Struct Multidisc Optim (2008) 37:13–28
DOI 10.1007/s00158-007-0210-7

RESEARCH PAPER

Eigenvector dimension reduction (EDR) method for sensitivity-free probability analysis

Byeng D. Youn · Zhimin Xi · Pingfeng Wang

Received: 29 March 2007 / Revised: 30 July 2007 / Accepted: 9 October 2007 / Published online: 13 February 2008

Abstract This paper presents the eigenvector dimension reduction (EDR) method for probability analysis that makes a significant improvement based on univariate dimension reduction (DR) method. It has been acknowledged that the DR method is accurate and efficient for assessing statistical moments of mildly nonlinear system responses in engineering applications. However, the recent investigation on the DR method has found difficulties of instability and inaccuracy for highly nonlinear system responses while maintaining reasonable efficiency. The EDR method integrates the DR method with three new technical components: (1) eigenvector sampling, (2) one-dimensional response approximation, and (3) a stabilized Pearson system. First, $2N+1$ and $4N+1$ eigenvector sampling schemes are proposed to resolve correlated and asymmetric random input variables. The eigenvector samples are chosen along the eigenvectors of the covariance matrix of random parameters. Second, the stepwise moving least squares (SMLS) method is proposed to accurately construct approximate system responses along the eigenvectors with the response values at the eigenvector samples. Then, statistical moments of the responses are estimated through recursive numerical integrations. Third, the stabilized Pearson system is proposed to predict probability density functions (PDFs) of the responses while eliminating singular behavior of the original Pearson system. Results for some numerical and engineering examples indicate that the EDR method is a very accurate, efficient, and stable probability analysis method in estimating PDFs, component reliabilities, and qualities of system responses.

Keywords Probability analysis · Eigenvector dimension reduction · Reliability · Quality · Sensitivity-free

B. D. Youn (✉) · Z. Xi · P. Wang
Department of Mechanical Engineering,
University of Maryland, College Park,
MD 20742, USA
e-mail: bdyoun@umd.edu

Z. Xi
e-mail: zxi@umd.edu

P. Wang
e-mail: pfwang@umd.edu

1 Introduction

A high-fidelity modeling has come true as computational mechanics has been sophisticated. Thus, probability analysis is of critical importance to understand random nature of physics in various engineering applications. However, a common challenge in probability analysis is a multi-dimensional integration to quantify probabilistic nature of system responses (e.g., fatigue life, corrosion, injury metrics) in various engineering applications (e.g., vehicle, airplane, electronics). Neither analytical multi-dimensional integration nor direct numerical integration is possible for large-scale engineering applications. Other than those approaches, existing approximate methods for probability analysis can be categorized into five groups: 1) sampling method; 2) expansion method; 3) the most probable point (MPP)-based method; 4) response surface approximate method; and 5) approximate integration method.

The sampling method is the most comprehensive but expensive method to use for estimating statistical moments,

reliability, and quality of system responses. Monte Carlo Simulation (MCS) (Varghese et al. 1996; Lin et al. 1997) is the most widely used sampling method, but demands thousands of computational analyses (e.g., finite element analysis (FEA), crash analysis, etc.). To relieve the computational burden, other sampling methods have been developed, such as quasi-MCS (Niederreiter and Spanier 2000; Sobol 1998), (adaptive) importance sampling (Engelund and Rackwitz 1993; Melchers 1989; Bucher 1988; Wu 1994), directional sampling (Bjerager 1988), etc. Nevertheless, sampling methods are considerably expensive. Thus, it is often used for verification of probability analysis when alternative methods are used.

The idea of the expansion method is to estimate statistical moments of system responses with a small perturbation to simulate input uncertainty. This expansion method includes Taylor expansion (Jung and Lee 2002), perturbation method (Kleiber and Hien 1992; Rahman and Rao 2001), Neumann expansion method (Yamazaki and Shinozuka 1988), etc. Taylor expansion and perturbation methods require high-order partial sensitivities to maintain good accuracy. The Neumann expansion method employs Neumann series expansion of the inverse of random matrices, which requires an enormous amount of computational effort. In summary, all expansion methods could become computationally inefficient or inaccurate when the number or the degree of input uncertainty is higher. Moreover, as it requires high-order partial sensitivities of system responses, it may not be practical for large-scale engineering applications.

The MPP-based method has been widely used to perform reliability analysis. Rotationally invariant reliability index is introduced through a nonhomogeneous transformation (Hasofer and Lind 1974). Probability analysis can be conducted in two different ways: performance-level (G-level) (Hasofer and Lind 1974) and probability-level (P-level) (Youn et al. 2004a; Du and Chen 2002) methods. It has been found that the P-level method is more efficient and stable than the G-level method (Youn et al. 2004a). However, the MPP-based method requires the first-order sensitivities of system responses. Moreover, it could generate relatively large error caused by some nonlinearity of system response and is not suitable for multiple MPP problems.

The response surface approximate method (Myers and Montgomery 1995) is often used with MCS to perform reliability analysis. A true system response is approximated based on limited design of experiment (DOE) samples and a response surface approximation method. Once the response surface is constructed, the MCS can be used for reliability analysis without extra expense except for the DOE samples. The accuracy of this method greatly depends on the accuracy of response surface. Besides, response surface method is not suitable for high-dimensional problems because of a course of dimensionality.

The approximate integration method is a direct approach to estimate the probability density function (PDF) (or statistical moments) through numerical integration. Numerical integration can be done in the input uncertainty domain (Rahman and Xu 2004; Seo and Kwak 2003) or the output uncertainty domain (Youn et al. 2005). Recently, the dimension reduction (DR) method (Rahman and Xu 2004; Xu and Rahman 2004) has been proposed and is known to be a sensitivity-free method. In the univariate DR method (Rahman and Xu 2004), it uses an additive decomposition of the responses that simplifies one multi-dimensional integration to multiple one-dimensional integrations. Generally, it can provide accurate lower moment of system responses such as mean. However, it may produce a relatively large error for the second-order or higher moments of nonlinear system responses. Otherwise, it could be expensive with large number of numerical integration points. In the general DR method (Xu and Rahman 2004), the theoretical error of univariate DR method can be reduced by considering multi-dimensional integrations. However, the computation effort is increased exponentially. Therefore, it is hard to afford a general DR calculation in most engineering applications.

This paper proposes the eigenvector dimension reduction (EDR) method, which is an enhancement of the univariate DR method. It has three technical elements: (1) eigenvector sampling, (2) one-dimensional response approximations, and (3) a stabilized Pearson system. The $2N+1$ and $4N+1$ eigenvector sampling schemes are proposed in the EDR method to resolve correlated and asymmetric random input variables while maintaining high accuracy and efficiency for sensitivity-free probability analysis. The Stepwise Moving Least Squares (SMLS) method is proposed for response approximation. The SMLS method integrates a Moving Least Squares (MLS) method (Youn and Choi 2004) with a stepwise regression scheme (Myers and Montgomery 1995). The one-dimensional response approximation allows the increase of integration points without demanding additional computation. Therefore, the EDR method improves numerical accuracy in calculating the statistical moments with no extra expense other than the eigenvector samples. The stabilized Pearson system is proposed to predict probability density functions (PDFs) of the responses while eliminating singular behavior of the original Pearson system.

In this paper, the univariate DR method is reviewed in Section 2. The EDR method is then developed based on the univariate DR method with three new technical components in Section 3. Finally, six examples demonstrate that the EDR method makes considerable improvements from the perspective of accuracy, efficiency, and stability compared with the univariate DR method and some traditional probability analysis methods.

2 Review of univariate dimension reduction method

2.1 Dimension reduction method using additive decomposition

In general, statistical moments (or PDF) of system responses (e.g., fatigue life, corrosion, injury metrics), $Y(\mathbf{X})$, can be calculated as

$$\mathrm{E}\{Y^m(\mathbf{X})\} = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} Y^m(\mathbf{x})\cdot f_{\mathbf{X}}(\mathbf{x})\cdot d\mathbf{x}, \quad m = 0,1,2,\cdots \tag{1}$$

In (1), a major challenge is a multi-dimensional integration over the entire random input ($\mathbf{X}$) domain. To resolve this difficulty, the univariate DR method uses an additive decomposition (Rahman and Xu 2004) that converts a multi-dimensional integration in (1) into multiple one-dimensional integrations. The additive decomposition, Y_a, is defined as

$$\begin{aligned} Y(X_1,\ldots,X_N) &\cong Y_a(X_1,\ldots,X_N) \\ &= \sum_{j=1}^{N} Y\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right) \\ &\quad -(N-1)Y(\mu_1,\ldots,\mu_N) \end{aligned} \tag{2}$$

To validate the use of the additive decomposition, the error incurred because of its use in determining the statistical moments must be small. To accomplish this, the Taylor series expansion of the actual function, $Y(\mathbf{x})$, in (3) is compared to the expansion of the additive decomposition, $Y_a(\mathbf{x})$, in (4).

