Struct Multidisc Optim (2009) 38:1–16
DOI 10.1007/s00158-008-0277-9

RESEARCH PAPER

Reliability-based design optimization of problems with correlated input variables using a Gaussian Copula

Yoojeong Noh · K. K. Choi · Liu Du

Received: 14 July 2007 / Revised: 24 March 2008 / Accepted: 22 April 2008 / Published online: 10 June 2008

Abstract The reliability-based design optimization (RBDO) using performance measure approach for problems with correlated input variables requires a transformation from the correlated input random variables into independent standard normal variables. For the transformation with correlated input variables, the two most representative transformations, the Rosenblatt and Nataf transformations, are investigated. The Rosenblatt transformation requires a joint cumulative distribution function (CDF). Thus, the Rosenblatt transformation can be used only if the joint CDF is given or input variables are independent. In the Nataf transformation, the joint CDF is approximated using the Gaussian copula, marginal CDFs, and covariance of the input correlated variables. Using the generated CDF, the correlated input variables are transformed into correlated normal variables and then the correlated normal variables are transformed into independent standard normal variables through a linear transformation. Thus, the Nataf transformation can accurately estimates joint normal and some lognormal CDFs of the input variable that cover broad engineering applications. This paper develops a PMA-based RBDO method for problems with correlated random input variables using the Gaussian copula. Several numerical examples show that the correlated random input variables significantly affect RBDO results.

Y. Noh · K. K. Choi (✉) · L. Du
Department of Mechanical and Industrial Engineering,
College of Engineering, The University of Iowa,
Iowa City, IA 52242, USA
e-mail: kkchoi@engineering.uiowa.edu

Y. Noh
e-mail: noh@engineering.uiowa.edu

L. Du
e-mail: liudu@engineering.uiowa.edu

Keywords Correlated input variables · Reliability-based design optimization (RBDO) · Inverse reliability analysis · Rosenblatt transformation · Nataf transformation · Gaussian copula

Nomenclature

NDV	Number of design variables
n	Number of input random variables
$\mathbf{d}$	Vector of design variables, $\mathbf{d}=\left[d_1, \cdots, d_{NDV}\right]$
$\mathbf{X}$	Vector of random variables, $\mathbf{X}=[X_1, \cdots, X_n]$
$\mathbf{x}$	Realization of vector $\mathbf{X}$, $\mathbf{x}=[x_1, \cdots, x_n]$
$\mathbf{U}$	Vector of independent standard normal variables, $\mathbf{U}=[U_1, \cdots, U_n]$
$\mathbf{u}$	Realization of vector $\mathbf{U}$, $\mathbf{u}=[u_1, \cdots, u_n]$
$\mathbf{Y}$	Vector of correlated standard normal variables, $\mathbf{Y}=[Y_1, \cdots, Y_n]$
$\mathbf{y}$	Realization of vector $\mathbf{Y}$, $\mathbf{y}=\left[y_1, \cdots, y_n\right]$
Ξ_i	Normalized random variable of the ith original variable X_i, $\Xi_i=(X_i-\mu_i)/\sigma_i$
ξ_i	Realization of Ξ_i, $\xi_i=(x_i-\mu_i)/\sigma_i$
$F_{X_i}(\cdot)$	Marginal CDF of X_i
$F_{X_i}(\cdot\mid\cdot)$	Conditional CDF of X_i
$C(\cdot)$	Copula with marginal CDFs
$C_\Phi(\cdot)$	Gaussian copula with marginal CDFs
$M(\cdot)$	Fréchet–Hoeffding upper bound
$W(\cdot)$	Fréchet–Hoeffding lower bound
$\prod(\cdot)$	Independent copula

$\Phi^{-1}(\cdot)$	Inverse normal CDF
$\phi(\cdot)$	Normal probability density function (PDF)
$f_{X_1\cdots X_n}(\cdot)$	Exact joint PDF of $\mathbf{X}$
$\tilde{f}_{X_1\cdots X_n}(\cdot)$	Approximate joint PDF of $\mathbf{X}$
ρ_{ij}	Correlation coefficient between X_i and X_j
ρ'_{ij}	Reduced correlation coefficient between Y_i and Y_j
$\mathbf{P}$	Covariance matrix of $\mathbf{X}$, $\{\rho_{ij}\}$
$\mathbf{P}'$	Reduced covariance matrix of $\mathbf{Y}$, $\{\rho'_{ij}\}$

1 Introduction

In many structural RBDO problems, the input random variables such as the material properties and fatigue properties are correlated (Socie 2003). To solve the RBDO problems with the correlated input variables, it is desirable to have a joint CDF of the input variables to transform the correlated variables into independent standard normal variables for the inverse reliability analysis in RBDO. However, since it has been well-known that the true input joint CDF requires infinite data to obtain, in the literature, most studies have assumed that all input random variables are independent.

To transform correlated variables into independent standard normal variables, there exist various methods: the Hermite polynomial transformation, the Winterstein approximation, the Rosenblatt transformation, and the Nataf transformation (Ditlevsen and Madsen 1996). The Hermite polynomial transformation expresses correlated variables as a linear combination of Hermite polynomials of the standard normal variables using covariance and estimated moments such as mean, variance, skewness, and kurtosis. The Winterstein approximation is a specific type of the Hermite polynomial, which uses a linear combination of three Hermite polynomials. However, the accuracy of these two transformations is directly determined by accurately estimated statistical moments, especially kurtosis and skewness, which are difficult to obtain when the available data are limited. On the other hand, the input marginal CDFs and their parameters, which are required by the Rosenblatt and Nataf transformations, can be more correctly determined using statistical methods based on samples than high moments.

In this paper, two commonly used transformation methods, the Rosenblatt and Nataf transformations, are studied for application to RBDO of problems with correlated input variables. The Rosenblatt transformation (Rosenblatt 1952) requires complete information about the input variables such as a joint CDF (Melchers 1999; Ditlevsen and Madsen 1996). Unlike the Rosenblatt transformation, which uses a given joint CDF, the Nataf transformation approximates the joint CDF using the Nataf model (Nataf 1962), which is identified as a Gaussian copula. In the Nataf transformation, since a copula, which is a link between a joint CDF and marginal CDFs, requires only the marginal CDFs and correlation parameters such as covariance to generate the joint CDF, the joint CDF can be easily generated in real industrial applications. Moreover, since the copula decouples the marginal CDFs and the joint CDF, the joint CDF type can be different from the marginal CDF types. That is, having normal marginal CDFs does not mean the joint CDF is normal (this situation sometimes occurs in real applications). Since the approximate joint CDF is obtained from the Gaussian copula, the linear Rosenblatt transformation (Rosenblatt 1952) can be used. That is, the Nataf transformation converts the correlated variables to the correlated standard normal variables, and then uses the linear transformation (Rosenblatt transformation) to transform the correlated standard normal variables to independent standard normal variables.

The copula has been widely used to obtain the joint CDF in fields such as actuarial science and statistics, but it has not been used in the engineering field for RBDO. This paper is the first to introduce the copula for RBDO. In particular, the Gaussian copula, which is one type of copula, is used in this paper to obtain the joint CDF of input variables for the inverse reliability analysis in RBDO.

However, since the Nataf transformation uses the Gaussian copula, it accurately approximates the joint normal or some lognormal CDFs with small coefficients of variation. It cannot accurately approximate a non-normal CDF. For instance, if the exact joint CDF is lognormal with a large coefficient of variation or exponential, then the Gaussian copula cannot accurately approximate the exact joint CDF.

