the rate of non-compliance to the performance specification or the complimentary event of reliability and is given by

$$P_f = P(g(X) \le 0) = \int_{g(x)\le 0} ... \int j_x(x) dx \tag{4}$$

Eq. (4) is the volume integral of $j_x(x)$ over the failure region. By definition $R + P_f = 1$. The quantitative evaluation of the true R or P_f often poses two major difficulties. One is the determination of the correct form of $j_x(x)$, which is often unavailable or difficult to obtain in practice because of insufficient data [6]. The other is the determination of the limit state surface $g(x) = 0$, which separates the failure region and safe region. When the computation of the performance function is an expensive function such as FEA, such a difficulty aggravates. Even if both $j_x(x)$ and $g(x)$ are given, the volume integral from Eq. (3) or Eq. (4) maybe difficult to compute. Therefore, direct calculation of reliability R or failure probability P_f from formula Eq. (3) or Eq. (4) may be impractical.

(Insert Fig. 1 here.)

Current reliability assessment methods include Monte Carlo Simulation (MCS) methods and variations, Most Probable Point (MPP)-based approximation methods, and emerging discriminative sampling based methods. A typical MCS method involves three steps: (1) repetitive sampling from the set of random variables according to their respective probability distributions, (2) obtaining the corresponding performance function values to the samples, and (3) identifying whether the failure has occurred [7]. The estimated probability of failure is then simply the number of failures, m_f, divided by the total number of simulations, m, or $P_f = m_f/m$, when m is sufficiently large. MCS is accurate and reliable for all types of problems. Therefore, its results are often used as a standard to test other methods [8]. However, MCS is computationally intensive because it requires a large m. For implicit and computationally intensive, or expensive, performance functions, the computation burden becomes unbearable. Some variants of MCS methods have been developed that can reduce the computation effort by about one order-of-magnitude. These methods include the Importance Sampling (IS),

Adaptive Importance Sampling (AIS), Latin Hypercube Sampling, and Directional Simulation methods [8-12]. The total number of sample points required by MCS methods, however, is still too large and thus the application of these methods is limited mostly to inexpensive performance functions.

The MPP-based approximation methods originate from the concept of "most probable point." MPP was from the phrase "most probable failure point" coded in [13]. Because of the difficulties to accurately determine the failure surface that separates the safe and failure region and to obtain the volume integral, a reliability index based on a linear approximation to the limit state function is thus used as a simplified solution for reliability assessment [14]. The reliability index is defined as the shortest distance from the origin to the failure surface in the transformed normal design space. The point on the failure surface that is closest to the origin is then referred as the MPP. The idea of using reliability index is equivalent to using a tangent line at MPP to represent the failure surface, which is also referred as First Order Reliability Method (FORM). In FORM, the distance between MPP, u^*, and the origin in the normalized U-space, $\beta = \| u^* \|$, specifies the reliability R or the complementary P_f.

$$P_f = \Phi(-\beta) = \Phi(-\| u^* \|) \tag{5}$$

where $\Phi(.)$ denotes the standard normal cumulative distribution function. In essence, FORM uses the tangent line as shown in Fig. 2 to approximate the nonlinear $g(u)=0$ surface. As one can see from Fig. 2, for the complex nonlinear $g(u)=0$ function(s), the approximation could induce a large error. For the function underneath the tangent line, if FORM is used, the actual safe region would be much smaller than estimated, which might result in a risky design. For the function above the tangent line, FORM will give a very conservative estimation. Various FORM-based methods are studied and analyzed in a recent work [15].

(Insert Fig. 2 here.)

If a quadratic curve is used for approximation, then the Second Order Reliability Method (SORM) and its variations are developed. There are a number of variations of SORM by introducing different quadratic terms. Wu et al. [16] defined the linear and quadratic Taylor's expansions of $g(x)=0$ in the X-space and referred them as mean-value based first-order and second-order methods, respectively. The definitions in the X-space are very useful when the transformation from the non-normal space to normal space causes extra curvatures to the limit state function, which has been observed in many engineering application problems, as the authors stated [16]. An interesting work from the same group is in [17], in which a parabolic curve with adjustable curvature is developed and integrated with the sampling process.

All these MPP-based approaches use analytic techniques to approximate the failure surface in order to alleviate the computational burden of direct MCS. This class of methods represents the mainstream in reliability assessment. However for problems with a nonlinear limit-state function, their accuracy is suspicious [14]. Also for problems with expensive functions, the search of MPP may not be an easy task. Moreover, these methods have tremendous difficulties in solving problems of multiple failure regions or multiple MPP's. It is because there is no guarantee that all MPP's can be identified and it is impossible to use one limit state to divide safe or failure regions, e.g., for Problem 2 (to be defined in Section 5) as shown in Fig. 3 with two failure regions and four limit state surfaces.

(Insert Fig. 3 here.)

More recently, an emerging class of methods are under development [18-20]. These methods deviate from MPP-based approximation methods and endeavour to achieve the balance between MCS and

approximation. Continuing on their previous work by using importance sampling (IS) based methods [11], Zou and colleagues developed an indicator response surface based method, in which MCS is only performed in a reduced region around the limit state [18]. A recent work in literature reduces the sampling cost by approximating the output distribution with an analytical function [19]. The authors recently developed a more flexible discriminative sampling method with high efficiency and accuracy [20]. The work in [20], however, bears two limitations: 1) the reliability assessment relies on the accuracy of the approximation model of the performance function; and 2) it is not designed to solve problems of multiple failure regions. This work significantly improves [20] by addressing these two limitations. A new concept, the failure surface frontier (FSF), is proposed in order to gain a deeper insight to the failure surface and regions. Based on FSF, reliability assessment can be objectively assessed and does not rely on the accuracy of the approximation model. It solves problems of multiple failure regions with high efficiency. It also can effectively solve problems of multiple MPP's or of very low probability failure regions.

With respect to the difficulties in computing Eq. (3) or (4), this work targets at RA problems with expensive performance function; $j_x(x)$ is assumed readily available; and the integral is to be computed through MCS on the obtained FSF.

2. Failure surface frontier

Definition of Failure Surface Frontier (FSF)

For a random vector $x = (x_1, x_2, ..., x_n)$ defined on the space $S_x \subseteq \Re^n$ with a joint probability density function (JPDF) $j_x(x)$, we assume that the probability center, x_c , the point of the maximum $j_x(x)$ value, is not a failure point. That is,

$$\text{for } \forall x \in S_x,\ j_x(x) \le j_x(x_c),\ x_c \in S_x, \text{ and } g(x_c)>0 \tag{6}$$

For independent variables with normal distributions, its x_c will be at the mean $\mu = (\mu_1, \mu_2, ..., \mu_n)$. For other cases, a global optimization process might need to be called on $j_x(x)$ to identify the probability center, x_c. For uniform $j_x(x)$, any point is a probability center. The one closest to the mean $\mu = (\mu_1, \mu_2, ..., \mu_n)$ that satisfies Eq. (6) will be chosen as x_c.

This assumption should be satisfied for a reasonable engineering problem. It is because if the probability center is a failure point, the product design must have serious problems and need to be addressed first before resorting to optimization. Based on the assumption and determination of x_c, the definition of FSF is given as follows.