$$\begin{aligned} I[Y(\mathbf{x})] = I[Y(\mathbf{0})] &+ \frac{1}{2!}\sum_{i=1}^{N}\frac{\partial^2 Y}{\partial x_i^2}(\mathbf{0})I\left[x_i^2\right] \\ &+ \frac{1}{4!}\sum_{i=1}^{N}\frac{\partial^4 Y}{\partial x_i^4}(\mathbf{0})I\left[x_i^4\right] \\ &+ \frac{1}{2!2!}\sum_{i<j}\frac{\partial^4 Y}{\partial x_i^2\partial x_j^2}(\mathbf{0})I\left[x_i^2x_j^2\right]+\cdots \end{aligned} \tag{3}$$

$$\begin{aligned} I[Y_a(\mathbf{x})] = I[Y(\mathbf{0})] &+ \frac{1}{2!}\sum_{i=1}^{N}\frac{\partial^2 Y}{\partial x_i^2}(\mathbf{0})I\left[x_i^2\right] \\ &+ \frac{1}{4!}\sum_{i=1}^{N}\frac{\partial^4 Y}{\partial x_i^4}(\mathbf{0})I\left[x_i^4\right] \\ &+ \frac{1}{6!}\sum_{i=1}^{N}\frac{\partial^6 Y}{\partial x_i^6}(\mathbf{0})I\left[x_i^6\right]\cdots \end{aligned} \tag{4}$$

It can be seen in (5) that the largest error occurs at the fourth even-order term, producing negligible error. In fact, the error produced by the additive decomposition is less than that of a second-order Taylor expansion method for probability analysis (Rahman and Xu 2004). The accuracy in the use of the additive decomposition is partly because the integration is being performed over a symmetric domain. This results in all of the odd-order terms in the integration to be zero.

$$I[Y(x)] - I[Y_a(x)] = \frac{1}{2!2!}\sum_{i<j}\frac{\partial^2 Y}{\partial x_i^2\partial x_j^2}+\cdots \tag{5}$$

In aid of the additive decomposition, probability analysis of system responses becomes much simpler. For reliability and quality assessment, the m^{th} statistical moments for the responses are considered in (6) as

$$\begin{aligned} E[Y^m(\mathbf{X})] &\simeq E\left[Y_a^m(\mathbf{X})\right] \\ &= E\left\{\left[\sum_{j=1}^{N} Y\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right) - (N-1)\cdot Y(\mu_1,\ldots,\mu_N)\right]^m\right\} \\ &= \int_{-\infty}^{\infty}\left[\sum_{j=1}^{N} Y\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right) - (N-1)\cdot Y(\mu_1,\ldots,\mu_N)\right]^m \cdot f_{X_j}(x_j)\cdot dx_j \end{aligned} \tag{6}$$

Using a binomial formula, (6) can be evaluated by executing one-dimensional integration recursively. In other words, uncertainty of system responses can be evaluated through multiple one-dimensional numerical integrations. So the challenge of the problem still remains how to carry out one dimensional integration effectively. Using numerical integration, one-dimensional integration will be performed with integration weights $w_{j,i}$ and points $x_{j,i}$ using (7).

$$\begin{aligned} &\mathrm{E}\left[\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{Y}^{\mathrm{m}}\left(\mu_1,\ldots,\mu_{\mathrm{j}-1},\mathrm{X}_{\mathrm{j}},\mu_{\mathrm{j}+1},\ldots,\mu_{\mathrm{N}}\right)\right] \\ &\quad\cong \sum_{\mathrm{j}=1}^{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{w}_{\mathrm{j,i}}\mathrm{Y}^{\mathrm{m}}\left(\mu_1,\ldots,\mu_{\mathrm{j}-1},\mathrm{x}_{\mathrm{j,i}},\mu_{\mathrm{j}+1},\ldots,\mu_{\mathrm{N}}\right) \end{aligned} \tag{7}$$

The number of integration points determines computational efficiency of the univariate DR method. In general, the univariate DR method uses $(n-1)\times N+1$ integration points where N is the number of input random parameters and n is the integration points along each random variable. It is suggested in the proposed Eigenvector Dimension Reduction (EDR) method that n must be maintained at 3 or, at most, 5, for large-scale engineering design problems.

2.2 One-dimensional numerical integration in the dimension reduction method

The DR method suggests the use of a moment-based quadrature rule (Rahman and Xu 2004; Xu and Rahman 2004) to perform the one-dimensional numerical integration in (6). Integration points and weights can be obtained by solving a linear system equation that requires the statistical information of the input parameters. The linear relationship is made between low- and high-order moments of the random input variables, as shown in (8).

$$\begin{bmatrix} \mu_{j,n-1} & -\mu_{j,n-2} & \mu_{j,n-3} & \cdots & (-1)^{n-1}\mu_{j,0} \\ \mu_{j,n} & -\mu_{j,n-1} & \mu_{j,n-2} & \cdots & (-1)^{n-1}\mu_{j,1} \\ \mu_{j,n+1} & -\mu_{j,n} & \mu_{j,n-1} & \cdots & (-1)^{n-1}\mu_{j,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{j,2n-2} & -\mu_{j,2n-3} & \mu_{j,2n-4} & \cdots & (-1)^{n-1}\mu_{j,n-1} \end{bmatrix} \begin{bmatrix} r_{j,1} \\ r_{j,2} \\ r_{j,3} \\ \vdots \\ r_{j,n} \end{bmatrix} = \begin{bmatrix} \mu_{j,n} \\ \mu_{j,n+1} \\ \mu_{j,n+2} \\ \vdots \\ \mu_{j,2n-1} \end{bmatrix} \tag{8}$$

Here, $\mu_{j,n}$ represents the nth raw moment considering the jth input variable and r is a moment vector. The solution of (8) can be manipulated to produce the resulting integration points and the weights.

2.3 Remarks on the Dimension Reduction (DR) method

A different statistical moment formula from (6) was developed in Xu and Rahman 2004 by replacing Y^m with Z, expressed as

$$\begin{aligned} \mathrm{E}_Z[Y^m(X)] &= \mathrm{E}[Z] \\ &\simeq \mathrm{E}_Z\left[\sum_{j=1}^{N} Z(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N) - (N-1)\cdot Z(\mu_1,\ldots,\mu_N)\right] \\ &\simeq \mathrm{E}_Z\left[\sum_{j=1}^{N} Y^m(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N) - (N-1)\cdot Y^m(\mu_1,\ldots,\mu_N)\right] \end{aligned} \tag{9}$$

Thus, it may eliminate a complicated process using a binomial formula. However, it is found that this formula could lead to larger error caused by the replacement of the power term before the additive decomposition. For example, when m=2, (6) gives the following formula as

$$\begin{aligned} \mathrm{E}\left[Y^2(X)\right] \simeq{}& \mathrm{E}\left[\sum_{j=1}^{N} Y^2\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right)\right] \\ &+ (N-1)^2 Y^2(\mu_1,\ldots,\mu_N) - 2(N-1)Y(\mu_1,\ldots,\mu_N) \\ &\mathrm{E}\left[\sum_{j=1}^{N} Y\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right)\right] \end{aligned} \tag{10}$$

However, (9) proposed for simplicity gives the different formula as

$$\begin{aligned} &\mathrm{E}_Z\left[Y^2(X)\right] \\ &\simeq \mathrm{E}_Z\left[\sum_{j=1}^{N} Y^2\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right)\right] \\ &- (N-1)\cdot Y^2(\mu_1,\ldots,\mu_N) \end{aligned} \tag{11}$$

A distinctive difference is found between two formulae as

$$\begin{aligned} \mathrm{E}\left[\mathrm{Y}^2(\mathrm{X})\right] - \mathrm{E}_Z\left[\mathrm{Y}^2(\mathrm{X})\right] \simeq{}& \left(\mathrm{N}^2-\mathrm{N}\right)\mathrm{Y}^2(\mu_1,\ldots,\mu_\mathrm{N}) \\ &- 2(\mathrm{N}-1)\mathrm{Y}(\mu_1,\ldots,\mu_\mathrm{N}) \\ &\mathrm{E}\left[\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{Y}\left(\mu_1,\ldots,\mu_{\mathrm{j}-1},\mathrm{X_j},\mu_{\mathrm{j}+1},\ldots,\mu_\mathrm{N}\right)\right], \end{aligned} \tag{12}$$

where the difference is an additional error induced by the different formulation in Xu and Rahman 2004.

3 Eigenvector Dimension Reduction (EDR) method

The univariate DR method is enhanced by incorporating three technical components: (1) eigenvector sampling, (2) the Stepwise Moving Least Squares (SMLS) method for efficient and accurate numerical integration, and (3) a stabilized Pearson system for PDF generation. Although the univariate DR method gives reasonably good results for probability analysis, the EDR method attempts to resolve the disadvantages of the DR method addressed in Section 1.

3.1 Eigenvector sampling

With the additively decomposed function in (2), the challenge of probability analysis still remains how to carry out one dimensional integration efficiently and accurately. Accuracy for probability analysis can be increased as the

number of integration points in recursive one-dimensional integration. However, the increase of integration points makes probability analysis prohibitively expensive for large-scale applications. To achieve both accuracy and efficiency in probability analysis, an eigenvector sampling scheme selects sample points along the eigenvectors of the covariance matrix ($\boldsymbol{\Sigma}$) of the system input random parameters ($\mathbf{X}$), and then one dimensional response surface (Section 3.2) will be created using the response values at the samples. The primary reason to choose samples along the eigenvectors is because the eigenvectors and eigenvalues contain information for statistical correlation and variation.