Even though application of the Gaussian copula is limited, it is still very broadly applicable since the normal CDF and lognormal CDF with a small coefficient of variation cover broad engineering applications. Therefore, in this paper, the Nataf transformation is used to develop an RBDO method for design problems with correlated random input variables. The amount of error that exists between true joint CDF and approximate joint CDF for some CDFs is investigated, as well as the conditions in which the Gaussian copula can or cannot be used to generate some joint non-normal CDF. Numerical examples are used to demonstrate the proposed method, and it is shown that the correlated random input variables significantly affect the RBDO results.

2 Reliability-based design optimization formulation

The RBDO problem can be formulated to

$$\begin{aligned}&\min.\ \ \mathrm{cost}(\mathbf{d})\\&s.t.\ \ P(G_i(\mathbf{X})>0)\le P_{F_i}^{Tar}, i=1,\cdots,NC\\&\qquad \mathbf{d}=\mu(\mathbf{X}), d_L\le\mathbf{d}\le d_U,\quad \mathbf{d}\in R^{NDV} and\, \mathbf{X}\in R^{NRV}\end{aligned} \tag{1}$$

where $\mathbf{X}$ is the vector of random variables; $\mathbf{d}$ is the vector of design variables; $G_i(\mathbf{X})$ represents the constraints; $P_{F_i}^{Tar}$ is a given target probability of failure for the ith constraint; and NC, NDV, and NRV are the number of probabilistic constraints, number of design variables, and number of random variables, respectively.

The probability of failure is estimated by a multiple integral of the joint PDF of the input variables over the failure region as

$$P(G_i(\mathbf{X})>0)=\int_{G_i(\mathbf{x})>0}\cdots\int f_{X_1\cdots X_n}(\mathbf{x})dx_1\cdots dx_n,\quad i=1,\cdots,NC \tag{2}$$

where $\mathbf{x}$ is the realization of the random vector $\mathbf{X}$. However, since it is difficult to compute these multiple integrals analytically, approximation methods such as the first order reliability method (FORM) or the second order reliability method (SORM) are used. Since FORM often provides adequate accuracy and is much easier to use than SORM, it is commonly used in RBDO. Since FORM and SORM require the transformation of the correlated random input variables into the standard normal variables, the Rosenblatt transformation or the Nataf transformation is used.

Using a performance measure approach (PMA+; Youn et al. 2005a, b), the ith probabilistic constraint can be defined from (1)

$$P\big[G_i(\mathbf{X})>0\big]-P_{F_i}^{Tar}\le 0\Rightarrow G_{p_i}\big(\mathbf{x}^*\big)\le 0 \tag{3}$$

where $G_{p_i}(\mathbf{x}^*)$ is the ith probabilistic constraint evaluated at the most probable point (MPP) $\mathbf{x}^*$ in X-space. Using FORM, (3) can be rewritten as

$$P\big[G_i(\mathbf{X})>0\big]-\Phi\big(-\beta_{t_i}\big)\le 0\Rightarrow G_{p_i}\big(\mathbf{x}^*\big)\le 0 \tag{4}$$

where $P_{F_i}^{Tar}=\Phi\left(-\beta_{t_i}\right)$ and β_{t_i} is the target reliability index.

To satisfy the feasibility of the probabilistic constraint, the MPP needs to be estimated for each constraint by solving the following optimization problem:

$$\begin{aligned}&\max.\ G_i(\mathrm{U})\\&\mathrm{s.t.}\quad \|\mathrm{U}\|=\beta_{t_i}\end{aligned} \tag{5}$$

where G_i is the ith constraint, which is transformed from X-space into the standard normal space, U-space. Using the estimated MPPs, the value of the probabilistic constraint can be estimated (Youn et al. 2005a, b).

3 Rosenblatt transformation

The Rosenblatt transformation is a well-known transformation method that maps the correlated variables onto the independent standard normal variables. It is defined by the following successive conditioning:

$$\begin{aligned}u_1&=\Phi^{-1}\left[F_{X_1}(x_1)\right]\\u_2&=\Phi^{-1}\left[F_{X_2}(x_2|x_1)\right]\\&\vdots\\u_n&=\Phi^{-1}\left[F_{X_n}(x_n|x_1,x_2,\cdots,x_{n-1})\right]\end{aligned} \tag{6}$$

where n is number of input variables, $F_{X_i}(x_i|x_1,x_2,\ldots,x_{i-1})$ is the CDF of X_i conditional on $X_1=x_1$, $X_2=x_2,\cdots,X_{i-1}=x_{i-1}$, and $\Phi^{-1}(\cdot)$ is the inverse CDF of the standard normal variables. Based on (6), when the joint CDF is known, the Rosenblatt transformation is exact. For independent input variables, the Rosenblatt transformation also can be used because the joint CDF is the multiplication of the marginal CDFs of each variable. In addition, analytically, the result of the Rosenblatt transformation is not affected by the ordering adopted for the variables $\mathbf{X}$ as shown in the following equation:

$$\begin{aligned}P_f&=\int_{g(\mathbf{X})>0}\ldots\int f_{X_1\ldots X_n}(x_1,\ldots,x_n)dx_1\ldots dx_n\\&=\int_{g(\mathbf{X})>0}\ldots\int f_{X_1}(x_1)f_{X_2}(x_2|x_1)\cdots f_{X_n}\\&\quad\times(x_n|x_1,\ldots x_{n-1})dx_1\ldots dx_n\\&=\int_{g(X)>0}\ldots f_{U_1\cdots U_n}(u_1,u_2,\ldots u_n)\\&\quad\times\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}dx_1\ldots dx_n\\&=\int_{g(U)>0}\ldots\int\phi(u_1)\cdots\phi(u_n)du_1\ldots du_n\end{aligned} \tag{7}$$

where $\phi(u_i)$ is the marginal PDF of the ith independent variable u_i.

Even though the Rosenblatt transformation is exact, since it requires the joint CDF, it can be used only for limited cases where all input variables are independent or a joint CDF is provided. In addition, the result of the Rosenblatt transformation is supposed to be theoretically independent of the ordering, but the estimation of probability of failure might be different for an input joint non-normal CDF due to the approximation error of FORM. This behavior is discussed in Section 5.

4 Nataf transformation

The Nataf transformation uses the Gaussian copula to transform correlated input variables into correlated standard normal variables and linear transformation to transform correlated standard normal variables into independent standard normal variables. Since the Gaussian copula originates from the copula family, it is beneficial to study copulas to understand the Nataf transformation.

4.1 Copula

A true joint CDF, which is necessary for the exact transformation, is usually unknown because it requires infinite number of data that are difficult to obtain in industrial applications. However, a copula only requires marginal CDFs and correlation parameters to obtain an approximate joint CDF, so that the joint CDF can be easily obtained from limited data.

The *copula* originates from a Latin word for "link" or "tie" that connects two different things. In statistics, the definition of copulas is stated by Roser (1999): "Copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval [0, 1]."

According to Sklar's theorem, if the random variables have a joint distribution $F_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ with marginal distributions, $F_{X_1}(x_1), \cdots,$ and $F_{X_n}(x_n)$, then there exists an n-dimensional copula C such that

$$F_{X_1 \ldots X_n}(x_1, \ldots, x_n) = C\big(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\big) \tag{8}$$

If marginal distributions are all continuous, then C is unique. Conversely, if C is an n-dimensional copula and $F_{X_1}(x_1), \cdots,$ and $F_{X_n}(x_n)$ are the marginal distributions, then $F_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ is the joint distribution (Roser 1999).