Definition: Failure surface frontier is a hyper-surface consisting of the set of non-dominated failure points on the limit states of a failure region. A non-dominated failure point is defined when this point cannot be moved closer to the probability center, x_c, within the same failure region along any one of its x_i, i =1,…,n, component direction.

For example, in Fig. 3, the failure surface frontiers are illustrated. All the star points (failure points) are either on the FSF or dominated by FSF. The other two limit state surfaces further away from the probability center are dominated by the FSF's, respectively, and thus they are not FSF's. Similarly, the $g_1(x)$=0 curve in Fig. 1 is a FSF but not $g_2(x)$=0.

Advantages of FSF Definition

Based on the concept of failure surface frontier (FSF), we view the region dominated by FSF as failure region while the other as safe region. For example, the left bottom region of $g_1(x)$ in Fig. 1 is the failure region and the right of $g_1(x)$ is the safe region.

The proposed FSF definition offers a few distinctive advantages:

1. FSF is a complete or part of the real failure surface, which could be convex, concave, or any other complex form. It is thus a more accurate representation of the limit-state surfaces for reliability assessment than MPP based linear or quadratic approximation models.
2. When failure regions are formed by one or multiple failure surfaces, FSF could be comprised of discontinuous surface segments from multiple limit states.
3. FSF sets apart the dominating failure surfaces from the rest. The proposition of FSF thus first frees our attention from the dominated failure surfaces, which are of much less significance in practice than FSF. Even though there could be safe regions in the "failure regions" determined by FSF, those areas are of much lower probability and would not likely be considered in real design practice. For example in Fig. 1, the safe region left to $g_2(x)$ will be viewed as a failure region by using FSF. Such an error is, however, on the conservative side. FSF therefore eliminates the need to obtain the dominated failure surfaces, which significantly reduces the computation effort for RA.
4. By using the proposed method in this work, FSF could be computed for expensive problems with multiple failure modes, multiple MPP's, or of extremely low probability failure regions. RA results based on FSF are of high accuracy and efficiency. Test details will be given in Section 4.

It is to be noted that all directional sampling based RA approaches are limited to star-shaped limit states, in which any ray from the origin intersects the limit state at most at one point [12]. The proposed definition of FSF shares some similarities with the star-shaped limit state assumption. In the proposed definition, only the non-dominated FSF segment from the limited states is of concern, so that any ray from the probability center, x_c, intersects the FSF at most at one point. The differences however, are many fold. First, directional sampling is performed in the transformed normal space while FSF is in the original variable space. Second, the probability center, x_c, is used instead of the transformed origin. The former offers more flexibility. Third, FSF facilitates the combined, even discontinuous, frontier from multiple limit states. From the proposed algorithm and test results in this work, the concept of FSF facilitates simpler and more aggressive sampling than directional sampling for RA on expensive functions.

FSF Computation

Mathematically, the definition of FSF resembles that of Pareto Frontier in the performance space in multi-objective optimization (MOO). An important concept from MOO, the fitness function, is to be introduced first to describe the non-dominated FSF points in the design space.

For a given set of points, a fitness function is defined as [21]

$$G_i = 1 - \max_{j \neq i}(\min(x_{s1}^i - x_{s1}^j,\ x_{s2}^i - x_{s2}^j,\ \cdots, x_{sk}^i - x_{sk}^j, \cdots,\ x_{sn}^i - x_{sn}^j)) \tag{7}$$

Where, G_i denotes the fitness value of the i^{th} point; x_{sk}^i is the scaled k^{th} random variable value of the i^{th} design, then $x_{sk}^i - x_{sk}^j$ is the scaled k^{th} random variable value difference between the i^{th} and j^{th} points, $k = 1,\ldots,n$; the 'min' operation is over all the random variables and the 'max' is over all other points

$j \neq i$ in the generation. The variables $x_{s1}, x_{s2}, ..., x_{sn}$, are normalized to a range between zero and one. For example, for x_{s1}^i,

$$x_{s1}^i = \frac{rawx_{1,i} - rawx_{1,\min}}{rawx_{1,\max} - rawx_{1,\min}} \tag{8}$$

Where, the subscript 's' of x_{s1}^i stands for scaled; $rawx_{1,i}$ denotes raw (un-scaled) value of the first random variable for the i^{th} point; $rawx_{1,max}$ denotes the maximum raw value of the first random variable among all points; and $rawx_{1,min}$ denotes the minimum raw value of the first random variable among all points. For normalized variables in [0 1], it is known that all the points on the frontier in a point set should have a fitness function value equal to or larger than 1, and points with a fitness value less than 1 are definitely dominated by the frontier points. By calculating the fitness values of points in the failure regions, FSF can be identified.

(Insert Fig. 4 about here.)

For example there are four points in Fig. 4 with normalized coordinates. Using their coordinates, their fitness values can be calculated by using Eq. (7). As one can see that only one point (0.9, 0.8) has a fitness value less than 1 and it is a dominated point. The other points are on the frontier or called dominating points. It is to be noted that the frontier could be of any shape, convex or non-convex. The frontier shown in Fig. 4 is apparently non-convex.

Given the concept of FSF, the goal of the proposed method is to identify all the FSF's in the design space, based on which reliability is assessed. In this work, the design space is translated so that the probability center x_c is at the origin of the new design space. Therefore, the fitness value of each failure point in the design space can be computed using its relative coordinates. Because the fitness value of FSF is used to guide the sampling process (which will be described later), for multiple failure regions in problems such as in Fig. 3, the fitness value is calculated respectively in each quadrant (for n=2) (we use "quadrant" as a general term for n>2 cases), so that all FSF's can be obtained simultaneously.

3. Proposed FSF-based reliability assessment algorithm

Failure surface frontier (FSF) is defined mostly to facilitate the efficient reliability assessment for expensive performance functions, for which the efficiency of RA is largely manifested by the total number of function evaluations. This work employs the kriging model as an approximation of the performance function. The kriging model, however, is not used as a surrogate of the failure surface for RA, but rather as a guide to sample more points on FSF.

The kriging model is defined below [22-24].

$$\hat{y}(X) = \sum_{i=1}^{k} a_i f_i(X) + z(X) \tag{9}$$

Kriging model consists of two parts. The first part is usually a simple linear regression of the data. The second part is a random process. The coefficients, a_i, are regression parameters. $f_i(x)$ is the regression model. The random process $z(X)$ is assumed to have mean zero and covariance, $v(x_1, x_2) = \sigma^2 R(x_2 - x_1)$. The process variance is given by σ^2 and its standard deviation is σ. The

spatial correlation function, $R(.)$, controls the smoothness of the model, influence of other nearby points, and differentiability of the approximation model, or metamodel. Kriging is flexible to approximate complex functions. It interpolates sample points, and the influence of other nearby points is controlled by the spatial correlation function. Based on these features, kriging is chosen for the proposed method. The proposed method, however, does not dictate the exclusive use of kriging as the metamodel.

A kriging toolbox is given by [25]. It provides regression models with polynomials of orders 0, 1, and 2, as well as seven spatial correlation functions for selection. This work uses the regression model with polynomials of order 0, and the Gaussian correlation model. A detailed description of kriging is in the corresponding author's previous work [23].

Given the concept of FSF, the goal of the proposed method is to assess reliability by means of FSF. For clarity, points evaluated by expensive processes are referred as evaluated points or expensive points; points calculated from metamodels are referred as cheap points.

