

Sampling-Based RBDO Using Score Function with Re-Weighting Scheme

Shih-Po Lin

University of California, Los Angeles

Lei Shi

Shanghai Jiao Tong University

Ren-Jye Yang

Ford Motor Co.

ABSTRACT

Sampling-based methods are general but time consuming for solving a Reliability-Based Design Optimization (RBDO) problem. In order to alleviate the computation burden, score function together with the Monte Carlo method was used to compute the stochastic sensitivities of reliability functions. In literature, re-weighting schemes were shown to converge faster than the regular Monte Carlo method. In this paper, a reweighting scheme together with score function is employed to perform sampling-based stochastic sensitivity analysis to improve the computational efficiency and accuracy. An analytical example is used to show the advantages of the proposed method. Comparisons to the conventional methods are made and discussed. Two RBDO problems are solved to demonstrate the use of the proposed method.

CITATION: Lin, S., Shi, L. and Yang, R., "Sampling-Based RBDO Using Score Function with Re-Weighting Scheme," *SAE Int. J. Mater. Manf.* 6(3):2013, doi:10.4271/2013-01-0377.

INTRODUCTION

Reliability-based design optimization (RBDO) has gained more attention recently. It has been widely applied to various engineering design problems to consider the variabilities of design variables [1,2,3,4,5,6,7]. Zhang et al. proposed an efficient probabilistic re-analysis approach to estimate the failure probability of a design for various probability distribution [8]. Papadopoulos et al. developed a novel methodology for computing the probability of structural failure efficiently by incorporating neural network and subset simulation techniques [9]. However, one of the major challenges in RBDO is that it requires double optimization loop to obtain an optimal design. In the literature, the most probable point (MPP) based reliability design optimization is commonly used but it has some challenges, e.g., it may be unable to find an MPP. It also has difficulty to find and deal with problems which have multiple MPPs. In addition, the sensitivities of the performance function are often unavailable for problems such as crashworthiness. The sampling-based RBDO method by using Monte Carlo (MC) method is general for obtaining the reliability of a complex system but requires intensive computation. Using MC method to

calculate the sensitivities of probabilistic responses, which are often obtained by using finite difference method (FDM), is very time consuming if not impossible. Besides, FDM may produce inaccurate sensitivities for the probabilistic responses, as an appropriate step size or perturbation size for each design variable needs to be determined, which is often arbitrary. A small step size may lead to inaccurate sensitivities due to the statistical noise and a large step size may result in a divergent solution.

Conventional MC method involves simulations of models at each sample which is computationally intensive. Even if a surrogate model is used, the computational cost is still significant due to repeated sampling. To circumvent the repetition of simulations, Sahin and Diwekar [14] introduced a re-weighting scheme which transforms an original probability density function (PDF) into a constant function so that the samples can be selected to be uniformly distributed in the feasible domain.

The score function is used to address those issues when applying FDM to calculate the sensitivities of probabilistic responses [10,11,12,13]. It uses the information of the

probability density function for obtaining the derivatives of the probabilistic responses analytically by using MC method.

In this research, a new sampling-based RBDO methodology using score function with re-weighting scheme is proposed. The re-weighting scheme is exploited to improve MC's accuracy and efficiency. In the following sections, the sampling-based RBDO using score function with re-weighting scheme is first presented, followed by a demonstration of an analytical example, and then two mathematical, RBDO examples are used to demonstrate the feasibility of the proposed method. Finally, the conclusion is given in the end.

SAMPLING-BASED RBDO

A general mathematical RBDO problem is formulated as:

minimize $F(\mathbf{d})$

subject to $P[G_j(\mathbf{X}) \geq 0] \leq P_{F_j}^{tar}$, $j = 1, \dots, nc$

$\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U$, $\mathbf{d} \in \mathbf{R}^{ndv}$, and $\mathbf{X} \in \mathbf{R}^n$

(1)

where F and G_j are the objective and constraint functions, respectively, $\mathbf{X}$ is the random design variable vector, $\mathbf{d}$ is the design vector, which is the mean value of $\mathbf{X}$, $P_{F_j}^{tar}$ is the target failure probability for j^{th} constraint, and nc , ndv , and n are the numbers of probabilistic constraints, design variables, and random variables, respectively. A reliability analysis requires to calculate the reliability which is defined as $1 - P_F$, where P_F is the probability of failure and is written as [12,13]:

$$P_F(\boldsymbol{\theta}) \equiv P_0[\mathbf{X} \in \Omega_F] = \int_{\mathbf{R}^n} I_{\Omega_F}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} = E_0[I_{\Omega_F}(\mathbf{X})] \quad (2)$$

where $\Omega_F \equiv \{\mathbf{x} : G(\mathbf{x}) \geq 0\}$ is the failure set for reliability analysis, $\boldsymbol{\theta}$ is a vector of distribution parameters, $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ is a joint probability density function of $\mathbf{X}$, E_0 is the expectation operator, and $I_{\Omega_F}(\mathbf{x})$ is an indicator function and expressed as:

$$I_{\Omega_F}(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega_F \\ 0, & \mathbf{x} \in \Omega \setminus \Omega_F \end{cases}; \quad \mathbf{x} \in \mathbf{R}^n \quad (3)$$

where Ω is the design space.