The eigenvector sampling scheme assist finding the samples using the eigenvectors and eigenvalues of the covariance of the system input random parameters. For efficiency, the EDR method employs either two (n=2) or four (n=4) samples along each eigenvector excluding the sample at the design point, depending on nonlinearity of the system responses. For N number of random variables, the EDR method demands $2N$+1 or $4N$+1 samples. To obtain the eigenvectors and eigenvalues, an eigenvalue problem for the covariance of the system input random parameters $\mathbf{X}$ can be formulated as

$$\boldsymbol{\Sigma}\mathbf{X} = \mathbf{1}\mathbf{X} \tag{13}$$

where $\mathbf{X}$ and λ are eigenvectors and eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$. The covariance matrix with the N random input variables is defined as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1N} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & \Sigma_{2N} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & \Sigma_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma_{N1} & \Sigma_{N2} & \Sigma_{N3} & \cdots & \Sigma_{NN} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1N} \\ \Sigma_{21} & \sigma_2^2 & \Sigma_{23} & \cdots & \Sigma_{2N} \\ \Sigma_{31} & \Sigma_{32} & \sigma_3^2 & \cdots & \Sigma_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma_{N1} & \Sigma_{N2} & \Sigma_{N3} & \cdots & \sigma_N^2 \end{bmatrix},$$

where the covariance between the input variables X_i and X_j can be defined as

$$\mathrm{Cov}(X_i, X_j) = \sum\nolimits_{ij} = \mathrm{E}\left[(X_i - \mu_i)\left(X_j - \mu_j\right)\right]$$

and μ_i and μ_j are the means of X_i and X_j. According to the definition, the covariance is symmetric with $\Sigma_{ij}=\Sigma_{ji}$ and σ_i^2 is the variance of any random variable X_i.

Depending on the random properties of system inputs, four different types of the random properties can be defined as: (1) uncorrelated and symmetric, (2) correlated and symmetric, (3) uncorrelated and asymmetric, and (4) correlated and asymmetric. The $2N$+1 eigenvector sampling scheme is first considered here. For any circumstance, the $2N$+1 eigenvector samples will be found at

$${}^1\mathbf{V}_i = \boldsymbol{\mu} - k\sqrt{\lambda_i}\mathbf{X}_i' \text{ and } {}^2\mathbf{V}_i = \boldsymbol{\mu} + k\sqrt{\lambda_i}\mathbf{X}_i' \tag{14}$$

where $\mathbf{X}_i'$ and λ_i are the ith eigenvector and eigenvalue, and k determines a sample location along the eigenvectors. The locations of the eigenvector samples dictate accuracy of one-dimensional response approximations. Subsequently, accuracies of one-dimensional response approximations determine accuracy of one-dimensional numerical integrations and, eventually, probability analysis in the EDR method.

If the k is too large, the accuracy of one-dimensional response approximations will be degraded on the inner side of two eigenvector samples ${}^1\mathbf{V}_i'$ and ${}^2\mathbf{V}_i'$; on the other hand, if k is too small, the accuracy of the response approximations will be descended on the outer side of eigenvector samples because of an extrapolation. As the response approximation is involved, it is nearly impossible to determine the optimum location (k) of the eigenvector samples with a reasonable justification. Thus, a parametric study is performed by using a set of mathematical examples, and two facts are observed: (1) the accuracy of the EDR appears to be the best with k=[2.5~3.5]; (2) the accuracy is nearly insensitive with any k value in the range. So, this paper uses k=3 for eigenvector sampling. For the different types of the system input random properties, the eigenvector samples are found as follows:

a. Uncorrelated and symmetric

If all random variables are statistically uncorrelated, all off-diagonal terms in the covariance matrix become zero. In this case, the eigenvectors are simply the original random variable axes. The eigenvector samples are obtained along the original random vectors at ${}^1\mathbf{V}_i = \boldsymbol{\mu} - 3\sqrt{\lambda_i}\mathbf{X}_i' = \boldsymbol{\mu} - 3\sigma_i\mathbf{X}_i'$ and ${}^2\mathbf{V}_i = \boldsymbol{\mu} + 3\sqrt{\lambda_i}\mathbf{X}_i' = \boldsymbol{\mu} + 3\sigma_i\mathbf{X}_i'$, where $\mathbf{X}_i'$ is the ith eigenvector where all elements are zero except the ith element is one.

b. Correlated and symmetric

If some random variables are statistically correlated, the eigenvector samples are obtained at ${}^1\mathbf{V}_i = \boldsymbol{\mu} - 3\sqrt{\lambda_i}\mathbf{X}_i'$ and ${}^2\mathbf{V}_i = \boldsymbol{\mu} + 3\sqrt{\lambda_i}\mathbf{X}_i'$ along the eigenvectors of the eigenvalue problem in (13).

c. Uncorrelated and asymmetric

If all random variables are statistically uncorrelated but asymmetrically distributed, the eigenvectors are still same as the original random variable axes. To facilitate the eigenvector sampling for asymmetrically distributed random input parameters, the random parameters are transformed into a standard-normally distributed random parameter (U), such as $\mathbf{T}$: $\mathbf{X}_i \rightarrow \mathbf{U}_i$ (Rackwitz and Fiessler 1978). The eigenvector samples are similarly obtained along

the eigenvectors in the transformed space at ${}^1\mathbf{U}_i = -3\mathbf{U}_i^{'}$ and ${}^2\mathbf{U_i} = +3\mathbf{U_i^{'}}$, where $\mathbf{U}_i^{'}$ is the i^{th} eigenvector where all elements are zero except the i^{th} element is one. Then, two eigenvector samples ${}^1\mathbf{V}_i$ and ${}^2\mathbf{V}_i$ will be found from ${}^1\mathbf{U}_i$ and ${}^2\mathbf{U}_i$ through the inverse transformation, $\mathbf{T}^{-1}$.

d. Correlated and asymmetric

If some random variables are both correlated and with asymmetric distributions, the eigenvectors of the covariance matrix of the system input random parameters must be first obtained, as illustrated in the part b. For the random variables with correlated and asymmetric distributions, the eigenvector samples will be chosen along the eigenvectors through the transformation given in the part c.

The $2N$+1 eigenvector samples are illustrated for the four different cases shown in Fig. 1 and the samples are used for constructing one-dimensional response approximation using the SMLS method in the following section. To enhance numerical accuracy in probability analysis, the $4N$+1 eigenvector samples will be selected with two extra samples located at ${}^3\mathbf{V}_i = \boldsymbol{\mu} - 1.5\sqrt{\lambda_i}\mathbf{X}_i^{'}$ and ${}^4\mathbf{V}_i = \boldsymbol{\mu} + 1.5\sqrt{\lambda_i}\mathbf{X}_i^{'}$.

3.2 Stepwise Moving Least Squares (SMLS) Method for numerical integration

The moving least square (MLS) method (Youn and Choi 2004) is improved by a stepwise selection of basis functions, referred to as the Stepwise Moving Least Squares (SMLS) method. The optimal set of basis terms is adaptively chosen to maximize numerical accuracy by screening the importance of basis terms. This technique is exploited for approximating the additively decomposed one-dimensional integrand in (6). The idea of a stepwise selection of basis functions comes from the stepwise regression method (Myers and Montgomery 1995). The SMLS method for one-dimensional response approximation proceeds in the following steps:

Step 1. Define a pool of basis elements and forced basis elements out of the pool. Set the total number of basis elements, *nb*, and sub-domain counter, m=0.

Step 2. Define the *m*th sub-domain surrounded by *nb* neighboring samples.

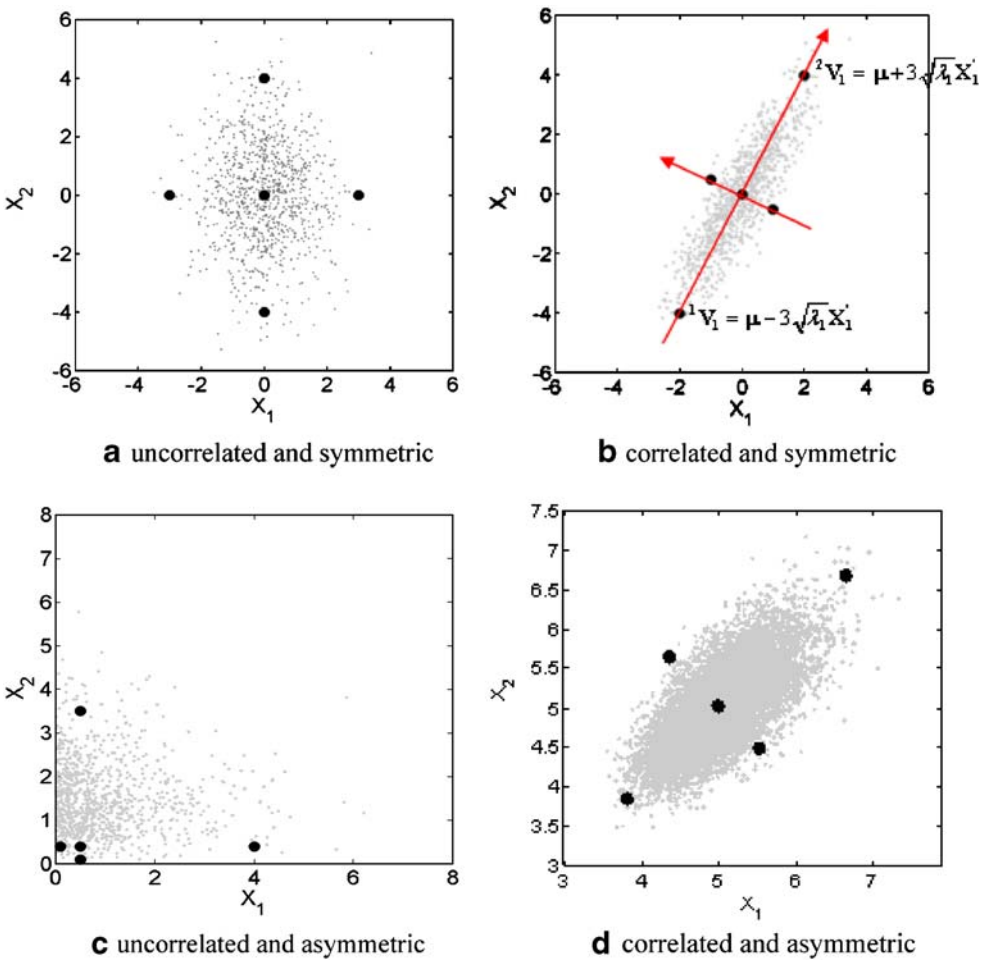

Fig. 1 Eigenvector samples for EDR method

Step 3. Find *nt* training points in all sub-domains, where training points are defined in the middle of every two samples.