Algorithm

Conceptually, the proposed algorithm consists of three major steps: 1) sampling in the original space and building a metamodel, 2) using the metamodel to generate many points, from which selecting a few (guided by FSF) for expensive evaluation, and 3) returning to Step 1 for re-sampling in the original space. For Steps 1) and 3), the sample points are generated in the original variable space; no variable transformation is performed. After the sample points are generated at Step 1), Step 2) is to locate FSF from these sample points. The fitness value as defined in Eq. (7) is a relative measure with respect to the rest of sample points in a point set. Therefore Eq. (8) depends on the value of other points in the same set. After applying Eqs. (7) and (8), each sample point (in the original variable space) receives its fitness value, as if each sample point is labelled with its fitness value. A few points are selected for expensive evaluation according to their fitness value. Then the sampling process continues in the original variable space. In the end, the expensive points will tend to concentrate on or near FSF. Step 1) and 3) employ ordinary random sampling processes, for which existing sampling methods can be applied for either correlated or uncorrelated variables . The novelty of the proposed method mainly lies in Step 2, which employs metamodeling and the concept of FSF to select a few from the sample points for expensive evaluation. Specifically, the algorithm consists of following detailed steps:

Step 1: Generate a large number of sample points in the design space according to the joint probability density function (JPDF) of random variables. As one can see that from these sample points, $(2n + 1)$ initial points are chosen; n is the number of random variables. The end points along each variable direction, as well as the point defined by the mean of all x_i components, are chosen as the initial points. Note the number of initial points does not increase exponentially with the number of variables.

Step 2: Evaluate the initial points by calling the expensive performance function, $f(x)$. Construct a kriging model based on the initial points. The kriging model is thus a metamodel of $f(x)$.

Step 3: Classify the evaluated points into failure and safe points according to their limit state function value. If failure points exist, identify the FSF points by Eq. (7) from these points. Recall that the fitness calculation was performed in each different quadrant.

Step 4: Randomly generate a large number of cheap points from the kriging model (e.g. 10^3). From these cheap points, new sample points are to be picked and be evaluated. We would like to avoid points that have an extremely low probability. Also we need randomly distributed sample points, from

which points of desired property will be picked. The questions at this step are then 1) *how to determine the sampling region?* and 2) *how to avoid extremely low probability points?* These questions are addressed by the following sub-algorithm:

4a. Generate a relatively large number of sample points in the design space according to the JPDF, and identify the min($j_x(x)$) from these sample points.

4b. Randomly generate a large number of sample points in the 6 sigma hyper-box and their $j_x(x)$ values can be computed. The points whose $j_x(x)$ value is less than min($j_x(x)$) are discarded from the point set. As a result, the left-over points will be randomly distributed and all have a higher $j_x(x)$ value than min($j_x(x)$).

The sub-algorithm was developed in order to balance the efficiency and accuracy. In cases that the accuracy is of prominent concern, Step 4a could be eliminated and one can directly sample within the 6 sigma hyper-box.

Step 5: Classify the cheap points into two classes (failure and safe points) using the kriging model prediction.

Step 6: Draw *one* sample from the cheap safe points to explore new failure regions.

- Randomly draw a number of points according to the inverse of their failure probability, i.e., low probability points will have higher chance of being sampled. This is because in engineering design failure points are often of low probability. Since the intention is to identify failure regions, we would like those regions to have higher chance of being explored.
- Calculate the minimum distance between each of the above points and the expensive evaluated safe points; pick the sample point that has the maximum minimum distance (i.e., maxmin distance) so that the new sample point is far from the expensive safe points.

Step 7: Combine the cheap failure points with the expensive failure points, calculate their fitness values in each quadrant, and identify the FSF points by Eq. (7) from the combined point set. One can choose all or some of the cheap FSF points according to their fitness values as new sample points. In addition, find the minimum $j_x(x)$ value of the combined FSF points and the minimum of all the evaluated FSF points. If the former is less than the later, and the point yielding the minimum $j_x(x)$ in the combined FSF is not yet in the new sample set, add it to the new sample set.

Step 8: Evaluate the new drawn points in Step 6 and Step 7 by calling the expensive process.

Step 9: If convergence criteria are satisfied, perform RA and the process terminates. Otherwise, go back to Step 2 with all evaluated points as initial points.

Convergence Criteria

This work applies two convergence criteria. The first criterion measures the closeness and distribution of the points on the failure surface frontier (FSF) at the last iteration. It is known that when points on the frontier are closely and evenly distributed, the fitness value of all frontier points tends to be 1 [26]. Therefore, in this work we set the average fitness value of all the current failure surface frontier points less one, $\overline{G} = \frac{1}{L}\sum_{i=1}^{L} G_i - 1$, close to 0 as a convergence criterion, where L is the number of frontier points in the last iteration. In practice, we use $\overline{G} \le 0.01$.

The second criterion checks the convergence of FSF. Define the number of expensive FSF points at the last iteration as N_f; and define N_n as the number of FSF points among N_f that are dominated by any of the new FSF points generated at the current iteration. Then the ratio $\lambda = N_n / N_f$, $\lambda \in [0, 1]$ is expected to approach zero as FSF converges. In the testing, $\lambda \le 0.05$ is applied. It is observed from testing that for most testing cases, the convergence history is similar to that shown in Fig. 8 (which will be discussed later).

For cases when the failure region is extremely small or does not exist, there is no new failure point being drawn after a number of iterations, or although new failure points can be drawn, it is very difficult to satisfy all of the two criteria simultaneously, the process will terminate after a prescribed maximum number of iterations has been reached.

Reliability Assessment

After FSF's have been identified, at Step 9 of the algorithm, the failure probability of the problem is to be evaluated.

Assume we now generate a large number of sample points. These sample points are judged by FSF's to see if they are in the safe or failure region. That is, if existing FSF's cannot dominate a sample point, the point is then in the safe region, otherwise it is in the failure region. These sample points are also evaluated by the final kriging model. If a sample point's model function value is larger than the safety threshold, it is considered safe; otherwise, failure. The combination of the two independent criteria will create two situations:

1. Both criteria are consistent, i.e., they simultaneously predict a point either safe or failure.
2. The two criteria are inconsistent, i.e., their predictions conflict.

In Situation 1, since there is no conflict, their prediction will be used as the final judgment for the point. In Situation 2, the point will be evaluated by calling the expensive performance function as the final judgment. From the error analysis point of view, an error can occur in Situation 1 when a point is in fact in the failure region but not dominated by existing FSF points. This happens because FSF might be incomplete. Another type of error in Situation 1 is the mis-judgment of a safe point as a failure point. This happens in cases such as the safe region left to $g_2(x) = 0$, as depicted in Fig. 1. The argument is such an error is in the conservative side and the points in such regions will unlikely be chosen as design candidates. Moreover, both types of errors are remedied in Situation 2, in which the point will be evaluated and used in the reliability assessment.

Because FSF is simply a set of expensive evaluated points on the limit state, it thus provides an objective measure for RA. The proposed FSF-based RA method thus alleviates the reliance on an accurate metamodel, which is hard to obtain and validate.

4. Tests of the Proposed Method

Five test problems are chosen from the literature. They are assumed to be expensive black-box functions and hence the efficiency of the reliability assessment method is measured by the total number of evaluations of the expensive performance function.