Since the joint CDF is expressed as a function of marginal CDFs, it is easy to obtain a joint CDF from marginal CDFs and correlation parameters. Moreover, since the copula decouples marginal CDFs and the joint CDF, the joint CDF generated from the copula can be expressed as any type of marginal CDFs. Thus, it is desirable to use the copula for constructing the joint CDF in real applications that may have correlated input variables with a joint CDF but with different types of marginal CDFs of the input variables.

Let $F_{X_i}(x_i) = u_i$ for $i = 1, \cdots, n$. Any copula $C(u_1, \cdots, u_n)$ lies between the Fréchet–Hoeffding lower and upper bounds for every $(u_1, \cdots, u_n)$ in I^n, and the bounds are themselves copulas, which are given as

$$\max(u_1 + \cdots + u_n - n + 1, 0) \le C(u_1, \cdots, u_n) \le \min(u_1, \cdots, u_n) \tag{9}$$

where $I^n = I \times I \times \cdots \times I \; (I = [0, 1])$.

For the two-dimensional case, (9) can be written as

$$\max(u_1 + u_2 - 1, 0) \le C(u_1, u_2) \le \min(u_1, u_2) \tag{10}$$

Let $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$, $M(u_1, u_2) = \min(u_1, u_2)$, and consider an independent copula $\prod(u_1, u_2) = u_1 \cdot u_2$. These copulas $W(u_1, u_2)$, $M(u_1, u_2)$, and $\prod(u_1, u_2)$ are graphically shown in Fig. 1 in u_1–u_2 space.

Three copulas can be easily compared by drawing these copulas along the diagonal direction $u_1 = u_2$ as shown in Fig. 2. The graph of any two-dimensional copula is a continuous surface within I^3 (Fig. 1), and along the horizontal, vertical, and diagonal directions, all copulas are non-decreasing functions and uniformly continuous on I (Roser 1999).

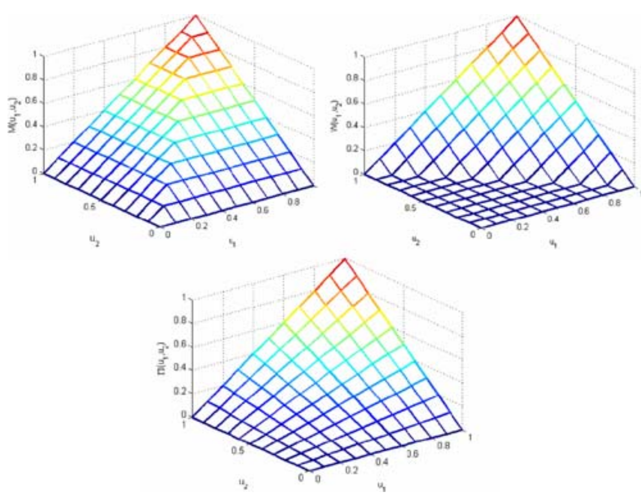

Fig. 1 Graph of the copulas W, M, and $\prod$

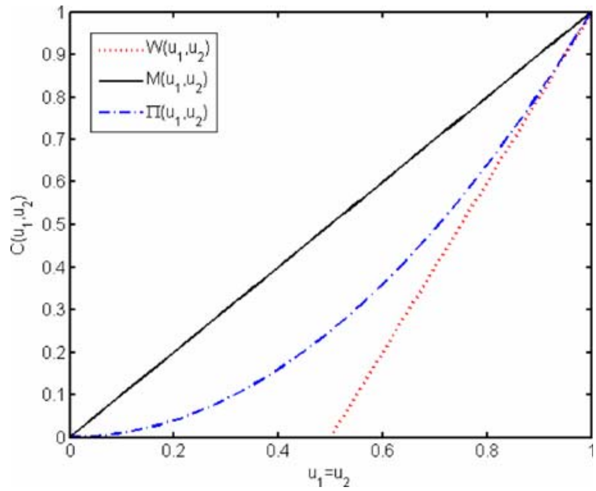

Fig. 2 Graph of copulas W, M, and $\prod$ along the diagonal direction

4.2 Gaussian copula

The Nataf transformation transfers correlated input variables $\mathbf{X}$ with marginal CDF $F_{X_i}(x_i)$ and covariance matrix $\mathbf{P} = \{\rho_{ij}\}$ into independent standard normal variables $\mathbf{U}$ (the covariance matrix of $\mathbf{U}$ is an identity matrix $\mathbf{I}$) through the multivariate correlated standard normal variables $\mathbf{Y}$ with reduced covariance matrix $\mathbf{P}' = \{\rho'_{ij}\}$, which is the covariance matrix of $\mathbf{Y}$, using the Gaussian copula.

The Gaussian copula is a link between a multivariate normal distribution and marginal distributions. The n-dimensional Gaussian copula with reduced covariance matrix $\mathbf{P}'$ is defined as

$$C_\Phi(\mathbf{u}) = \Phi_{\mathbf{P}'}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)), \quad \mathbf{u} \in I^n \tag{11}$$

where u_i can be any arbitrary marginal CDF $F_{X_i}(x_i)$. The Gaussian copula in (11) might be confused with the joint normal CDF, which has been widely known, in (12).

$$\Phi_{\mathbf{P}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} \frac{\exp\left[-(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{P}^{-1} (\mathbf{x}-\boldsymbol{\mu})/2\right]}{(2\pi)^{n/2} |\mathbf{P}|^{1/2}} dx_1 \cdots dx_n \tag{12}$$

The difference between the Gaussian copula and the joint normal CDF is that the Gaussian copula allows having different marginal CDF types from the joint CDF type whereas the joint normal CDF does not. Thus, the Gaussian copula is indeed distinguished from the joint normal CDF.

The first step of the Nataf transformation is to transfer the margin of $\mathbf{X}$ into the standard normal margin $\mathbf{Y}$ using

$$y_i = \Phi^{-1}\left[F_{X_i}(x_i)\right], \; i = 1, \cdots, n. \tag{13}$$

The covariance matrix of $\mathbf{Y}$, which is called as the reduced covariance matrix $\mathbf{P}'$ is unknown. The second step of the Nataf transformation is to estimate the reduced covariance matrix. If the multivariate standard normal variable $\mathbf{Y}$ has the joint PDF (probability density function) $\phi(\mathbf{y}, \mathbf{P}')$, the covariance matrix $\mathbf{P} = \{\rho_{ij}\}$ of the correlated input variables $\mathbf{X}$ should be defined as

$$\rho_{ij} = E[\Xi_i \Xi_j] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi_i \xi_j \phi(y_i, y_j; \rho'_{ij}) dy_i dy_j \tag{14}$$

where $\Xi_i = \left(X_i - \mu_{X_i}\right)/\sigma_{X_i}$ is the normalized random variable of X_i, ξ_i is the realization of Ξ_i, and ρ_{ij} is the correlation coefficient between X_i and X_j (Melchers 1999; Ditlevsen and Madsen 1996). However, since the iterative process is very tedious and unknowns are within the double integral, (14) is approximated by

$$\rho'_{ij} = R_{ij}\rho_{ij} \tag{15}$$

to obtain the reduced correlation coefficient. In (14), R_{ij} is approximated by

$$R_{ij} = a + bV_i + cV_i^2 + d\rho_{ij} + e\rho_{ij}^2 + f\rho_{ij}V_i + gV_j + hV_j^2 + k\rho_{ij}V_j + lV_iV_j \tag{16}$$

where V_i and V_j are the coefficients of variation $(V = \sigma/\mu)$ for each variable, and the coefficients depend on the types of input variables. For different types of input variables, the corresponding coefficients are given by Liu and Der Kiureghian (1986), Melchers (1999), and Ditlevsen and Madsen (1996). The maximum error of the estimated correlation coefficient obtained from (15) is normally much less than 1%, and even if the exponential CDF or negative correlation is involved, the maximum error in the correlation coefficient is at most up to 2% (Melchers 1999). Therefore, the approximation provides adequate accuracy with less computational effort.