Step 4. Approximate responses at training points using the MLS method (see Appendix) as

$\widehat{\mathbf{Y}}(\mathbf{d}) = \mathbf{h}^T(\mathbf{d})\mathbf{M}^{-1}(\mathbf{d})\mathbf{B}(\mathbf{d})\mathbf{Y}$ where $\mathbf{M} = \mathbf{H}^T\mathbf{W}(\mathbf{d})\mathbf{H}$, $\mathbf{B} = \mathbf{H}^T\mathbf{W}(\mathbf{d})$ where $\mathbf{W}$ is the weight matrix and $\mathbf{H}$ is the basis matrix

Step 5. Filter the basis elements adaptively in the *m*th subdomains by ranking the magnitudes of the coefficients. The basis element with the maximum coefficient will be selected and add to the forced basis elements as the current basis elements. This process will be repeated until the total number of required basis elements (*nb*) is reached.

Step 6. Set $m=m+1$ and go to Step 2 if $m \leq ns$, where *ns* is the total number of subdomains. Otherwise go to Step 7.

Step 7. Construct one-dimensional response surface using sample responses.

3.2.1 Example of SMLS Method

As the objective is to approximate one-dimensional response accurately, a highly nonlinear one-dimensional response example is used to show accuracy of the SMLS. For the purpose of the EDR method, the response would be treated as the integrand used in the EDR method. The exact response is explicitly expressed as

$$Y(X) = X^2(2 + \sin(2X))/4, \quad 1 \leq X \leq 7 \tag{15}$$

Six subdomains ($m=6$) are defined and six training points ($nt=6$) are used. Seven basis terms ($nb=7$) are used where two (1 and X) are the forced basis terms. Including the forced basis terms, the pool of basis terms are $\{1, X, X^2, X^3, X^4, X^5, X^6, X^7, \sin X, \cos X, \exp(X)\}$. In addition to the ordinary polynomial basis, the sinusoidal and exponential basis terms are used because they are good for nonlinear representation. For example, at $X=4$ the selected basis terms are [1, X, sin$X$, cos$X$, X^2, X^3, X^4] with the corresponding coefficients [−124.2285, 617.9624, −151.2387, −97.4442, −382.9456, 74.3153, −4.4639]. As shown in Fig. 2, the SMLS method approximates the response very accurately in aid of the adaptive selection of basis elements in different subdomains. In Table 1, the normalized error is measured as

$$e = \frac{1}{nt}\sum_{i=1}^{nt}\left(\frac{\left(\widehat{y}_i - y_i^t\right)^2}{\left(y_i^t\right)^2}\right), \tag{16}$$

where the total trial points, $nt=61$. $\widehat{y}_i$ and y_i^t are approximate and true responses, respectively, at the *i*th trial points.

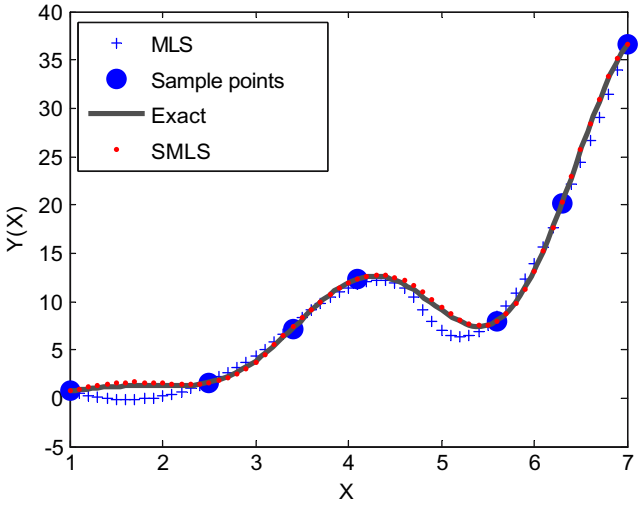

Fig. 2 Response approximation using stepwise moving least square method

In general, for numerical calculation of statistical moments, the integration domain is relatively small because the contribution of PDF is negligible for $X<\mu-6\sigma$ and $X>\mu+6\sigma$. For some bounded random distributions, such as uniform and beta distributions, the integration domain will be limited from lower to upper bounds. Therefore, although probability analysis is applied for nonlinear system responses, they are less nonlinear in a local region than those in a global region. Thus, the SMLS method can approximate system responses very accurately in the integration domain. The SMLS method allows the increase in the number of numerical integration points without requiring actual simulations (or experiments) for system response evaluations. Responses at all integration points are approximately obtained from approximated one-dimensional responses, $\widehat{Y}(x_{j,i})$, instead of system responses, $Y(x_{j,i})$ through actual system evaluations, as shown in (17).

$$\begin{aligned} &\mathrm{E}\left[\sum_{j=1}^{N} Y\left(\mu_1,\ldots,\mu_{j-1},X_j,\mu_{j+1},\ldots,\mu_N\right)\right] \\ &\cong \sum_{j=1}^{N}\sum_{i=1}^{n} w_{j,i}\, Y\left(\mu_1,\ldots,\mu_{j-1},x_{j,i},\mu_{j+1},\ldots,\mu_N\right) \\ &\cong \sum_{j=1}^{N}\sum_{i=1}^{n} w_{j,i}\, \widehat{y}\left(\mu_1,\ldots,\mu_{j-1},x_{j,i},\mu_{j+1},\ldots,\mu_N\right) \end{aligned} \tag{17}$$

Thus, a large number ($n=20\sim30$) of integration points can be used to increase numerical accuracy in assessing statistical moments of the responses without requiring actual

Table 1 Normalized errors of the MLS and SMLS

Method	MLS	SMLS
Normalized error	3.2200	0.0130

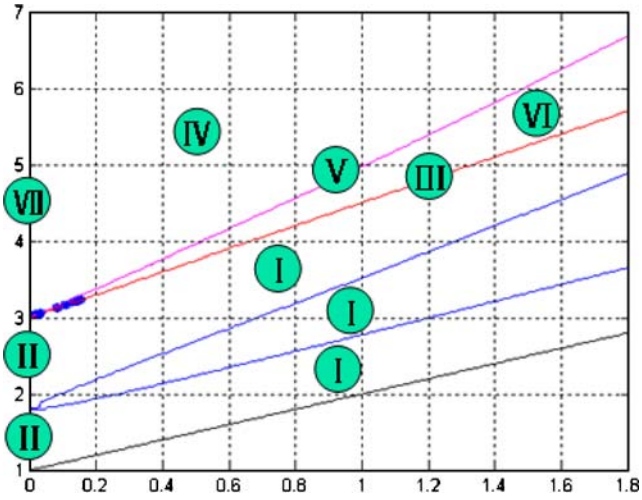

Fig. 3 Pearson curve

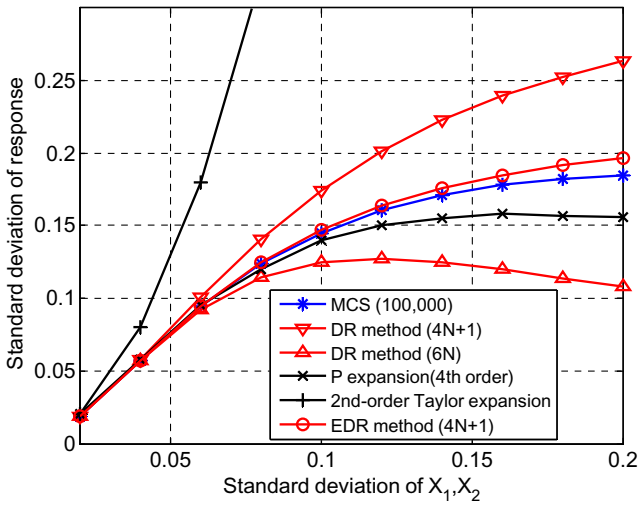

Fig. 5 Standard deviations of response with different input standard deviations

system evaluations. So, numerical accuracy in estimating statistical moments is improved considerably while high efficiency is remained, as only $2N+1$ or $4N+1$ simulations or experiments are required.

a. Moment-based quadrature rule In the DR method, a moment-based quadrature rule (Rahman and Xu 2004) was proposed for one-dimensional numerical integration because of its good accuracy and efficiency, compared with other integration methods. However, it may still produce a relatively large error for the second order or higher moments of nonlinear responses as will be shown in the later examples. In the EDR method, however, as large amount of integration points could be used without actual simulations (or experiments) for system response evaluations, accuracy of moment-based quadrature rule could be improved substantially. Thus, moment-based quadrature rule could still be used in the EDR method. However, moment-based quadrature rule could have two problems as

1) The number of integration points should be predetermined. So it is hard to decide an optimal number of the points to maximize accuracy of the EDR method for probability analysis.
2) Larger amount of integration points could result in a singular moment matrix in (8), so it may fail to find the corresponding integration points. Specifically, a large number of integration points require the use of higher order statistical raw moments. As the order of the moments is increased, the matrix in (8) becomes singular because of the higher-order moments asymptotically approaching zero.