Problem 1 [11] has a performance function $f(x_1, x_2) = x_1^3 + x_2^3$. Its limit state function is $g(x_1, x_2) = x_1^3 + x_2^3 - 18$. The distribution of random variables is $x_1 \sim N(10,5), x_2 \sim N(9.9,5)$. This

problem has only one failure mode and has been widely used. It provides a good reference for quantitative comparison.

Problem 2 is given by [27] and used by [11; 18]. It is formulated from a tuned vibration absorber (TVA) system. The amplitude of the system is the performance function described as below:

$$f(\beta_1,\beta_2)=\frac{|1-(\frac{1}{\beta_2})^2|}{\sqrt{[1-R(\frac{1}{\beta_1})^2-(\frac{1}{\beta_1})^2-(\frac{1}{\beta_2})^2+\frac{1}{\beta_1^2\beta_2^2}]^2+4\varsigma^2[\frac{1}{\beta_1}-\frac{1}{\beta_1\beta_2^2}]^2}} \tag{10}$$

Where R is the mass ratio of the absorber to the original system, ς is the damping ratio of the original system, and β_1 and β_2 are the ratios of the natural frequency of the original system and vibration absorber with respect to the excitation frequency, respectively. In this work, R and ς are treated as deterministic variables with R = 0.01 and ς = 0.01; only β_1 and β_2 are random variables with a distribution $\beta_1 \sim N(1, 0.025)$ and $\beta_2 \sim N(1, 0.025)$. The objective of the design problem is to reduce the risk of the amplitude being larger than a certain value, under the uncertainties of the parameters. The limit state function for this example is $g(x_1,x_2)=28.0-f(x_1,x_2)$. Variables x_1, x_2 are used in the function for consistency instead of β_1 and β_2, respectively. This problem has multiple failure regions, as shown in Fig. 3. It was referred as an "extremely difficult" problem [18].

Problems 3 is modified from [28]. The limit state function is given as

$$g(x_1,x_2)=\begin{cases}3-\sqrt{x_1^2+x_2^2} & |x_1|\le 3\sqrt{2},\ x_2>0\\ 3\sqrt{2}-|x_1|-x_2 & otherwise\end{cases} \tag{11}$$

Where both x_1 and x_2 are independent, identically and standard normally distributed.

Problems 4 is from [28]. The limit state function is given as

$$g(x_1,x_2)=3-x_2+(4x_1)^4 \tag{12}$$

Where both x_1 and x_2 are independent, identically and standard normally distributed.

Problem 5 is a practical engineering example from [14]. The settlement of a point A in Fig. 5 caused by the construction of a structure can be shown to be primarily caused by the consolidation of the clay layer. Suppose the contribution of settlement due to secondary consolidation is negligible. For a normally loaded clay, the settlement S is given by

$$S = \frac{C_c}{1+e_0} H \log \frac{p_0+\Delta p}{p_0} \tag{13}$$

where C_c is the compression index of the clay; e_0 is the void ration of the clay layer before loading; H is the thickness of the clay layer; p_0 is the original effective pressure at point B (mid-height of the clay layer) before loading; and Δp is the increase in pressure at point B caused by the construction of the structure; "log" denotes logarithm to the base 10. Because of the non-uniform thickness and lack of

homogeneity of the clay layer, the settlement predicted by the empirical formula could be subject to model error, which may be corrected by a factor N.

(insert Fig. 5 about here)

Suppose satisfactory performance requires that the settlement be less than 2.5 inches and the variables have the statistics as shown in Table 1.

(Insert Table 1 here)

Since the probability distributions of the variables are unknown, they will be assumed to be normally distributed. Determine the probability of excessive settlement at point A in Fig 5.

$$g(X) = 2.5 - \frac{C_c}{1+e_0} H \log \frac{p_0+\Delta p}{p_0} \tag{14}$$

The test results and comparison with related methods are listed in Tables 2-3 respectively. For Monte Carlo simulation, the required number of samples, n_s, and %error is determined by [14; 29]

$$\%error = 100 \times z_{\alpha/2} \sqrt{\frac{1-p_f}{n_s \times p_f}} \tag{15}$$

where $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$, $\Phi^{-1}(.)$ denotes the inverse function of the standard normal cumulative distribution. α reflects the confidence level. Usually the 95% confidence ($\alpha = 0.05$) is taken. For a fair comparison, the sample size for Problem 1 and 2 is taken from the references [11; 18], respectively. The sample sizes for Problems 3 and 4 are determined from Eq. (15) with a small %error.

Table 2 list the test results by using the proposed approach for the 5 test problems, respectively. Table 3 compares the test results with the state-of-the-art in the literature for the corresponding test problem. For Problem 1, which is a relatively simple problem, the proposed method achieved the highest accuracy and efficiency. For Problem 2, the MPP-based methods failed to find the two MPP's. The multi-modal AIS based MPP also failed to identify both failure regions, and severely underestimated the probability of failure. The failure boundary for Problem 3 is convex and coincides with the contour of the performance function with an infinite number of MPP's [28] . Therefore FORM or SORM will have difficult or lead to significant risk for this problem. It also presents a problem for MPP-based importance sampling methods as there is no single MPP. Problem 4 has a highly concave limit state curve [28]. Due to the low probability of the failure region, it will result in low efficiency for MCS or MPP-based importance sampling methods. Ref. [28] applied the IS on kernel density method to solve both Problems 3 and 4. Although the kernel density based IS method [28] is not based on MPP, it needs a large number of samples to obtain the first estimate of the kernel density. With the available data given by [28], the proposed method out wins for both Problems 3 and 4 in efficiency, as shown in Tables 2 and 3. For a larger scale problem, i.e., Problem 5 with six design variables, the proposed method also outperforms in efficiency the Monte-Carlo simulation method to achieve the same accuracy.

(Insert Figs. 6-7 about here.)

Fig. 6 plots all the evaluated expensive points in the design space for Problem 2. It is visually apparent that significantly more sample points are generated on or close to the limit states. Fig. 7 plots the evaluation result for 10^4 random sample points for Problem 2 using the proposed method against the evaluation directly from the performance function. Similar plots have been obtained for other test problems. The plots in Fig. 6 and Fig. 7 indicate that the chance of sampling in Problem 2 is almost identical for both failure regions. This proves the capability of the proposed method for problems with multiple failure regions. Fig. 8 plots the convergence history for Problem 1 as an example. Similar convergence behaviour for other problems is also observed.

5. Discussions

For expensive performance functions, metamodels are commonly used [20]. The reliability assessment (RA), however, often depends on the accuracy of the metamodel. This method fundamentally differs from this common practice by introducing an objective means, failure surface frontier (FSF), for the purpose of RA. The metamodel is employed in this work mainly as a guide for iterative sampling. It is used in RA only as a supplementary tool to identify controversial points, at which the expensive performance function is called for evaluation. Therefore, the proposed RA method does not rely on the accuracy of the metamodel.

The importance sampling (IS) based methods have two types. One is the IS on design points and the other is IS on kernel sampling density [11; 28; 29]. IS on the design point is to move the center of sampling to a point on the limit state, which is usually the conventional MPP. This method was promoted by [30]. The method, however, bears a few problems:

1. In general, the search for design points occupies a considerable portion of the total computational effort.
2. If there are multiple MPPs, the search for multiple design points requires more sophisticated algorithms for the optimization problem. Moreover, since it is not known a priori, multiple points search has to be carried out for every problem.
3. The application of design points to IS becomes more difficult or inefficient in situations such as noisy limit state functions, relatively flat PDFs along or in the neighborhood of the limit state surfaces, and highly concave/convex limit state surfaces. When such situations are not properly handled in IS, the estimate can have a large variance or even become practically biased [28] .