As stated in Section 2, the reliability analysis is carried out using the transformed standard independent normal variables $\mathbf{U}$. Since the relationship between the correlated input variables $\mathbf{X}$ and the correlated standard normal variables $\mathbf{Y}$ is given in (13), the next step is to transform the correlated standard normal variables $\mathbf{Y}$ to the independent standard normal variables $\mathbf{U}$ using a linear transformation.

Consider the following linear equation:

$$\mathbf{Y} = \mathbf{AU} + \mathbf{B} \tag{17}$$

where $\mathbf{Y}\sim N(\mathbf{0},\mathbf{I})$ has the reduced correlation matrix $\mathbf{\Sigma_Y} = \mathbf{P}'$ and $\mathbf{U}\sim N(\mathbf{0},\mathbf{I})$ has the covariance matrix $\mathbf{\Sigma_U} = \mathbf{I}$. The mean of $\mathbf{Y}$ can be calculated as

$$\mathbf{E}[\mathbf{Y}] = \mathbf{E}\left[\mathbf{AU}+\mathbf{B}\right] = \mathbf{AE}[\mathbf{U}] + \mathbf{B} = \mathbf{B} = \mathbf{0} \tag{18}$$

In the same way, the covariance matrix of $\mathbf{Y}$ can be calculated as

$$\mathbf{P}' = \mathbf{\Sigma_Y} = Var\left[\mathbf{AU}+\mathbf{B}\right] = Var\left[\mathbf{AU}\right] = \mathbf{A\Sigma_U A^T} = \mathbf{AA^T} \tag{19}$$

Since the covariance matrix of $\mathbf{Y}$ is positive definite, $\mathbf{P}'$ can be decomposed into the lower and upper triangular matrix using Cholesky factorization. Therefore, the matrix $\mathbf{A}$ can be expressed as a lower triangular matrix. If the joint input CDF is normal, the Rosenblatt transformation becomes a linear transformation (Rosenblatt 1952). In the Nataf transformation, since the correlated normal variables $\mathbf{Y}$ has a joint normal CDF, the Nataf transformation can be viewed as a combination of the Gaussian copula and the linear Rosenblatt transformation. Instead of using the Rosenblatt transformation, eigenvalues and eigenvectors of the reduced covariance matrix can be used to transform the correlated normal variables into the independent normal variables. In that case, the transformation matrix $\mathbf{A}$ consists of the eigenvectors of the reduced covariance matrix and this linear transformation is called as an orthogonal transformation (Madsen et al. 1986). However, since the two linear transformations provide the same RBDO results, only the Rosenblatt transformation was used in this paper.

Using (13) and (17), the relationship (Nataf transformation) between the correlated input variables $\mathbf{X}$ and the independent standard variables $\mathbf{U}$ is obtained as

$$\begin{aligned} x_1 &= F_{X_1}^{-1}(\Phi(a_{11}u_1)) \\ x_2 &= F_{X_2}^{-1}(\Phi(a_{12}u_1 + a_{22}u_2)) \\ &\vdots \\ x_n &= F_{X_n}^{-1}(\Phi(a_{1n}u_1 + a_{2n}u_2 + \cdots + a_{nn}u_n)) \end{aligned} \tag{20}$$

where the entries a_{ij} in the lower triangular matrix $\mathbf{A}$ is expressed in terms of the reduced correlation coefficients as shown in (21).

$$a_{ij} = \begin{cases} \sqrt{\left(1 - \sum_{k=1}^{i-1} a_{ik}^2\right)}, & i = j \\ \left(\rho'_{ij} - \sum_{k=1}^{j-1} a_{ik}a_{jk}\right) / a_{jj}, & i > j \end{cases} \tag{21}$$

Since the correlated variables can be expressed in terms of the independent standard normal variables, the reliability analysis can be carried out by substituting (20) into the constraint function G_i in (1). As previously stated, if only the covariance matrix and marginal CDF are available, Gaussian copula needs to be used to construct the joint CDF of the input random variables.

4.3 Applicability of Gaussian copula

An advantage of Nataf transformation is that it involves a Gaussian copula that can generate a joint CDF for various types of correlated input variables based on limited information. As a result, many industrial problems with various types of the correlated input variables can be solved using RBDO. However, it is noted that Gaussian copula can approximate joint normal or some lognormal CDFs more accurately than other joint non-normal CDFs.

For the first case, consider mixed normal $X_1 : N\left(\mu_1, \sigma_1^2\right)$ and lognormal $X_2 : LN\left(\lambda_2, \xi_2\right)$ input random variables. The exact joint PDF of normal and lognormal variables is given as

$$\begin{aligned} f_{X_1X_2}(x_1, x_2) &= \frac{1}{2\pi\sigma_1\xi_2\sqrt{1-\eta_2^2}x_2} \\ &\times \exp\left\{-\frac{1}{2}\left[\frac{\left(\frac{\ln x_2 - \lambda_2}{\xi_2}\right) - \eta_2\left(\frac{x_1-\mu_1}{\sigma_1}\right)}{\sqrt{1-\eta_2^2}}\right]^2 \right. \\ &\left. -\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right\} \end{aligned} \tag{22}$$

where $\xi_2 = \sqrt{\ln\left(1+\left(\frac{\sigma_2}{\mu_2}\right)^2\right)} = \sqrt{\ln\left(1+\kappa_2^2\right)}$, $\lambda_2 = \ln\mu_2 - \frac{1}{2}\xi_2^2$, and $\eta_2 = \frac{\rho_{12}\kappa_2}{\xi_2}$. The approximate joint PDF can be obtained by differentiating the Gaussian copula given in (11):

$$\begin{aligned} \tilde{f}_{x_1\cdots x_n}(x_1, \ldots, x_n) &= \frac{\partial^2 C_\Phi}{\partial x_1 \partial x_2} = \frac{\partial^2 C_\Phi}{\partial y_1 \partial y_2} \cdot \frac{\partial y_1}{\partial x_1} \cdot \frac{\partial y_2}{\partial x_2} \\ &= \phi\left(y_1, y_2, \rho'_{12}\right) \frac{f_{x_1}(x_1) f_{x_2}(x_2)}{\phi(y_1)\phi(y_2)} \end{aligned} \tag{23}$$

where $\frac{\partial^2 C_\Phi}{\partial y_1 \partial y_2} = \phi\left(y_1, y_2, \rho'_{12}\right)$, $\frac{\partial y_i}{\partial x_i} = \frac{f(x_i)}{\phi(y_i)}$, $f_{X_i}(x_i)$ is the marginal PDF of X_i, $\phi(y_i)$ is the normal PDF of Y_i for i = 1, 2, and $\phi\left(y_1, y_2, \rho'_{12}\right)$ is the joint normal PDF of Y_1 and Y_2. The reduced correlation coefficient ρ'_{12} between the correlated standard normal variables Y_1 and Y_2 is obtained from the correlation coefficient ρ_{12} between X_1 and X_2 (Melchers 1999; Ditlevsen and Madsen 1996)

$$\rho'_{12} = \frac{\kappa_2}{\sqrt{\ln\left(1+\kappa_2^2\right)}}\rho_{12} = \eta_2 \tag{24}$$