STOCHASTIC SENSITIVITY BY SCORE FUCTION

The sensitivity of the failure probability P_F with respect to θ_i is derived by using the score function in this section. Four assumptions, which are known as the regularity conditions [11], are made as:

1. The probability density function $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ is continuous.

2. The parameter $\theta_i \in \Theta_i \subset \mathbf{R}$, $i = 1, \dots, m$, where m is the number of distribution parameter, Θ_i is an open interval on $\mathbf{R}$.

3. The partial derivate $\partial f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}) / \partial \theta_i$ exists and is finite for all $\mathbf{x}$ and θ_i . In addition, $P_F(\boldsymbol{\theta})$ is a differentiable function of $\boldsymbol{\theta}$.

4. There exists a Lebesgue integrable dominating function $r(\mathbf{x})$ such that $\left| I_{\Omega_F}(\mathbf{x}) \frac{\partial f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i} \right| \leq r(\mathbf{x})$ for all θ_i , where $r(\mathbf{x})$ is integrable in the sense of Lebesgue, that is $\int |r(\mathbf{x})| d\mathbf{x} < \infty$.

Taking the partial derivative of P_F in equation (2) with respect to the θ_i yields:

$$\frac{\partial P_F(\boldsymbol{\theta})}{\partial \theta_i} = \frac{\partial}{\partial \theta_i} \int_{\mathbf{R}^n} I_{\Omega_F}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} \quad (4)$$

After interchanging differential and integral operators and using the Lebesgue dominated convergence theorem, equation (4) is rewritten as:

$$\begin{aligned} \frac{\partial P_F(\boldsymbol{\theta})}{\partial \theta_i} &= \int_{\mathbf{R}^n} I_{\Omega_F}(\mathbf{x}) \frac{\partial f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i} d\mathbf{x} \\ &= \int_{\mathbf{R}^n} I_{\Omega_F}(\mathbf{x}) \frac{\partial \ln f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i} f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} \\ &= E_0[I_{\Omega_F}(\mathbf{x}) \frac{\partial \ln f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i}] \end{aligned} \quad (5)$$

since $I_{\Omega_F}(\mathbf{x})$ is not a function of θ_i , the partial derivative of the log function of the joint PDF is known as the first-order score function for the parameter θ_i [11] and is denoted as:

$$s_{\theta_i}(\mathbf{x}; \boldsymbol{\theta}) = \frac{\partial \ln f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i} \quad (6)$$

Hence, the first-order sensitivity of $P_F(\boldsymbol{\theta})$ can be written as:

$$\frac{\partial P_F(\boldsymbol{\theta})}{\partial \theta_i} = E_0[I_{\Omega_F}(\mathbf{x}) s_{\theta_i}(\mathbf{x}; \boldsymbol{\theta})] \quad (7)$$

Note that in equation (6), the first-order score function is derived from the sensitivity of the joint distribution, which can be analytically obtained. The probabilistic response $P_F(\boldsymbol{\theta})$ and its sensitivity $\partial P_F(\boldsymbol{\theta}) / \partial \theta_i$ are formulated as the expectation of stochastic quantities with respect to the same density function. It facilitates their concurrent evaluations in a single stochastic analysis by using MC method. In this study, the vector of distribution parameters $\boldsymbol{\theta}$ are replaced by $\boldsymbol{\mu}$ for the sensitivity of the reliability function, since $\boldsymbol{\mu} = \{\mu_1, \dots, \mu_n\}^T$ is the mean of $\mathbf{X}$. As the score functions $s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu})$; $i = 1, \dots, n$ only depend on the parameters of the probability

distribution of random design variables $\mathbf{X} = \{X_1, X_1, \dots, X_n\}^T$, the sensitivity of the probabilistic response is expressed as:

$$\frac{\partial P_F(\boldsymbol{\mu})}{\partial \mu_i} = \int_{\Omega_x} I_{\Omega_F}(\mathbf{x}) s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu}) f_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\mu}) d\mathbf{x} = E[I_{\Omega_F}(\mathbf{x}) s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu})] \quad (8)$$

Equation (8) can be estimated by using MC method as:

$$\frac{\partial P_F(\boldsymbol{\mu})}{\partial \mu_i} \cong \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L I_{\Omega_F}[\mathbf{x}^{(l)}] s_{\mu_i}(\mathbf{x}^{(l)}; \boldsymbol{\mu}) \quad (9)$$

where L is MC sample size. Detailed derivation should refer to Rahman [11] and Lee et al. [13].