The IS on kernel sampling density is to use sample points to construct a kernel density estimator of the optimal IS density. The main drawback is that the points used to construct the kernel sampling density are simulated by the basic Monte Carlo procedure, so the probability of having samples generated in the failure region is equal to the failure probability, which is usually small in practical applications. The simulation of points lying in the failure region thus requires a very large number of sample points and so the method is computationally expensive. This work has chosen the version from [28] for the comparison (See Tables 6 and 8), which is considered the most efficient approach of this category by [29].

Comparatively, the proposed RA method based on the concept of FSF has a few distinctive features:

1. There is no need to perform transformation of design space. Since the method works always in its original design space, no transformation and thus approximation of the variable distribution is needed. This eliminates the possible error in transformation.

2. The method uses FSF as an objective means for reliability assessment instead of relying on the accuracy of the metamodel.
3. The proposed method does not call any MCS process directly on the expensive performance function.
4. The proposed strategy is simple and easy to understand by practitioners. Little tuning of parameters is required with no need to guess any parametric family of densities as in the case of IS methods.
5. The proposed method works for challenging problems with multiple failure regions, multiple (infinite) number of MPP's, and extremely small failure probabilities.

It is to be noted that the proposed method is a discriminative sampling method, in some sense an intuitive method. It is, therefore, lacks of a rigorous approach to quantify its error as that in Eq. (15) for Monte Carlo simulation. In addition, as dimensionality increases, a large set of evaluated FSF points is needed to represent the FSF's so that the number of evaluation increases. The efficiency of the method for high-dimensional problems rely on a few things, 1) the efficiency of constructing the approximation model (in this case, kriging), 2) the variance of each design variable as it determines the size of design space, and 3) the complexity of the failure surface. An approximation of FSF from the obtained FSF points might be used to improve efficiency. Future study will be on the reliability assessment for high-dimensional problems with expensive performance functions.

6. Conclusions

The work aims at reliability assessment on expensive performance functions. The proposition of the concept of failure surface frontier (FSF) in this work provides a better alternative to existing linear or quadratic approximation of expensive limit state surfaces. It also brings our attention to FSF, the set of most significant failure points from the entire failure surfaces (limit states). The final reliability assessment can be performed by employing FSF as an objective measure, instead of relying on the accuracy of approximation.

Practically, this work was a significant improvement to our recent work [20] to tackle problems of multiple failure regions, multiple MPP's, and low probability failure regions. The proposed algorithm based on the concept of FSF is able to solve these problems effectively and efficiently.

Test and comparison results show that the proposed method bears high accuracy and efficiency for four test problems with representative features. Its performance on high dimensionality problems needs to be further studied.

Acknowledgements

Financial support from the Natural Science and Engineering Research Council (NSERC) is highly appreciated. Valuable comments from anonymous reviewers are gratefully acknowledged which result in a much improved manuscript.

References

[1] Du, X., Sudjianto, A. and Chen, W., 2003, "An Integrated Framework for Probabilistic Optimization Using Inverse Reliability Strategy," *ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois, USA, September 2-6.

[2] Du, X. and Sudjianto, A., 2003, "Reliability-Based Design with the Mixture of Random and Interval Variables," *ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois, USA, September 2-6.
[3] Liu, H., Chen, W., Sheng, J. and Gea, H.C., 2003, "Application of the Sequential Optimization and Reliability Assessment Method to Structural Design Problems," *ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois, USA, September 2-6.
[4] Wu, Y.T. and Wang, W., 1998, "Efficient Probabilistic Design by Converting Reliability Constraints to Approximately Equivalent Deterministic Constraints," Transactions of the SDPS, Journal of Integrated Design and Process Science, 2(4), pp. 13-21.
[5] Youn, B.D., Choi, K.K. and Park, Y.H., 2003, "Hydrid Analysis Method for Reliability-Based Design Optimization," Transactions of ASME, Journal of Mechanical Design, 125, pp. 221-232.
[6] Ang, A.H.-S. and Tang, W.H., 1975, *Probability Concepts in Engineering Planning and Design, Vol. 1 Basic Principles*, Wiley, New York.
[7] Ross, S.M., 2002, *Simulation*, Academic Press, San Diego, California.
[8] Kloess, A., Mourelatos, Z. and Meernik, P., 2004, "Prababilistic Analysis of An Automotive Body-Door System," International Journal of Vehicle Design, 34(2), pp. 101 - 125.
[9] Ditlevsen, O., Olesen, R. and Mohr, G., 1987, "Solution of A Class of Load Combination Problems by Directional Simulation," Structural Safety, 4, pp. 95-109.
[10] Walker, J.R., 1986, "Practical Application of Variance Reduction Techniques in Probabilistic Assessments," *The Second International Conference on Radioactive Waste Management*, Winnipeg, Manitoba, Canada.
[11] Zou, T., Mahadevan, S., Mourelatos, Z. and Meernik, P., 2002, "Reliability Analysis of Automotive Body-door Subsystem," Reliability Engineering and System Safety, 78, pp. 315-324.
[12] Nie, J. and Ellingwood, B.R., 2005, "Finite Element-Based Structural Reliability Assessment Using Efficient Directional Simulation," Journal of Engineering Mechanics, 131(3), pp. 259-267.
[13] Freudenthal, A.M., 1956, "Safety and Probability of Structural Failure," ASCE Transactions, 121, pp. 1337-1397.
[14] Ang, A.H.-S. and Tang, W.H., 1984, *Probability Concepts in Engineering Planning and Design, Vol. 2 Decision, Risk and Reliability*, Wiley, New York.
[15] Mahadevan, S., Chiralaksanakul, A., 2005, “First-Order Approximation Methods in Reliability-Based Design Optimization,” Transactions of ASME, Journal of Mechanical Design, 127, pp. 851-857.
[16] Wu, Y.T., Millwater, H.R. and Cruse, T.A., 1990, "Advanced Probabilistic Structural Analysis Method for Implicit Performance Functions," AIAA Journal, 28(9), pp. 1663-1669.
[17] Wu, Y.T., 1994, "Computational Methods for Efficient Structural Reliability and Reliability Sensitivity Analysis," AIAA Journal, 32(8), pp. 1717-1723.
[18] Zou, T., Mourelatos, Z., Mahadevan, S. and Tu, J., 2003, "An Indicator Response Surface-Based Monte Carlo Method for Efficient Component and System Reliability Analysis," *ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois USA, September 2-6.
[19] Pitchumani, R., Mawardi, A., 2005, “SAMS: Stochastic Analysis With Minimal Sampling – A Fast Algorithm for Analysis and Design Under Uncertainty,” Transactions of ASME, Journal of Mechanical Design, 127, pp. 558-571.