Using (23) and (24), the approximate joint PDF of the normal and lognormal variables is

$$\begin{aligned}\tilde{f}_{X_1X_2}(x_1,x_2) &= \frac{f_{X_1}(x_1)\cdot f_{X_2}(x_2)}{\phi(y_1)\cdot\phi(y_2)}\phi\left(y_1,y_2,\rho_{12}^{'}\right)\\ &= \frac{1}{2\pi\sigma_1\xi_2\sqrt{1-\rho_{12}^{'2}}x_2}\exp\left\{-\frac{y_1^2-2\rho_{12}^{'}y_1y_2+y_2^2}{2\left(1-\rho_{12}^{'2}\right)}\right\}\\ &= \frac{1}{2\pi\sigma_1\xi_2\sqrt{1-\rho_{12}^{'2}}x_2}\exp\left\{-\frac{\left(y_2-\rho_{12}^{'}y_1\right)^2+\left(1-\rho_{12}^{'2}\right)y_1^2}{2\left(1-\rho_{12}^{'2}\right)}\right\}\\ &= \frac{1}{2\pi\sigma_1\xi_2\sqrt{1-\eta_2^2}x_2}\exp\left\{-\frac{1}{2}\left[\frac{y_2-\eta_2y_1}{\sqrt{1-\eta_2^2}}\right]^2-\frac{1}{2}y_1^2\right\}\end{aligned} \tag{25}$$

where $y_1=\frac{x_1-\mu_{x_1}}{\sigma_{x_1}}$, $y_2=\frac{\ln x_2-\lambda_2}{\xi_2}$. In (25), $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$ are given as

$$f_{X_1}(x_1)=\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right\}$$
$$f_{X_2}(x_2)=\frac{1}{\xi_2x_2\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\ln x_2-\lambda_2}{\xi_2}\right)^2\right\}. \tag{26}$$

Note that (25) is the same as (22). Thus, the Gaussian copula can construct the exact joint PDF of normal and lognormal variables in the two-dimensional case.

As a second case, consider two lognormal variables, $X_1 \sim LN(\lambda_1, \xi_1)$ and $X_2 \sim LN(\lambda_2, \xi_2)$. The exact joint lognormal PDF is given as

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{2\pi\xi_1\xi_2\sqrt{1-\rho_{12}^2}x_1x_2}\times\exp\left\{-\frac{y_1^2-2\rho_{12}y_1y_2+y_2^2}{2\left(1-\rho_{12}^2\right)}\right\} \tag{27}$$

where $y_i=\frac{\ln x_i-\lambda_i}{\xi_i}$, $\xi_i=\sqrt{\ln\left(1+\kappa_i^2\right)}$ and $\lambda_i=\ln\mu_i-\frac{1}{2}\xi_i^2$ for i = 1,2. Differentiating the Gaussian copula given in (11), the approximate joint PDF of two lognormal variables can be obtained as

$$\begin{aligned}\tilde{f}_{X_1X_2}(x_1,x_2) &= \frac{f_{X_1}(x_1)\cdot f_{X_2}(x_2)}{\phi(y_1)\cdot\phi(y_2)}\phi\left(y_1,y_2,\rho_{12}^{'}\right)\\ &= \frac{1}{2\pi\xi_1\xi_2\sqrt{1-\rho_{12}^{'2}}x_1x_2}.\\ &\quad\times\exp\left\{-\frac{y_1^2-2\rho_{12}^{'}y_1y_2+y_2^2}{2\left(1-\rho_{12}^{'2}\right)}\right\}\end{aligned} \tag{28}$$

The reduced correlation coefficient between the lognormal variables is obtained as

$$\rho_{12}^{'}=\frac{\ln\left(1+\rho_{12}\kappa_1\kappa_2\right)}{\sqrt{\ln\left(1+\kappa_1^2\right)\cdot\ln\left(1+\kappa_2^2\right)}} \tag{29}$$

Since (27) and (28) have the same formulation except for the correlation coefficients, the Gaussian copula can accurately approximate a joint CDF of the lognormal variables if the difference between the reduced correlation coefficient and the original correlation coefficient is small.

Assume $\mu_1=\mu_2=1.0$ and $\sigma_1=\sigma_2=0.3$. Figure 3 shows the original correlation coefficient and the reduced correlation coefficient obtained from (29). If two lognormal input variables are positively correlated, the difference between the reduced correlation coefficient and the original correlation coefficient is small. However, for negative correlation coefficients, the original correlation coefficient and the reduced correlation coefficient is rather different. Thus, if two variables are positively correlated or independent, the joint CDF can be accurately estimated using the Gaussian copula, but for the negatively correlated input variables, the joint CDF may not be accurate.

To investigate how the relative error between the reduced correlation coefficient and the original correlation coefficient affects the accuracy of the estimated joint CDF, Fig. 4 shows the relative error between the exact joint CDF and the approximate joint CDF obtained from the Gaussian copula for different

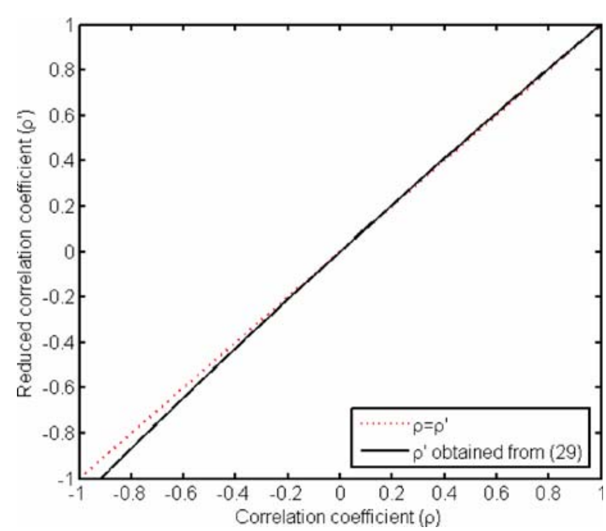

Fig. 3 Relative error between original correlation coefficient and reduced correlation coefficient

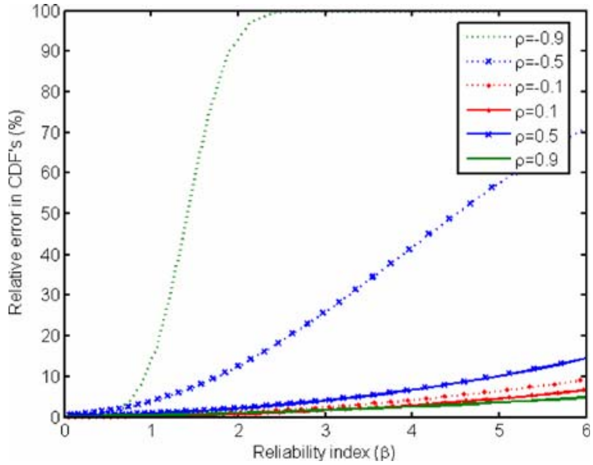

Fig. 4 Relative error between exact joint CDF and approximate joint CDF obtained from Nataf transformation

correlation coefficients along the diagonal direction ($x_1 = x_2$). The relative error is calculated by

$$\textit{Relative error}\ (\%) = \frac{\left|\tilde{F}_{X_1X_2}(x_1, x_2) - F_{X_1X_2}(x_1, x_2)\right|}{F_{X_1X_2}(x_1, x_2)} \times 100 \tag{30}$$

where $\tilde{F}_{X_1X_2}(x_1, x_2)$ and $F_{X_1X_2}(x_1, x_2)$ are the approximate and exact joint CDFs of the lognormal variables, respectively. As shown in Fig. 4, the more negatively the lognormal variables are correlated, the more significant are the relative errors contained in the approximate CDF.