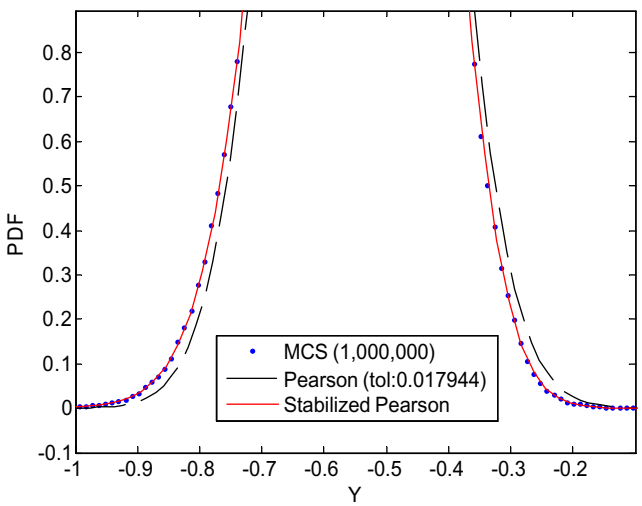

Fig. 4 Comparison of PDF

b. Adaptive Simpson rule This paper suggests an adaptive Simpson rule (Yamazaki and Shinozuka 1988) as an alternative integration method. It gives more freedom on selection of probability distribution types for system input random variables. Adaptive Simpson's rule uses an adaptive way to estimate the error from calculating a definite integral using Simpson's rule. If the error is larger than a user-specified tolerance, the integration interval is divided into subintervals, and Simpson's rule is applied to each subinterval. The adaptive Simpson rule generally demands

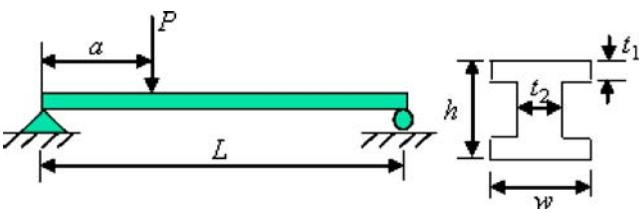

Fig. 6 Loading condition and structure of an I beam

Table 2 Statistical properties of random variables in beam example

Variable	Type	Mean	Std. Dev.	Lower bound	Upper bound	Mode
P	Normal	6,070	200	–	–	–
L	Beta	120	6	100	150	–
a	Uniform	–	–	50	80	–
S	Lognormal	17,0000	4760	–	–	–
h	Triangular	–	–	2.25	2.38	2.30
w	Weibull	2.9665	0.0750	–	–	–
t_1	Normal	0.1600	0.0208	–	–	–
t_2	Lognormal	0.2600	0.0208	–	–	–

a large number of integration points to preserve good accuracy by specifying the tolerance. The SMLS method enables the increase in the number of integration points to as many as possible with no additional computation. Unlike the DR method, the EDR method has no restriction to choose numerical integration schemes, although this paper uses the adaptive Simpson rule for the one-dimension integration.

3.3 A stabilized Pearson system

The Pearson system (Johnson et al. 1995) can be used to construct the PDF of a random response (Y) based on its first four moments (mean, standard deviation, skewness, and kurtosis). The detail expression of the PDF can be achieved by solving the differential equation as

$$\frac{1}{p(Y)}\frac{\mathrm{d}p(Y)}{\mathrm{d}Y} = -\frac{a+Y}{c_0 + c_1 Y + c_2 Y^2},$$

where a, c_0, c_1, and c_2 are four coefficients determined by the first four moments of the random response (Y) and expressed as

$$c_0 = (4\beta_2 - 3\beta_1)(10\beta_2 - 12\beta_1 - 18)^{-1}\mu_2$$
$$a = c_1 = \sqrt{\beta_1}(\beta_2 + 3)(10\beta_2 - 12\beta_1 - 18)^{-1}\sqrt{\mu_2},$$
$$c_2 = (2\beta_2 - 3\beta_1 - 6)(10\beta_2 - 12\beta_1 - 18)^{-1}$$

where β_1 is the square of skewness (x-axis in Fig. 3), β_2 is the kurtosis (y-axis in Fig. 3), and μ_2 is the variation. The

Table 3 Comparison of statistical moments

	Mean	Std. Dev.	Skewness	Kurtosis
MCS (1,000,000)	−49883	12961	0.0083	3.1479
EDR ($4N$+1)	−49860	12815	0.0050	2.9840
Error (%)	0.0460	1.1312	39.9880	5.2077

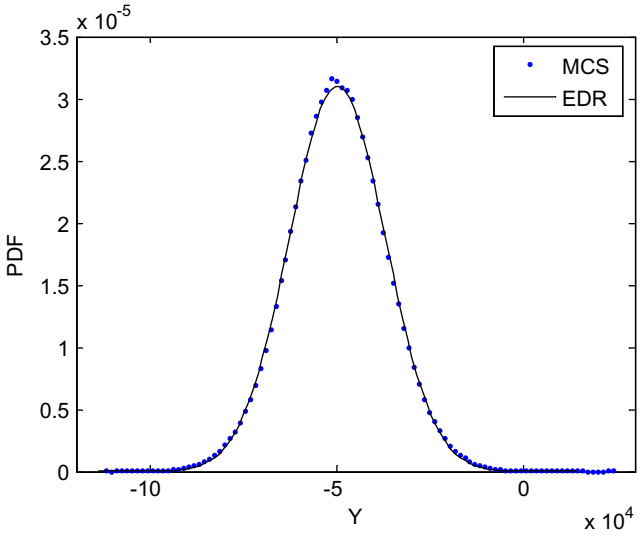

Fig. 7 PDFs using the EDR and MCS method

mean value is always treated as 0 in the Pearson system, and later it can be shifted to the true mean value once the differential equation is solved. Basically, the differential equation can be solved based on the different assumptions of the four coefficients a, c_0, c_1, and c_2. For example, if $c_1 = c_2 = 0$, this equation can be solved with a normal distribution, which corresponds to [0, 3] point in Fig. 3, and the type 1 in Pearson system corresponds to both roots of $c_0 + c_1 Y + c_2 Y^2$ being real. For more detail information, readers can refer to Johnson et al. 1995.

Generally, there are seven distribution types in the Pearson system based on the four coefficients, and among some types, subtypes are present. Normally, PDF can be successfully constructed based on the first four moments. However, the Pearson system can fail to construct the PDF, especially when the statistical moments in the Pearson

Table 4 Properties of design and random variables of vehicle side impact model

Random Variables	Distr. Type	Std Dev.	d^L	d	d^U
X_1 (B-pillar inner)	Normal	0.050	0.500	1.000	1.500
X_2 (B-pillar reinforce)	Normal	0.050	0.500	1.000	1.500
X_3 (Floor side inner)	Normal	0.050	0.500	1.000	1.500
X_4 (Cross member)	Normal	0.050	0.500	1.000	1.500
X_5 (Door beam)	Normal	0.050	0.500	1.000	1.500
X_6 (Door belt line)	Normal	0.050	0.500	1.000	1.500
X_7 (Roof rail)	Normal	0.050	0.500	1.000	1.500
X_8 (Mat. B-pillar inner)	Normal	0.006	0.192	0.300	0.345
X_9 (Mat. Floor side inner)	Normal	0.006	0.192	0.300	0.345
X_{10} (Barrier height)	Normal	10.0	–[a]	–[a]	–[a]
X_{11} (Barrier hitting)	Normal	10.0	–[a]	–[a]	–[a]

[a] 10th and 11th random variables are not regarded as design variables

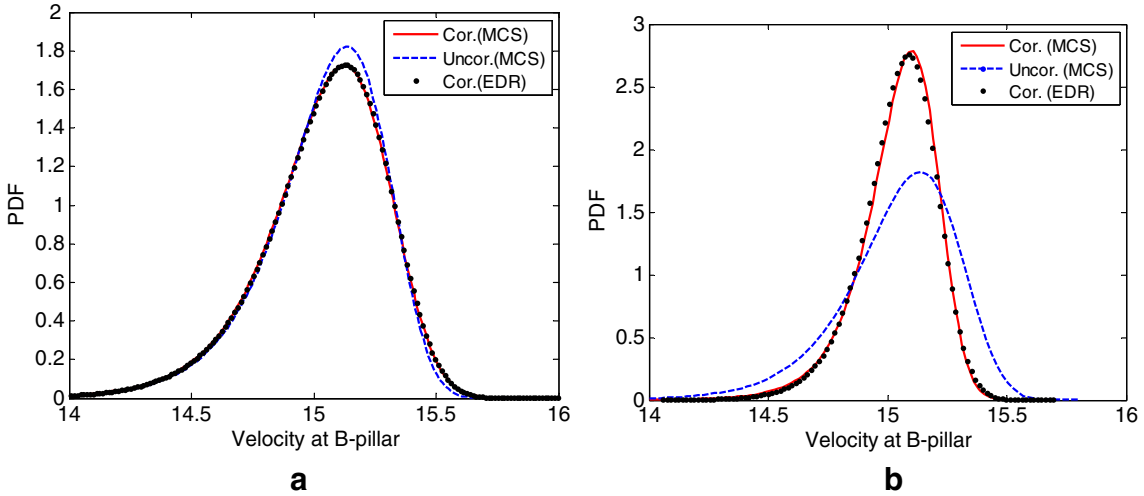

Fig. 8 PDF comparison of system response with correlation. **a** PDF comparison of system response ($\rho_{3,7} = 0.8, \rho_{5,6} = 0.7, \rho_{9,10} = 0.4$). **b** PDF comparison of system response ($\rho_{10,11} = 0.7$)

curve fall into the region that several distribution types merge, as shown in Fig. 3. The horizontal axis is for the square of skewness (β_1) and vertical axis is for the kurtosis (β_2). The solid dots stand for the locations having an instability problem while constructing the PDF. The trouble lies at the calculation of coefficients of a specific distribution type, which results in a numerical instability.