[20] Wang, G.G., Wang, L. and Shan, S., 2005, "Reliability Assessment Using Discriminative Sampling and Metamodeling," *2005 SAE World Congress*, Cobo Center, Detroit, MI, USA, April 11-14.
[21] Schaumann, E.J., Balling, R.J. and Day, K., 1998, "Genetic algorithms with multiple objectives," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis.
[22] Martin, J.D. and Simpson, T.W., 2003, "A Study on the Use of Kriging Models to Approximate Deterministic Computer Models," *ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois, USA, September 2-6.
[23] Wang, G.G. and Simpson, T.W., 2004, "Fuzzy Clustering Based Hierarchical Metamodeling for Space Reduction and Design Optimization," Journal of Engineering Optimization, 36(3), pp. 313-335.
[24] Jones, D.R., Schonlau, M. and Welch, W.J., 1998, "Efficient Global Optimization of Expensive Black Box Functions," Journal of Global Optimization, 13, pp. 455-492.
[25] Lophaven, S.N., Nielsen, H.B. and Søndergaard, J., 2002, *DACE: A Matlab Kriging Toolbox*, Technical University of Denmark, Kgs. Lyngby, Denmark.
[26] Shan, S. and Wang, G.G., 2005, "An Efficient Pareto Set Identification Approach for Multi-objective Optimization on Black-box Functions," Transactions of the ASME, Journal of Mechanical Design (In press).
[27] Chen, S., Nikolaidis, E., Cudney, H.H., Rosca, R. and Haftka, R., 1999, "Comparison of Probabilistic and Fuzzy Set Methods for Designing under Uncertainty," *40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural dynamics, and Materials Conference and Exhibit*, St. Louis, MO, April 12-15.
[28] Au, S.K. and Beck, J.L., 1999, "A new adaptive importance sampling scheme for reliability calculations," Structural Safety, 21, pp. 135-158.
[29] Schuëller, G.I., Pradlwarter, H.J. and Koutsourelakis, P.S., 2003, "A comparative study of reliability estimation procedures for high dimensions," *16th ASCE Engineering Mechanics Conference*, University of Washington, Seattle, USA, July 16-18.
[30] Du, X. and Chen, W., 2000, "Towards a better understanding of modeling feasibility robustness in engineering design," Transactions of ASME, Journal of Mechanical Design, 122, pp. 385-394.

List of Tables

List of Figures

Table 1 Statistics data for Problem 5

	Mean	C.O.V.	Standard Deviation
N	1.0	0.10	0.10
C_c	0.396	0.25	0.099
e_0	1.19	0.15	0.1785
H	168 in.	0.05	8.4
p_0	3.72ksf	0.05	0.186
Δp	0.50ksf	0.20	0.10

Table 2 Test Results for 5 problems

Example #	No. of iterations	No. of function evaluations	Convergence Criteria 1	Convergence Criteria 2	No. of failure points	No. of safe points	No. of errors compared with MCS
1	36	215	0.0082	0	617	11,983	0
2	24	285	0.0086	0.0125	326	29,674	3
3	27	284	0.0098	0.0460	382	99,618	2
4	52	215	0.0335	0	912	4,999,088	0
5	31	2213	0.0371	.0938	361	3639	0

Table 3 Test Result Comparison for 5 Problems

Example #	Methods	Probability of failure	Difference from MCS result (%)	No. of function evaluations
1[1]	FORM with response surface	0.0256	352	97
	SORM with response surface	0.0162	186	97
	Original multi-modal AIS	0.00617	9.0	450
	Multi-modal AIS	0.00575	1.6	560
	MCS	0.0051	0	120,000
	FSF	0.0051	0	215
2[2]	FORM	No convergence	-	-
	Multi-modal AIS	0.00407	58	850
	IRS-Based Monte Carlo	0.00963	0.7	206
	MCS	0.0108	0	30,000
	FSF	0.0109	0.9	285
3	Multi-modal AIS on kernel (Ref. [26]) based on Markov simulation	Not Provided	Not Provided	500 (Kernel sampling density construction) + N (the number of samples used to obtain the estimate in the AIS)
	MCS	0.0038	0	100,000
	FSF	0.0038	0	284
4	Multi-modal AIS on kernel (Ref. [26]) based on Markov simulation	Not provided	Not provided	500 (Kernel sampling density construction) + N (the number of samples are used to obtain the estimate in the AIS)
	MCS	0.0001824	0	5,000,000
	FSF	0.0001824	0	215
5	MPP method (Ref. 14)	0.102	Not provided	Not provided
	MCS	0.0902	0	4,000
	FSF	0.0902	0	2213

[1] The results in the first four rows are copied from Ref. [11]. The results in the last two rows come from our computation. The total number of MC sample points are chosen as the same as that in [11], i.e., 120,000.

[2] The results in the first three rows come from Ref. [17]. The results in the last two rows come from our computation. The samples size for MC is the same as that in [17], i.e., 30000.