RE-WEIGHTING SCHEME

Conventional sampling scheme requires intensive computation due to repetition of model simulations. A reweighting scheme was proposed to circumvent the problem of repeated model simulations [14]. The re-weighting scheme aims to reduce the computational burden by transforming an original probability density function into a constant function so that the samples can be selected to be uniformly distributed in the feasible domain. Hesterberg presents various reweighting techniques for estimating the expected value of an output distribution without evaluating the model for the input distribution [18]. By using the re-weighting scheme, the equation (8) is rewritten as:

$$\frac{\partial P_F(\boldsymbol{\mu})}{\partial \mu_i} = \int_{\Omega_x} I_{\Omega_F}(\mathbf{x}) s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu}) \frac{f_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\mu})}{f_{\mathbf{x}}^*(\mathbf{x}; \boldsymbol{\mu})} f_{\mathbf{x}}^*(\mathbf{x}; \boldsymbol{\mu}) d\mathbf{x} \quad (10)$$

where $f_{\mathbf{x}}^*(\mathbf{x}; \boldsymbol{\mu})$ is PDF of the uniform base distribution [14,18]. A weighting function, which is the ratio of the probability density function, is defined as:

$$w(\mathbf{x}; \boldsymbol{\mu}) = \frac{f_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\mu})}{f_{\mathbf{x}}^*(\mathbf{x}; \boldsymbol{\mu})} \quad (11)$$

Substituting equation (11) to equation (10), the sensitivities of $P_F(\boldsymbol{\mu})$ is rewritten as:

$$\begin{aligned} \frac{\partial P_F(\boldsymbol{\mu})}{\partial \mu_i} &= \int_{\Omega_x} I_{\Omega_F}(\mathbf{x}) s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu}) w(\mathbf{x}; \boldsymbol{\mu}) f_{\mathbf{x}}^*(\mathbf{x}; \boldsymbol{\mu}) d\mathbf{x} \\ &= E[I_{\Omega_F}(\mathbf{x}) s_{\mu_i}(\mathbf{x}; \boldsymbol{\mu}) w(\mathbf{x}; \boldsymbol{\mu})] \end{aligned} \quad (12)$$

By using MC method, the sensitivities of $P_F(\boldsymbol{\mu})$ is finally obtained as:

$$\frac{\partial P_F(\boldsymbol{\mu})}{\partial \mu_i} \cong \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L I_{\Omega_F}[\mathbf{x}^{(l)}] s_{\mu_i}(\mathbf{x}^{(l)}; \boldsymbol{\mu}) \hat{w}(\mathbf{x}^{(l)}; \boldsymbol{\mu}) \quad (13)$$

where $\hat{w}(\mathbf{x}^{(l)}; \boldsymbol{\mu})$ is equal to $\frac{w(\mathbf{x}^{(l)}; \boldsymbol{\mu})}{\sum_{l=1}^L w(\mathbf{x}^{(l)}; \boldsymbol{\mu})}$ such that $\sum_{l=1}^L \hat{w}(\mathbf{x}^{(l)}; \boldsymbol{\mu}) = 1$.

In general, repeated model simulations are required to approximate the distribution, as the selection of samples by using MC sampling method is based on the probabilistic distribution of the random variables. In re-weighting scheme, the base probability distribution can be uniform so that various well-developed uniform sampling methods [17] such as Latin Hypercube sampling (LHS) and Hammersley Sequence sampling (HSS) can be used. In this paper, HSS is employed as the sampling approach, as it shows a rather accurate estimate than LHS method [14].

NUMERICAL EXAMPLES

In this section, three numerical examples are used to examine the feasibility of proposed methodology. One is for sensitivity analysis and the other two are for sampling-based RBDO.

Sensitivities of Probabilistic Responses

Consider three 2D high-nonlinear functions [16],

$$G_1(\mathbf{x}) = 1 - \frac{X_1^2 X_2}{20}$$

$$X_1 \sim N(5.3, 0.3^2); X_2 \sim N(0.9, 0.3^2)$$

$$G_2(\mathbf{x}) = 0.7361 + (Y - 6)^2 + (Y - 6)^3 - 0.6(Y - 6)^4 + Z$$

$$\text{where } \begin{Bmatrix} Y \\ Z \end{Bmatrix} = \begin{bmatrix} 0.9063 & 0.4226 \\ 0.4226 & -0.9063 \end{bmatrix} \begin{Bmatrix} X_1 \\ X_2 \end{Bmatrix}$$

$$X_1 \sim N(5, 0.3^2); X_2 \sim N(0.8, 0.3^2)$$

$$G_3(\mathbf{x}) = 1 - \frac{80}{X_1^2 + 8X_2 + 5}$$

$$X_1 \sim N(6.5, 0.3^2); X_2 \sim N(5.5, 0.3^2)$$

For re-weighting scheme, the lower and upper bounds of the design variables are required to be predefined for practical use, as the base distribution is uniform. An appropriate domain of the design variables should be chosen by truncating the tails of the marginal PDF, which contribute less to the stochastic systems. In this paper, $\boldsymbol{\mu} - 4\boldsymbol{\sigma}$ and $\boldsymbol{\mu} + 4\boldsymbol{\sigma}$ are defined as the lower and upper bounds for the design variables, respectively. Figure 1 shows the sensitivity results of the probabilistic responses with respect to the design variable μ_1 and μ_2 by using score function with and without re-weighting scheme. Both of these two methods agree well when the sample size is greater than 10^7 . Notice that the conventional approach shows manifest oscillation compared to the proposed method with reweighting scheme. The re-weighting scheme with HSS shows better stability and faster convergence than the conventional approach with MC method.