If the correlation coefficients are positive, even if the lognormal input variables are highly correlated, the relative error in CDF is less than 15%, as shown in Fig. 4. The maximum error (15%) occurs near $\rho = 0.5$. The reason is that, for a positive high correlation coefficient, even though the difference between the reduced correlation coefficient and the original correlation coefficient is small, the correlation significantly affects estimation of the joint CDF. On the other hand, for positive low correlation coefficient values, even if the values of the original correlation coefficient and the reduced correlation coefficient are rather different, the effect of the correlation is negligible in estimating the joint CDF. Thus, the mid-range correlation ($\rho = 0.5$) causes a maximum error when the joint CDF is approximated.

Therefore, the Gaussian copula is very effective for building a joint CDF for the problem with positive correlated and independent lognormal input variables, but not applicable to the ones with negative correlated lognormal input variables. From the above result, it might be possible to change the sign of input variables to convert the negative correlation into the positive correlation to make the Gaussian copula accurately approximate lognormal CDF with negative correlation. However, it may not be a simple process.

In this example, the coefficient of variation, which is the ratio of the standard deviation to the mean, is moderately small (i.e., $\sigma_1/\mu_1 = \sigma_2/\mu_2 = 0.3$), and thus the difference between the exact and approximate CDF values is small. As a result, the shape of the lognormal distribution is very similar to the normal distribution. On the other hand, if the coefficient of variation is large, the lognormal distribution is rather similar to the exponential distribution. Thus, the Gaussian copula may accurately estimate the lognormal distribution with a small coefficient of variation but not accurately estimate non-normal distributions that are rather different from the normal distribution.

In fact, if the input variables are all normal, i.e., marginal CDFs and joint CDF are normal, the linear transformation, such as the orthogonal transformation, can be directly used. However, if the marginal CDFs or joint CDF are not normal, then the orthogonal transformation is not applicable since it might be significantly erroneous. To deal with various types of input variables, Youn et al. (2007) categorized the input types as four cases and used a linear transformation involving the eigenvalues and eigenvectors of the covariance of the input variables. However, since the linear transformation is only applicable when the joint CDF is normal, it is accurate for only normal and some lognormal CDFs, as explained in Fig. 4, but cannot accurately estimate the joint non-normal CDF.

This paper shows the applicable range of the Gaussian copula that is most commonly used in real applications and leads to a new investigation of other types of copulas that can be used to generate the joint non-normal CDF. The generated joint CDF using a copula can be used in the Rosenblatt transformation. The next research topic will address this issue.

5 Accuracy of reliability analyses using two transformations

In this section, a mathematical example is used to demonstrate how the Rosenblatt and Nataf transformations yield the reliability analysis results when the input variables are correlated with a joint exponential CDF. This example was introduced by Hohenbichler and Rackwitz (1981), and was discussed by Madsen et al. (1986).

Consider the following joint PDF of the exponential input variables:

$$f_{X_1X_2}(x_1,x_2)=\begin{cases}(x_1+x_2+x_1x_2)\exp(-x_1-x_2-x_1x_2), \\ \qquad x_1,\ x_2\geq 0 \\ 0, \quad otherwise\end{cases} \tag{31}$$

where the correlation coefficient is $\rho=-0.40366$. The constraint function is given as

$$G(\mathbf{x})=18-3x_1-2x_2 \tag{32}$$

Since the joint PDF can be written in two different ways as shown in (33),

$$f_{X_1X_2}(x_1,x_2)=f_{X_1}(x_1)f_{X_2}(x_2|x_1)=f_{X_2}(x_2)f_{X_1}(x_1|x_2) \tag{33}$$

the independent standard normal variables also can be expressed as correlated input variables in two different ways using the Rosenblatt transformation.

$$\begin{aligned}u_1&=\Phi^{-1}\left[F_{X_1}(x_1)\right]=\Phi^{-1}\left[1-\exp(-x_1)\right]\\ u_2&=\Phi^{-1}\left[F_{X_2}(x_2|x_1)\right]\\ &=\Phi^{-1}\left[1-(1+x_2)\exp\left[-(x_2+x_1x_2)\right]\right]\end{aligned} \tag{34}$$

and

$$\begin{aligned}u_1&=\Phi^{-1}\left[F_{X_2}(x_2)\right]=\Phi^{-1}\left[1-\exp(-x_2)\right]\\ u_2&=\Phi^{-1}\left[F_{X_1}(x_1|x_2)\right]\\ &=\Phi^{-1}\left[1-(1+x_1)\exp\left[-(x_1+x_1x_2)\right]\right]\end{aligned} \tag{35}$$

Similarly, using the Nataf transformation, the independent standard normal variables can be obtained in terms of the correlated input variables in two different ways.

$$\begin{cases}u_1=\Phi^{-1}\left[1-e^{-x_1}\right]\\ u_2=\dfrac{\Phi^{-1}\left[1-e^{-x_2}\right]-\rho'\Phi^{-1}\left[1-e^{-x_1}\right]}{\sqrt{1-\rho'^2}}\end{cases} \tag{36}$$

and

$$\begin{cases}u_1=\Phi^{-1}\left[1-e^{-x_2}\right]\\ u_2=\dfrac{\Phi^{-1}\left[1-e^{-x_1}\right]-\rho'\Phi^{-1}\left[1-e^{-x_2}\right]}{\sqrt{1-\rho'^2}}\end{cases} \tag{37}$$

where $\rho'\approx R\rho=-0.556$ with $R=1.229-0.376\rho+0.153\rho^2$.

As can be seen in (34) through (37), the Rosenblatt transformation uses the complete information such as the joint CDF of the input variables, while the Nataf transformation uses only marginal CDFs and covariance of input variables for transformation from $\mathbf{X}$ to $\mathbf{U}$.

When the Nataf and Rosenblatt transformations are used, the linear constraint function in (32) becomes highly nonlinear as shown in Fig. 5. In Fig. 5, the solid line (ordering 1) and the dashed line (ordering 2) indicate the constraint functions obtained from the Rosenblatt transformation for the ordering 1 and 2 of the input variables. The dotted line (ordering 1) and the dash-dot line (ordering 2) indicate the constraint functions obtained from the Nataf transformation for two different orderings of the input variables.

When the Rosenblatt and Nataf transformations are used in reliability analysis, almost the same MPP points are obtained in U space for the ordering 1, but different MPP points are obtained for the ordering 2 due to the nonlinearity of the constraint function. Accordingly, results such as MPP points and the probability of failure become different for different transformations and different orderings of the input variables as shown in Table 1.