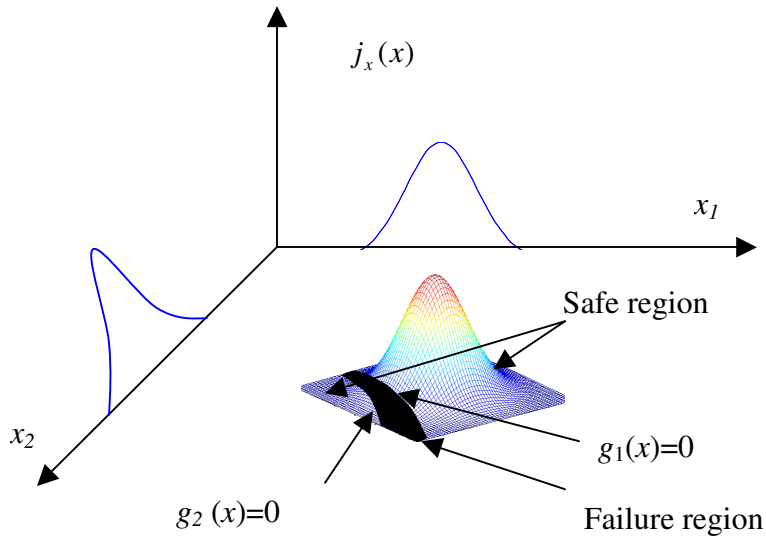

Fig. 1 An illustration of the failure surface and failure regions

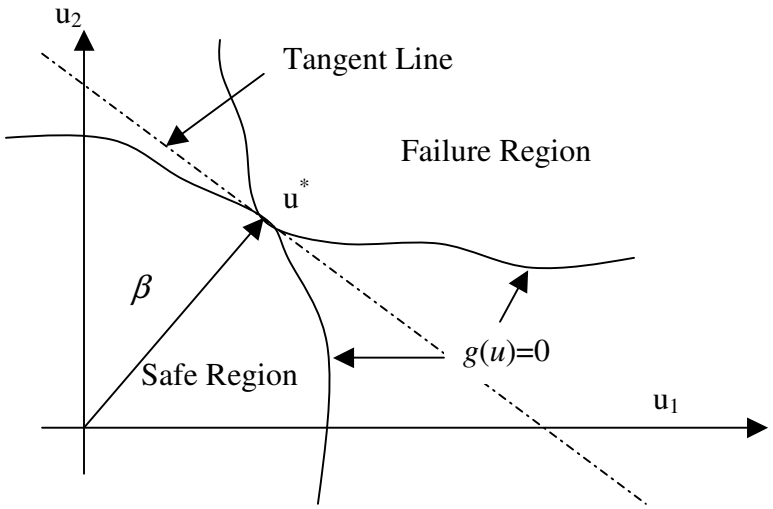

Fig. 2 Tangent line to $g(u)=0$ in the U-space

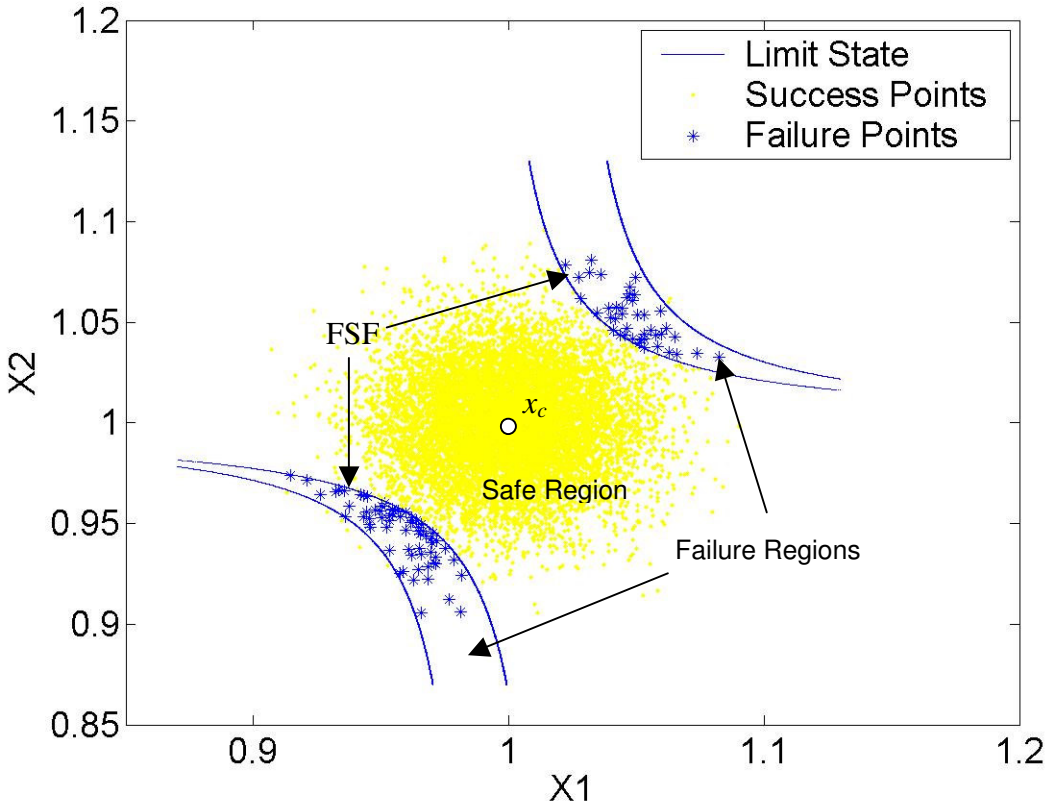

Fig. 3 A problem with 2 failure regions and many limit state surfaces

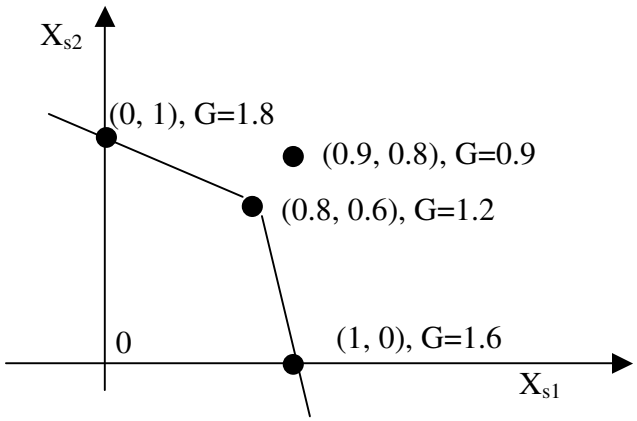

Fig. 4 An illustration of fitness function values