For the distributions (type II, III, V, VII) with an equality condition, it is rare that statistical moments meet the condition tightly. To resolve instability of the Pearson system, the condition is relaxed with a tolerance bound. In this study, 0.001 is used for the tolerance bound. For instance, the PDF should belong to type 6 based on the first four moments (−0.5491, 0.1085, −0.1573, 3.0464). However, numerical singularity is met because of larger numbers of n_1=3273.5 and n_2=−3930.2 in type 6 as

$$p(Y) = K(a_1 - Y)^{n_1}(a_2 - Y)^{n_2}, \qquad Y < a_2$$

The Pearson system fails to calculate the coefficient, K, as $p(Y)$ approaches $0 \cdot \infty$. By relaxing the tolerance bounds, type 3 can be selected, but the singularity problem still remains. Finally, a normal distribution is selected to approximate the PDF by increasing the tolerance value to 0.0118. However, as shown in Fig. 4, the Pearson system produces a noticeable error, compared to MCS with one million samples.

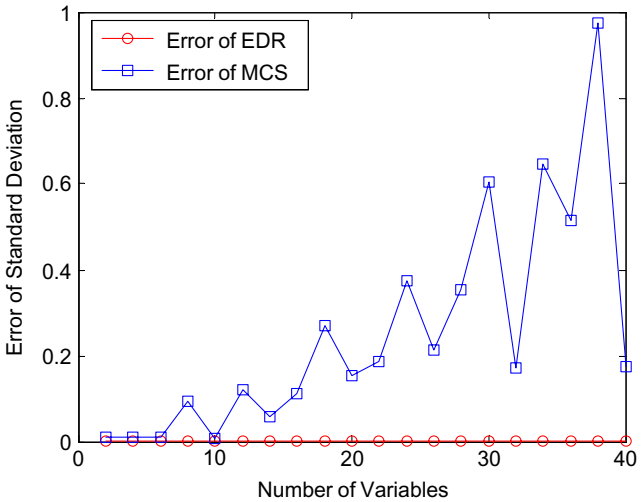

Fig. 9 Comparison of errors using MCS and the EDR method

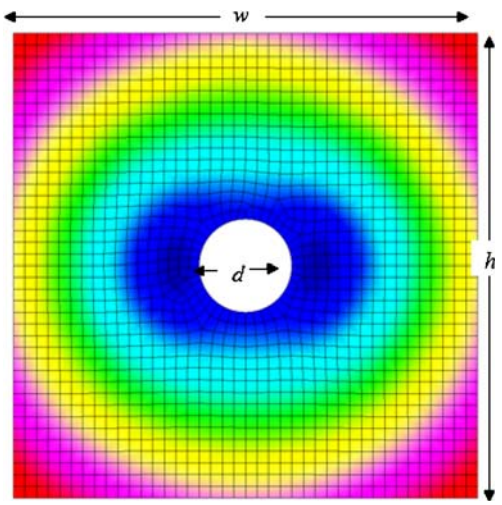

Fig. 10 Plate FE model

Table 5 Random properties in plate model

Random Variable	Mean	Standard deviation	Distribution type
h (shape)	500.0	25.0	Normal
w (shape)	500.0	25.0	Normal
d (shape)	100.0	5.0	Normal

In the EDR method, a stabilized Pearson system is proposed to avoid instability without relaxing tolerances. Two PDFs are generated by fixing the first three statistical moments and slightly increasing or decreasing the original kurtosis until two PDFs are successfully constructed. Then these two PDFs are used to approximate the PDF with original kurtosis. Suppose that the Pearson system fails to construct a PDF with the first four moments (m_1, m_2, m_3, m_4). Detail procedures follow as

Step 1: The first three moments are kept constant and gradually decrease the m_4 by a small decrement (Δm_4=0.01) until a PDF can be successfully constructed. $p_1(Y)$ and $m_{4,1}$ are denoted as the PDF and the corresponding kurtosis value, respectively.

Step 2: The first three moments are kept constant and gradually increase the m_4 by a small increment (Δm_4=0.01) until a PDF can be successfully constructed. $p_2(Y)$ and $m_{4,2}$ are denoted as the PDF and the corresponding kurtosis value, respectively.

Step 3: To build the PDF over the entire domain of the random response (Y), the response domain is discretized as Y_i, i=1 to l. At every value Y_i, the PDF value, $p(Y_i)$, is obtained using two hyper-PDFs $p_1(Y_i)$ and $p_2(Y_i)$, where they are obtained with the kurtosis $m_{4,1}$ and $m_{4,2}$, respectively. With two hyper-PDF values having the kurtosis $m_{4,1}$ and $m_{4,2}$, the PDF $p(Y_i)$ with the actual kurtosis m_4 ($m_{4,1} < m_4 < m_{4,2}$) can be approximated using SMLS without a singularity. It is found that the PDF $p(Y_i)$ is accurately generated because the amount of the kurtosis perturbation is relatively small.

The perturbation size (Δm_4=0.01) of a kurtosis is used to preserve a relatively small perturbation. Basically, the smaller the difference between $m_{4,1}$ and $m_{4,2}$, the more accurate the PDF approximation for the actual kurtosis m_4. Once the probability distribution ($p(Y)$) for system response is obtained, the distribution is explicitly given. So, reliability is computed through a numerical integration for

$$\text{Reliability} = \int_{-\infty}^{0} p(Y)dY$$

4 Examples

Six examples are used to show the effectiveness of the EDR method. In these examples, either $2N$+1 or $4N$+1 eigenvector samples are used based on the degree of response nonlinearity. A systematic selection of $2N$+1 or $4N$+1 is out of the scope of this paper; however, it will be discussed in the future research.

4.1 Mathematical example

The following nonlinear mathematical example (Rahman and Xu 2004) is used to compare accuracy and efficiency of different probability analysis methods such as the DR method, Taylor expansion, EDR method, etc.

$$G = \exp\left(-\frac{1}{1+100X_1^2+2X_2^2+X_1^2X_2^2}\right),$$

where X_j~Normal(0,σ^2), j=1, 2 are two independent and identically normal distribution. The MCS is conducted with

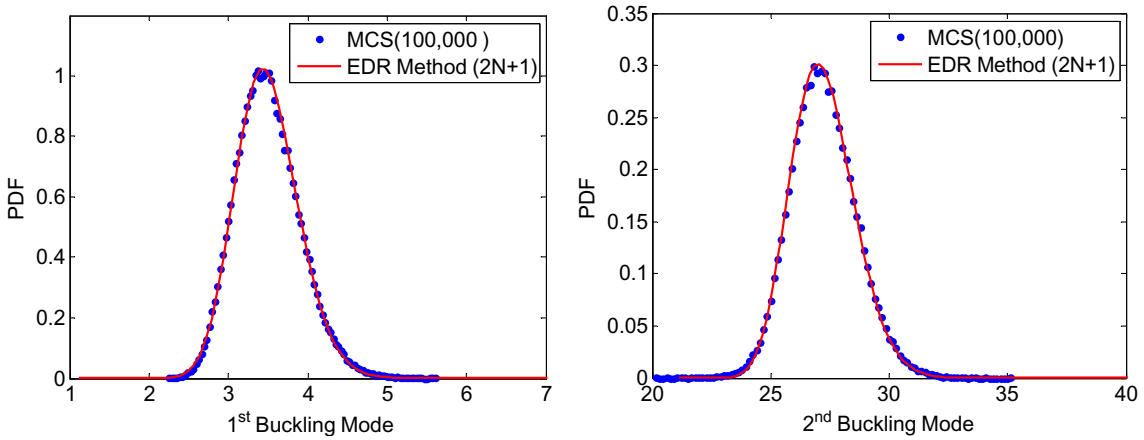

Fig. 11 PDF of Monte Carlo simulation and the EDR method

Table 6 Results of buckling example

Method	First buckling mode				Second buckling mode			
	Mean	Std. Dev.	Skewness	Kurtosis	Mean	Std. Dev.	Skewness	Kurtosis
MCS	3.5031	0.3948	0.3234	3.1692	27.2504	1.3570	0.3209	3.3813
EDR	3.4943	0.3966	0.3038	3.2273	27.2195	1.3374	0.3283	3.1536
Error, %	0.2512	0.4559	6.0606	1.8333	0.1135	1.4425	2.2932	6.7329

100,000 samples, while the DR and EDR methods operate with $4N$+1 samples (five samples in each eigenvector direction). In addition, the DR method is performed with $6N$ samples (six samples in each eigenvector direction) to illustrate the stability problem of the DR method in Section 3.2. In addition, the second-order Taylor series and fourth-order Perturbation method (P-method) are compared. As shown in Fig. 5, the standard deviations of the response are displayed with different standard deviations of the inputs. The EDR method approximates the standard deviation of system response very accurately, compared with the MCS result. However, the approximated standard deviation using the DR method with $4N$+1 samples is overestimated and underestimated with $6N$ samples when the standard deviations of the input variables increase. The fourth-order P-method shows some degree of error in estimating the standard deviations of the response except when input standard deviations are extremely small. And the second-order Taylor expansion shows large error, even the input standard deviation are very small.