As stated in Section 3, the Rosenblatt transformation should be theoretically independent of orderings of the input variables and provide accurate reliability analysis results when the exact joint input CDF is given. However, as shown in Table 1, the probability of failure result from the ordering 1 is accurate, but the result from the ordering 2 is inaccurate because the joint non-

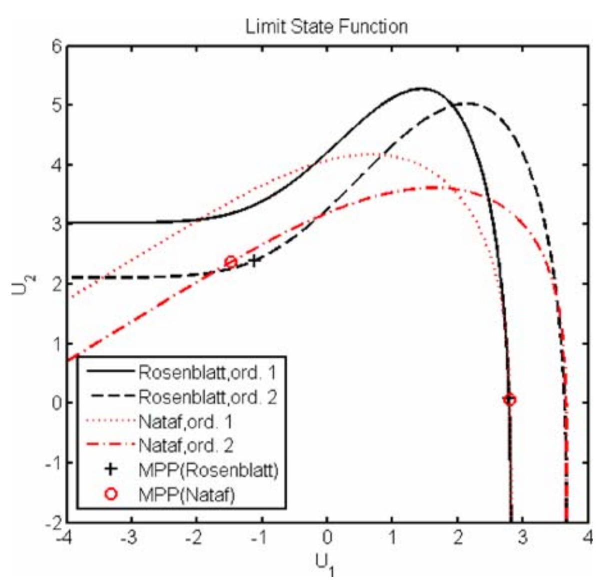

Fig. 5 Constraint functions for different transformations and orderings

Table 1 Reliability analysis results for different transformations and orderings

	Rosenblatt transformation		Nataf transformation		Independent
	Ord. 1	Ord. 2	Ord. 1	Ord. 2	
X_1^{MPP*}	5.911	5.907	5.954	5.954	4.969
X_2^{MPP*}	0.125	0.140	0.069	0.070	1.545
U_1^{MPP*}	2.781	−1.123	2.795	−1.493	2.460
U_2^{MPP*}	0.064	2.399	0.064	2.364	0.795
β	2.782	2.649	2.796	2.796	2.585
P_f	2.703E - 3	4.037E - 3	2.587E - 3	2.587E - 3	4.865E - 3
Error (%)	8.05	37.31	12.01	12.01	65.48
MCS	2.940E-3				

X^{MPP*} MPP in X space, U^{MPP*} MPP in U space, *MCS* Monte Carlo simulation

normal CDF yields a high nonlinearity of the constraint function near the MPP, which cannot be accurately estimated by FORM. On the other hand, in the Nataf transformation, since the linear transformation does not change nonlinearity of the constraint functions near the MPPs, the probability of failure results are very similar for different orderings. Table 1 also provides the results obtained if we assume that two input variables are independent when in fact these two input variables are correlated. As shown in the last column of the table, the assumption that two correlated variables are independent could lead to wrong results such that the MPP points, reliability index, and probability of failure have significant errors compared with the Monte Carlo simulation result. On the other hand, if the correlation in the input variables is considered in the reliability analysis, the errors are reduced. Thus, it is very important to consider the correlation in carrying out the reliability analysis and RBDO.

Even though the Nataf transformation results do not depend on the orderings of the input variables, unlike those of the Rosenblatt transformation, there still exists the approximation error between the generated joint CDF and the true joint CDF, as well as the FORM error in estimating the probability of failure. Thus, the Nataf transformation should be used carefully for approximating the joint non-normal CDF. When the exact joint non-normal CDF is given, the Rosenblatt transformation should be used. However, the nonlinearity of the transformation causes difficulty in accurate estimation of the probability of failure, especially when FORM is used. To resolve the inaccuracy of the FORM, a higher order approximation methods such as the SORM needs to be used. A recently developed MPP-based dimension reduction method (DRM; Lee et al. 2008), which reduces the FORM error, can offer a method that reduces the ordering effect. This MPP-based DRM is currently being investigated.

6 Limitation of Gaussian copula

It is noted that the Gaussian copula is applicable to approximate the joint normal CDF and joint lognormal CDF with a small coefficient of variation. It is informative to study whether it is also applicable to other joint non-normal CDFs whose CDF shapes are rather different from the joint normal CDF, since design variables in real industrial problems could include other types of input distributions.

For instance, for two exponential input variables with a joint exponential CDF, which could be used to analyze the reliability of an electronic system, the original joint exponential PDF is given as (Kotz et al. 2000)

$$f_{X_1X_2}(x_1,x_2) = \begin{cases} \{(1+\theta x_1)(1+\theta x_2)-\theta\} \\ \times\exp\left[-(x_1+x_2+\theta x_1\cdot x_2)\right], \\ \quad x_1\geq 0\,,\, x_2\geq 0 \\ 0, \quad otherwise \end{cases} \tag{38}$$

The corresponding exponential joint CDF is

$$F_{X_1X_2}(x_1,x_2) = \begin{cases} 1-e^{-x_1}-e^{-x_2}+\exp\left[-(x_1+x_2+\theta x_1x_2)\right], \\ \quad x_1\geq 0,\; x_2\geq 0 \\ 0, \quad otherwise \end{cases} \tag{39}$$

where the mean and standard deviation are $\mu_1=\mu_2=1.0$ and $\sigma_1=\sigma_2=1.0$, respectively. The correlation coefficient for the exponential variables is

$$\rho_{ij} = -1+\frac{1}{\theta}e^{1/\theta}Ei\left(\frac{1}{\theta}\right) \tag{40}$$

where $Ei(z)=\int_1^\infty \left(e^{-tz}/t\right)dt$. If the parameter θ varies from 0.0 to 1.0, the correlation coefficient ρ ranges from 0.0 to −0.40366. In this example, it is assumed that $\theta=1.0$ and, thus, $\rho=-0.40366$.

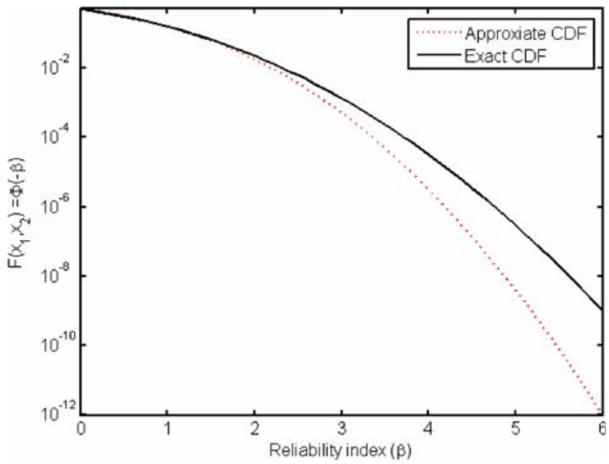

Fig. 6 Exact and approximate exponential CDFs

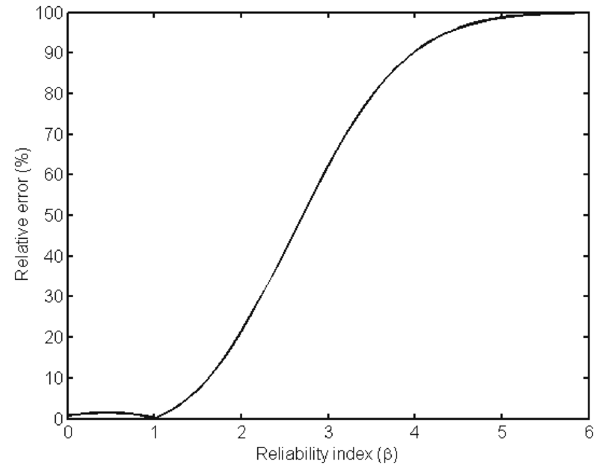

Fig. 7 Relative error of joint exponential CDF obtained from the Gaussian copula at different reliability index levels