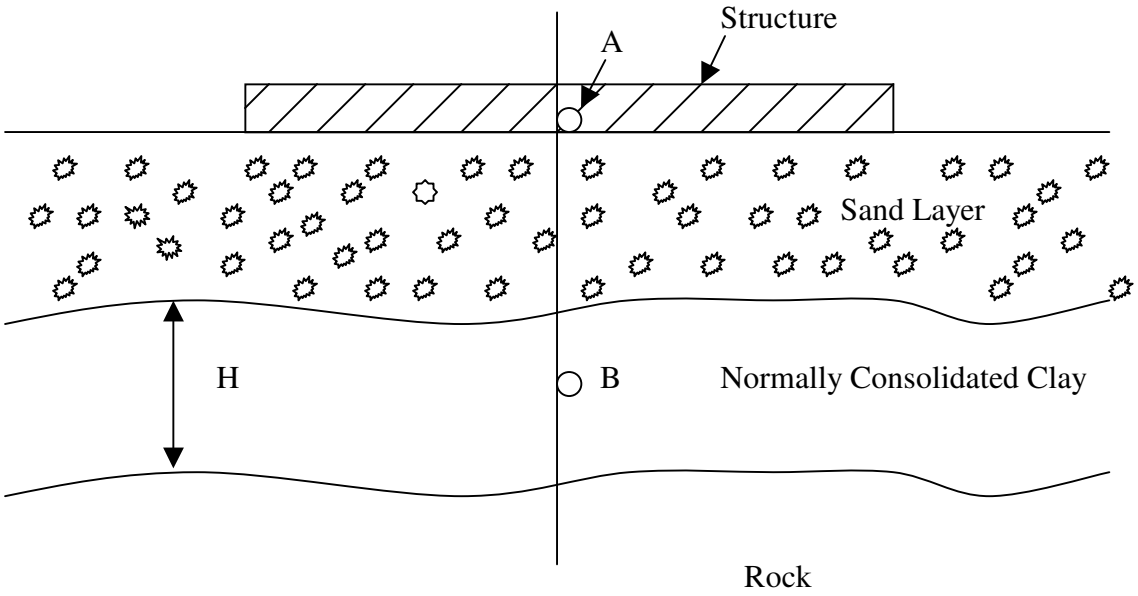

Fig. 5 Soil profile for Problem 5

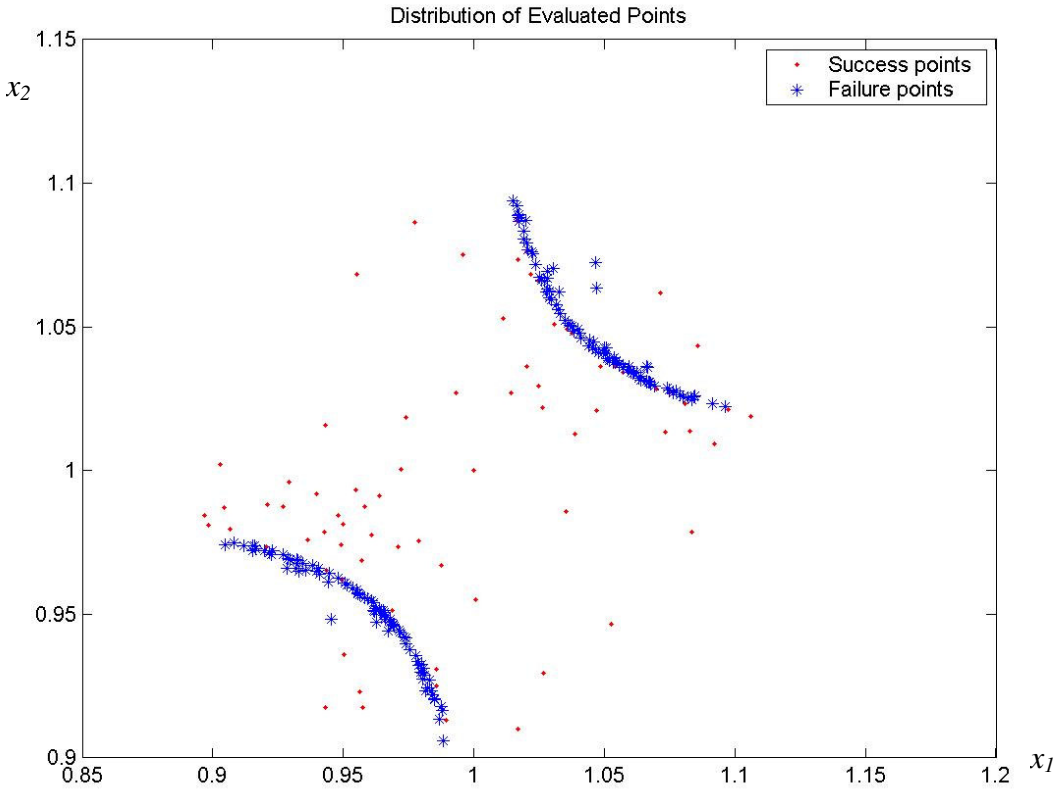

Fig. 6 Evaluated sampling points for Problem 2

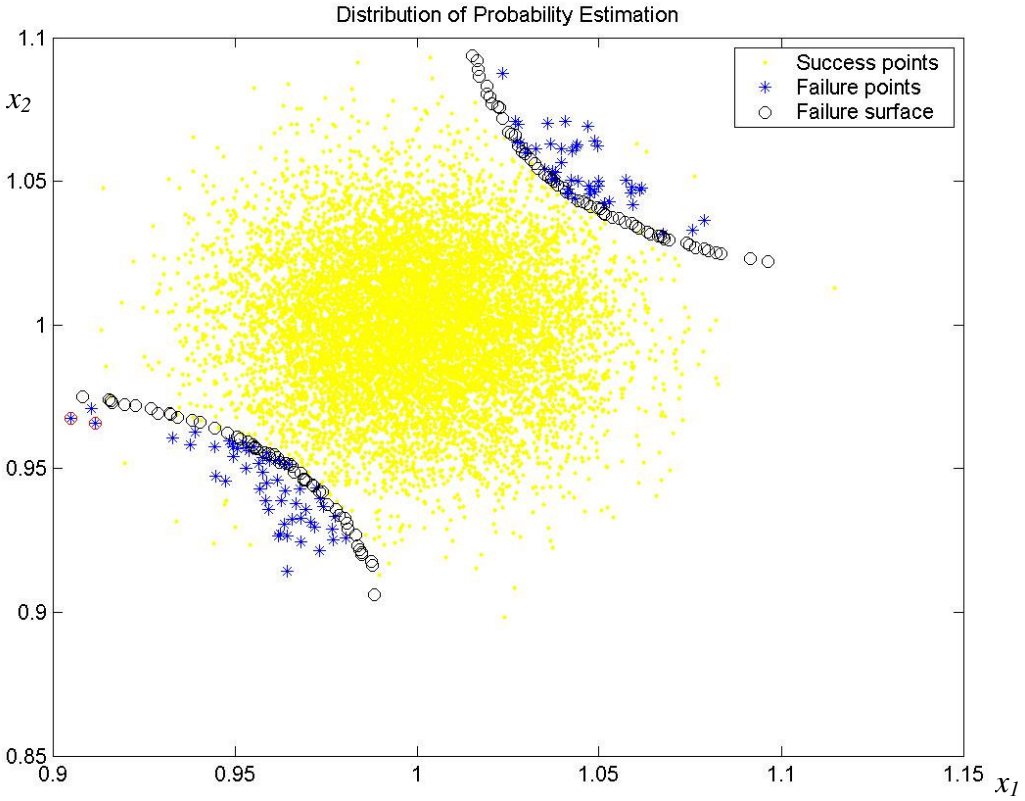

Fig. 7 Reliability assessment from 10^4 sample points for Problem 2

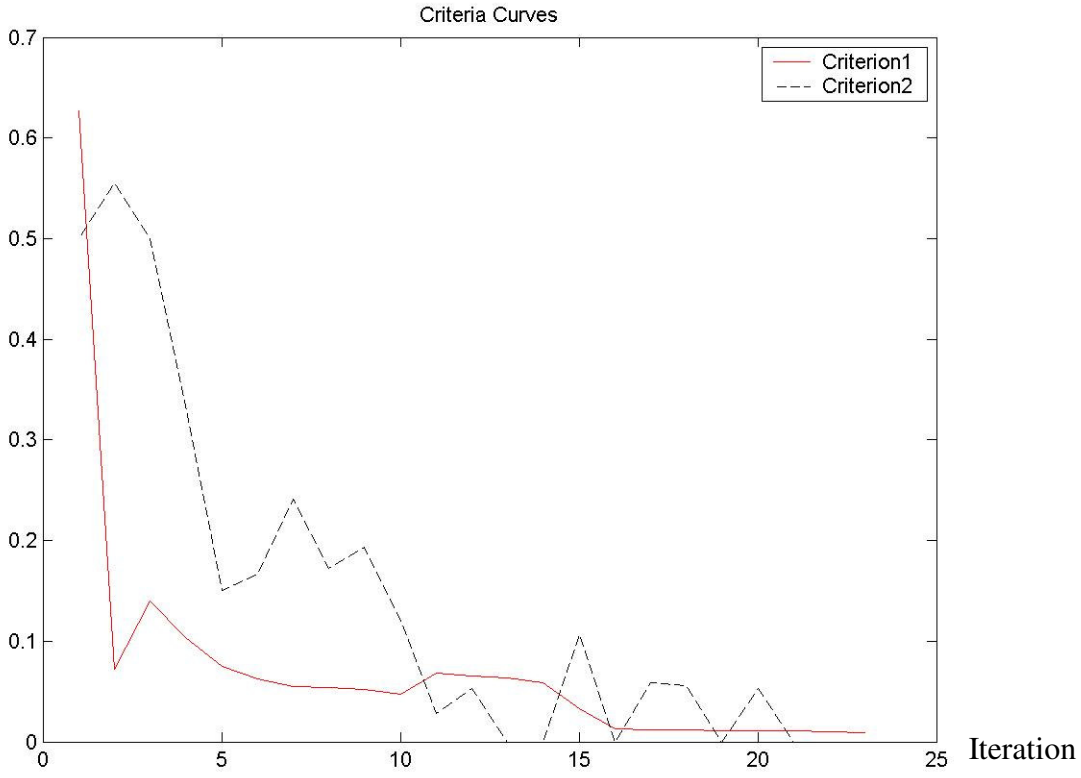

Fig. 8 Convergence history of two criteria for Problem 1