4.2 I beam example

An I-beam example (Huang and Du 2006) is used to demonstrate that the EDR method is capable of handling all kinds of input uncertainties such as symmetric, asymmetric, bounded, and unbounded distributions. An I beam is subject to a concentrate force P with a distance a away from the fixed end as shown in Fig. 6. The maximum stress can be expressed as:

$$\sigma_{\max} = \frac{Pa(L-a)h}{2LI},$$

where

$$I = \frac{wh^3 - (w-t_2)(h-2t_1)^3}{12}$$

The beam is safe only if the maximum stress is less than a target value S. A system response can be defined as $Y=\sigma_{\max}-S$ with the safety domain $Y<0$. The uncertainty properties of eight random variables are shown in Table 2. Many distribution types such as normal, lognormal, uniform, etc. are considered as they are commonly met in engineering problems.

$4N$+1 eigenvector samples are used to approximate the eight one-dimensional responses accurately using SMLS. Any numerical integration method can be used to calculate the statistical moments of system response without extra computation effort except for the $4N$+1 eigenvector samples. The statistical moments of system response achieved by the EDR method and 1,000,000 MCS are compared in Table 3. The percentage error of statistical moments is quite small except for the skewness because of the small value. Based on the approximated statistical moments, stabilized Pearson system is used to approximate the PDF of system response. In Fig. 7, the PDF directly achieved from MCS are compared with the one constructed by the stabilized Pearson system. The reliability value calculated by the EDR method and MCS are 99.9943 and 99.9827%, respectively. The results with similar accuracy can be achieved using bivariate DR method in the reference paper (Huang and Du 2006). However, bivariate DR method used 277 function evaluations, which are much more than 33 function evaluations used in the EDR method.

Table 7 Results of component reliability analysis

Reliabilities	FORM	SORM	EDR	MCS
G_1	1	1	1	1
G_2	1	1	1	1
G_3	0.9989	0.9989	0.9989	0.9989
G_4	0.9000	0.9136	0.9026	0.9026
G_5	1	1	1	1
G_6	1	1	1	1
G_7	1	1	1	1
G_8	0.9000	0.8723	0.7140	0.7067
G_9	0.9897	0.9905	0.9905	0.9900
G_{10}	0.9000	0.9025	0.9794	0.9714
Function Eval.	47	47	23	100,000
Sensitivity Eval.	47	47	0	0
Hessian Eval.	10	10	0	0

4.3 Side impact crash problem

Vehicle side impact responses (Youn et al. 2004b) are considered for system performances with statistical corre-

lation. The properties of the design and random variables are shown in Table 4. In this example, the velocity of front door at B-pillar is studied. The system performance can be expressed as

$$G = 16.45 - 0.489X_3X_7 - 0.843X_5X_6 + 0.0432X_9X_{10} - 0.0556X_9X_{11} - 0.000786X_{11}^2$$

Two studies are performed with different set of statistical correlation. In the first study, among these input variables, [X_3, X_7], [X_5, X_6], and [X_9, X_{10}] are assumed to have statistical correlation coefficient 0.8, 0.7, and 0.4, respectively. The EDR method employing 2N+1 (15) analyses is carried out to approximate the first four statistical moments of system performance and construct the PDF. The MCS with 100,000 samples is also carried out for both correlated and uncorrelated cases and the PDFs are correspondingly constructed. Figure 8a displays the results of the first case. In this case, there is only slight difference between correlated and uncorrelated cases. In the second study, as X_{10} and X_{11} are the variables having the maximum variation, they are assumed to have a statistical correlation coefficient $\rho_{X_{10}X_{11}} = 0.7$. Unlike the previous, this case shows the significant effect of statistical correlation on the system response, as shown in Fig. 8b. In both cases, the EDR method can predict the PDF of the system response with statistical correlation very efficiently and accurately.

4.4 Dimension dependency study

A mathematical multi-dimension problem (Rahman and Xu 2004) is considered for the accuracy study with the increase of random variables. In this example, the standard deviation of system response is used for the accuracy study. The input random variables are assumed to follow normal distribution as $X_k \sim$ Normal(0, 1). The multi-dimension problem is expressed as

$$G = \sum_{k=1}^{N} kX_k^2 - X_k$$

First, the analytical solution for the standard deviation of the response G is solved for the increasing number of random variables up to 40. Then, the EDR method with 2N+1 and MCS with 100,000 samples are separately carried out to approximate the standard deviation of response G. Finally, their absolute errors with respect to the analytical solution are calculated, as shown in Fig. 9. This result clearly indicates that accuracy of EDR is independent with the number of random variables. Accuracy of MCS, however, is dependent on the random variables.

4.5 Plate buckling

Buckling is a very important design issue occurring in many engineering disciplines, such as mechanical, aerospace, civil, etc. Structural buckling often leads to catastrophic failures. Thus, it is crucial to accurately estimate the effects of uncertainties inherent in a design upon the critical buckling load. As shown in Fig. 10, a highly non-linear buckling example is considered with three-shape design variable: the height (h) and width (w) of the plate and the hole diameter (d). The statistical information regarding these variables is presented in Table 5. A morphing

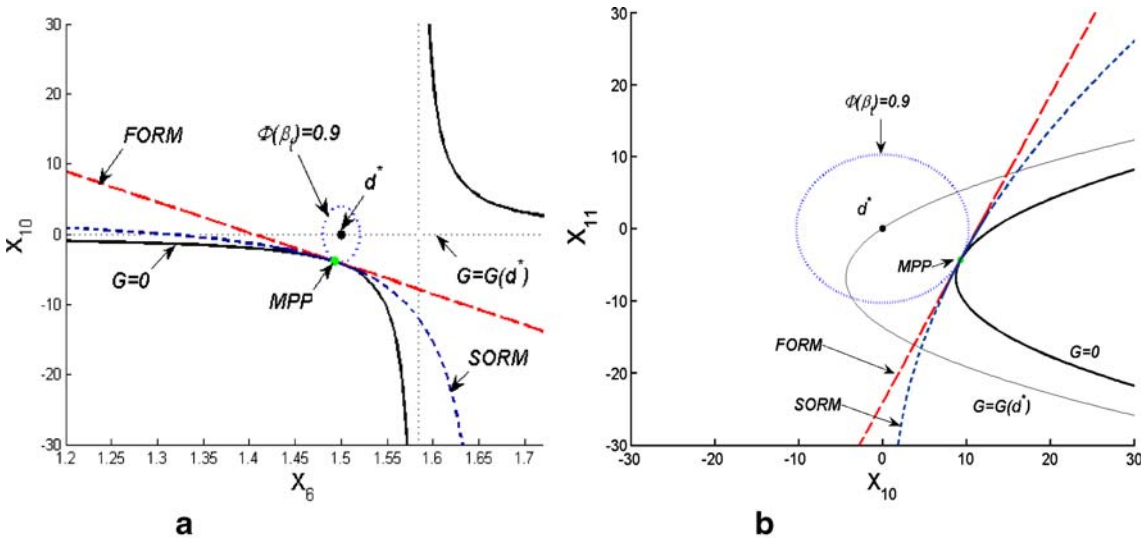

Fig. 12 FORM and SORM reliability analysis in hyper-plane (a): G_8; (b): G_{10}. **a** Failure surface for G_8. **b** Failure surface for G_{10}

technique in the HyperWorks 7.0 software package is used to deal with the shape variables (h, w, and d) in the FEA model. The plate is modeled using plane stress quad4 elements, consisting of 1,681 nodes, 1,571 elements, and 9,798 DOF. A unit load is applied along the top edge of the plate, while the bottom edge of the plate remains fixed in all six directions. The plate is made of Aluminum 6061, where E=67.6 GPa and ν=0.3.

The $2N$ +1 samples (seven buckling analyses) are used for this problem. As shown in Fig. 11, there is a good agreement of statistical moments for the first two buckling modes between the MCS and the EDR method. Moreover, Table 6 displays the resulting statistical information of the response from the EDR method and the MCS with 100,000 samples. It is found that the EDR method performs the uncertainty propagation analysis accurately.

4.6 Comparison of EDR, FORM, and SORM for reliability

In practice, reliability is one of the important engineering metrics to determine how well a product or process is designed. The most common method for reliability analysis is the First-Order Reliability Method (FORM) or Second-Order Reliability Method (SORM) because of their reasonable accuracy and efficiency. This study aims at comparing the EDR method with both FORM and SORM for reliability analysis. For reliability analysis and design, it will be shown that the EDR method is far more efficient than the other two, as one EDR execution takes care of reliability analyses for all constraints without requiring sensitivity of system responses.

The same example used in 4.3 is used here to compare reliability results from the EDR method, FORM, SORM, and MCS at the optimum design using FORM (Youn et al. 2004b). With 90% target reliability, the optimum design point is obtained at $[d^*]^T$=[0.500, 1.327, 0.500, 1.262, 0.623, 1.500, 0.500, 0.345, 0.192, 0.000, 0.000]^T. At the optimum design, reliabilities for ten constraints is verified using three other different methods: SORM, EDR, and MCS, with 100,000 samples. It is found in Table 7 that FORM yields large errors in reliability estimation especially for G_8 and G_{10} constraints. Although the errors can be slightly reduced to some extent using SORM, its accuracy is deficient. However, the EDR method predicts the reliability very accurately. The reason that both FORM and SORM have large error is mainly caused by highly nonlinear responses, as shown in Fig. 12. The dashed and dotted lines show the first-/second-order approximations of failure surfaces used in FORM and SORM for two active constraints G_8 and G_{10} at the optimum design. Inaccurate approximations of failure surfaces lead to the significant errors of FORM and SORM, whereas the EDR method can precisely estimate the failure domains. Nonetheless, it is found that the EDR method is far more efficient than both FORM and SORM.

5 Conclusion

This paper proposes the Eigenvector Dimension Reduction (EDR) method for probability analysis that makes a significant improvement, based on the univariate Dimension Reduction (DR) method. In the univariate DR method, to improve accuracy of probability analysis, a large number of integration points must be involved. Moreover, while increasing the number of integration points, the univariate DR method may become singular and inefficient. To resolve those difficulties, the EDR method is proposed with the three new technical elements: (1) eigenvector sampling, (2) the Stepwise Moving Least Squares (SMLS) method for efficient and accurate numerical integration, and (3) a stabilized Pearson system. First, the 2N+1 and 4N+1 eigenvector sampling schemes are proposed for probability analysis to maintain high accuracy without requiring sensitivity of system performances. Second, the Stepwise Moving Least Squares (SMLS) method is used to accurately approximate the responses, which allow one-dimensional numerical integration with no extra cost other than simulations or experiments at the eigenvector samples. Both moment-based quadrature rule and adaptive Simpson rule can be used for numerical integration. Third, the stabilized Pearson system is proposed to eliminate a singular behavior of the original Pearson system while accurately predicting Probability Density Functions (PDFs) of engineering system performances. In summary, compared with the univariate DR method, the EDR method makes considerable improvements from the perspective of accuracy, efficiency, and stability. The EDR method is far more efficient than traditional probability analysis method such as FORM and SORM, as one EDR execution takes care of reliability analyses for all constraints without requiring sensitivity of system responses. The EDR method could be more accurate than FORM and SORM for highly nonlinear limit state function or limit state function involving inflection points. However, the EDR method may not be good for problems with substantial contribution of high-order mixed terms. In addition, the EDR method may be less accurate than FORM/SORM for large probability levels (e.g., more than 99.9%).

Acknowledgments Research is partially supported by the STAS contract (TCN-05122) (sponsor: the U.S. Army TARDEC) and by the International Joint R&D Program (sponsor: Ministry of Commerce, Industry and Energy, Republic of Korea).

Appendix

Moving Least Squares (MLS) Method

For a given probability level, probabilistic responses at design points can be assessed after following the previous two steps. Any design-of-experiments (DOE) technique can be exercised to obtain design points over the entire design space. Then, the moving least squares (MLS) method can be used to build a stochastic response surface.

For a given probability level, j, the MLS method (Youn and Choi 2004) can be used to approximate stochastic response as

$$\widehat{Y}_j(\mathbf{d}) = \sum_{i=1}^{NB} h_i(\mathbf{d}) a_i(\mathbf{d}) \equiv \mathbf{h}^T(\mathbf{d})\mathbf{a}(\mathbf{d}) \tag{A.1}$$

A stochastic response for the given probability level can be approximated at any design point d as

$$\widehat{Y}_j(\mathbf{d}, \overline{\mathbf{d}}) = \sum_{i=1}^{NB} h_i(\overline{\mathbf{d}}) a_i(\mathbf{d}) = \mathbf{h}^T(\overline{\mathbf{d}})\mathbf{a}(\mathbf{d}), \tag{A.2}$$

where NB is the number of terms in the basis, $h_i(\overline{\mathbf{d}})$ are monomial basis functions evaluated at a set of given sample design points $\overline{\mathbf{d}}$, and $a_i(\mathbf{d})$ are their coefficients, which are functions of the design parameter $\mathbf{d}$.

To compute the coefficient vector $\mathbf{a}(\mathbf{d})$, a weighted residual is defined as

$$\begin{aligned} J &= \sum_{I=1}^{NS} w(\mathbf{d}-\mathbf{d}_I)\left[\widehat{Y}_p(\mathbf{d},\mathbf{d}_I) - \widehat{Y}_p(\mathbf{d}_I)\right]^2 \\ &= \sum_{I=1}^{NS} w(\mathbf{d}-\mathbf{d}_I)\left[\sum_i h_i(\mathbf{d}_I)a_i(\mathbf{d}) - \widehat{Y}_p(\mathbf{d}_I)\right]^2 \end{aligned} \quad \text{or} \tag{A.3}$$

$$J = (\mathbf{H}\mathbf{a} - \mathbf{Y}_p)^T \mathbf{W}(\mathbf{d})(\mathbf{H}\mathbf{a} - \mathbf{Y}_p)$$

where NS is the number of sample points, $w(\mathbf{d}-\mathbf{d}_I)$ is a weight function with a compact support, and

$$\begin{aligned} \mathbf{Y}_j &= [\, Y_j(\mathbf{d}_1) \quad Y_j(\mathbf{d}_2) \quad \ldots \quad Y_j(\mathbf{d}_{NS})\,]^T, \\ \mathbf{H} &= [\, \mathbf{h}(\mathbf{d}_1) \quad \mathbf{h}(\mathbf{d}_2) \quad \ldots \quad \mathbf{h}(\mathbf{d}_{NS})\,]^T, \\ \mathbf{W} &= \mathrm{diag}[\, w(\mathbf{d}_1) \quad w(\mathbf{d}_2) \quad \ldots \quad w(\mathbf{d}_{NS})\,] \end{aligned} \tag{A.4}$$

An appropriate support size at any data point $\mathbf{d}_I$ is selected so that a sufficient number of neighboring data points are included to avoid a singularity. A variable weight over the compact support provides local averaging of the response approximated by the MLS method.

The minimum of the weighted residual, J, by $\partial J/\partial \mathbf{a}=0$, yields the coefficient $\mathbf{a}(\mathbf{d})$ in (A.1) as represented by

$$\mathbf{a}(\mathbf{d}) = \mathbf{M}^{-1}(\mathbf{d})\mathbf{B}(\mathbf{d})\mathbf{Y}_j \text{ where } \mathbf{M} = \mathbf{H}^T\mathbf{W}(\mathbf{d})\mathbf{H}, \quad \mathbf{B} = \mathbf{H}^T\mathbf{W}(\mathbf{d}) \tag{A.5}$$

Substituting (A.5) into (A.1), the approximation $\widehat{Y}_j(\mathbf{d})$ with function information can then be expressed as

$$\widehat{Y}_j(\mathbf{d}) = \mathbf{h}^T(\mathbf{d})\mathbf{M}^{-1}(\mathbf{d})\mathbf{B}(\mathbf{d})\mathbf{Y}_j. \tag{A.6}$$

References

Bjerager P (1988) Probability integration by directional simulation. J Eng Mech 114:1285–1302

Bucher CG (1988) Adaptive sampling—an iterative fast Monte Carlo procedure. Struct Saf 5:119–126

Du X, Chen W (2002) Sequential optimization and reliability assessment method for efficient probabilistic design. Proceedings of the DETC-DAC34127 ASME Design Engineering Technical Conferences, Montreal, Canada, September 2002

Engelund S, Rackwitz R (1993) A benchmark study on importance sampling techniques in structural reliability. Struct Saf 12:255–276

Hasofer AM, Lind NC (1974) Exact and invariant second-moment code format. J Eng Mech 100:111–121

Huang B, Du X (2006) Uncertainty analysis by dimension reduction integration and saddlepoint approximations. J Mech Des 128:26–33

Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions. Wiley, New York

Jung DH, Lee BC (2002) Development of a simple and efficient method for robust optimization. Int J Numer Methods Eng 53:2201–2215

Kleiber M, Hien TD (1992) The stochastic finite element method. Wiley, New York, NY

Lin CY, Huang WH, Jeng MC, Doong JL (1997) Study of an assembly tolerance allocation model based on Monte Carlo simulation. J Mater Process Technol 70:9–16

Melchers RE (1989) Importance sampling in structural systems. Struct Saf 6:6–10

Myers RH, Montgomery DC (1995) Response surface methodology: process and product in optimization using designed experiments. Wiley & Sons, New York, NY, USA

Niederreiter H, Spanier J (2000) Monte Carlo and quasi-Monte Carlo methods. Springer, Berlin

Rackwitz R, Fiessler B (1978) Structural reliability under combined random load sequences. Comput Struct 9:489–494

Rahman S, Rao BN (2001) A perturbation method for stochastic meshless analysis in elastostatics. Int J Numer Methods Eng 50:1961–1991

Rahman S, Xu H (2004) A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics. Probab Eng Mech 19:393–408

Seo HS, Kwak BM (2003) An improved reliability analysis using design of experiments and an application to tolerance design. The 5th world congress of structural and multidisciplinary optimization, Lodo do Jesolo, Italy, May 19–23, 2003

Sobol IM (1998) On quasi-Monte Carlo integrations. Math Comput Simul 47:103–112

Varghese P, Braswell RN, Wang B, Zhang C (1996) Statistical tolerance analysis using FRPDF and numerical convolution. Comput-Aided Des 28(9):723–732

Wu Y-T (1994) Computational methods for efficient structural reliability and reliability sensitivity analysis. AIAA J 32:8

Xu H, Rahman S (2004) A generalized dimension-reduction method for multidimensional integration in stochastic mechanics. Int J Numer Methods Eng 61:1992–2019

Yamazaki F, Shinozuka M (1988) Neumann expansion for stochastic finite element analysis. J Eng Mech 114:1335–1354

Youn BD, Choi KK (2004) A new response surface methodology for reliability-based design optimization. Comput Struct 82:241–256

Youn BD, Choi KK, Du L (2004a) Adaptive probability analysis using an enhanced hybrid mean value (HMV+) method. J Struct Multidiscipl Optim 28(4):319–333

Youn BD, Choi KK, Gu L, Yang R-J (2004b) Reliability-based design optimization for crashworthiness of side impact. J Struct Multidiscipl Optim 26(3–4):272–283

Youn BD, Choi KK, K Yi (2005) Performance moment integration for quality assessment in reliability-based robust design optimization. Mechanics Based Design of Structures and Machines 33:185–213