Differentiating the joint CDF obtained from the Gaussian copula given in (11), an approximate joint PDF can be obtained as

$$\begin{aligned}\tilde{f}_{X_1X_2}(x_1,x_2) &= \frac{f_{X_1}(x_1)\cdot f_{X_2}(x_2)}{\phi(y_1)\cdot\phi(y_2)}\phi\left(y_1,y_2,\rho'\right)\\ &= \frac{e^{-x_1}\cdot e^{-x_2}}{\phi(y_1)\cdot\phi(y_2)}\phi\left(y_1,y_2,\rho'\right)\end{aligned} \tag{41}$$

where $Ei(z)=\int_1^\infty\left(e^{-tz}/t\right)dt$, $y_2=\Phi^{-1}\left[1-e^{-x_2}\right]$, and $\rho'\approx R\rho=-0.556$. The approximate joint exponential CDF can be estimated by integrating the approximate joint PDF in (41) as

$$\tilde{F}_{X_1X_2}(x_1,x_2)=\int_0^{x_2}\int_0^{x_1}\tilde{f}(x_1,x_2)\,dx_1dx_2 \tag{42}$$

The approximate joint exponential CDF obtained from the Gaussian copula is compared with the exact joint exponential CDF at different reliability index levels, where the reliability index is obtained from

$$\beta=-\Phi^{-1}(F(x_1,x_2)) \tag{43}$$

by comparing these CDFs along the line $x_1=x_2$, as shown in Fig. 6.

From Fig. 6, where the vertical axis is drawn in log scale, the difference of the exact and approximate CDF values is small, but the relative error is significant in the interval from $\beta=2.0$ to 6.0 as shown in Fig. 7. As the reliability index increases, the relative error of the approximate joint CDF increases rapidly.

Further, at a certain target reliability index, e.g., $\beta=3.0$, the relative error is significant for most values of the correlation coefficient (Fig. 8). Thus, when the joint CDF is exponential, the Gaussian copula can be used for the small target reliability index (i.e., less than 2.0), but it may not be appropriate for the large target reliability index due to the large relative error in the CDF.

Since the exponential variables have a large coefficient of variation, which is 1, the approximated joint CDF using the Gaussian copula contains significant error compared to the true exponential CDF. On the other hand, the lognormal variables have a wide range of coefficient of variations. Thus, the Gaussian copula can accurately approximate the joint lognormal CDF with a small coefficient of variation.

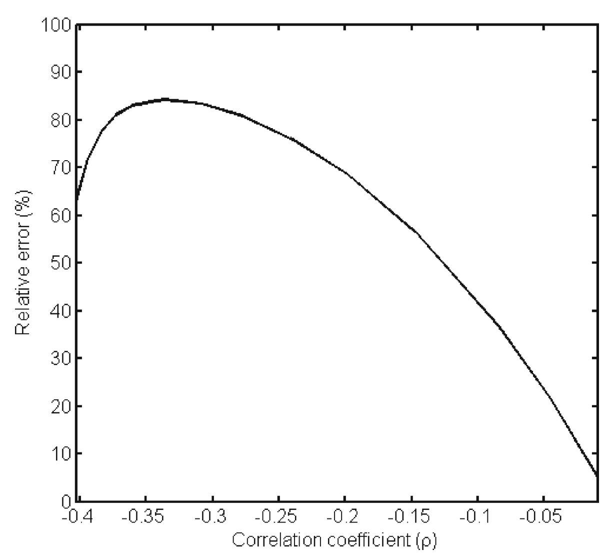

Fig. 8 Relative error versus correlation coefficient for $\beta=3.0$

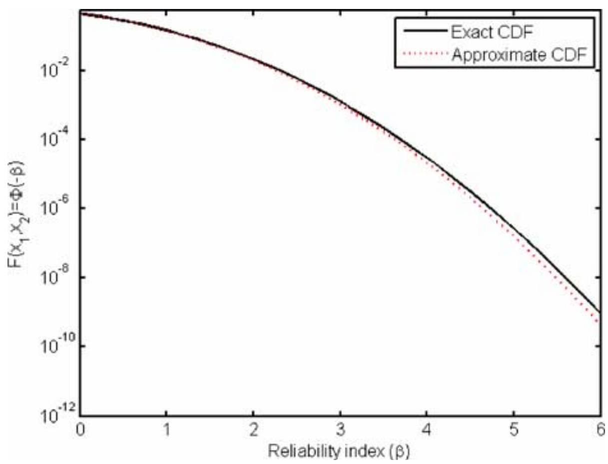

Fig. 9 Exact and approximate lognormal CDFs

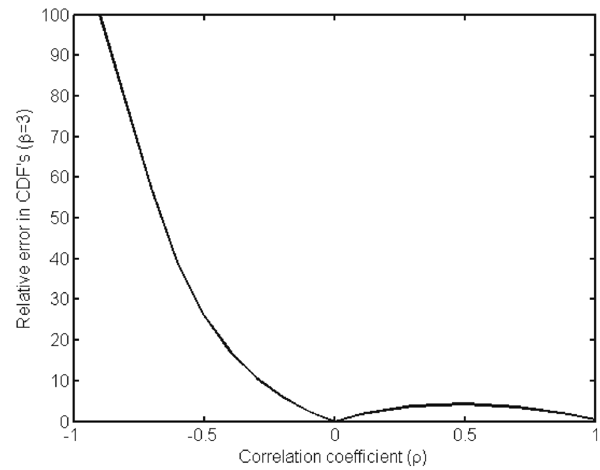

Fig. 11 Relative error in CDF versus correlation coefficients

Similarly, consider the joint lognormal CDF in (27) in Section 4.3. Figure 9 shows a comparison of the exact joint CDF and the approximate joint lognormal CDF at different reliability index levels. Assume that the correlation coefficient is -0.40366 and the coefficient of variation is 0.3. As shown in Fig. 9, the difference between the approximate joint CDF and exact joint CDF is small for up to $\beta = 3$. The trend can be more clearly observed by the relative error shown in Fig. 10.

The relative error of the approximate joint lognormal CDF also can be calculated for the whole range of the correlation coefficient at certain target reliability index, e.g., $\beta = 3$. Figure 11 shows that the relative error is significant for low values of negative correlation coefficients, but that it is small for correlation coefficients between -0.3 and 1. Thus, the Gaussian copula is applicable for the problem with positively correlated lognormal variables and some values of negative correlation coefficients with small target reliability index if the coefficient of variation is small.

Instead of using the Gaussian copula, it might be possible to use an exponential copula to generate a joint exponential CDF; like using the Gaussian copula to generate a joint normal CDF. In fact, significant research has been carried out to develop exponential copulas. However, the currently developed exponential copulas do not seem to be applicable to RBDO because they are not continuous for multi-variables and do not have a wide range of correlation coefficients (Gembel 1965; Marshall and Olkin 1967; Freund 1961; Basu and Sun 1997; Block and Basu 1974; Raftery 1984).

Consider the joint exponential CDF in (39). The joint exponential CDF has a more limited admissible range of correlation coefficients than the Gaussian copula. As shown in Fig. 12, the exponential CDF covers only the correlation coefficient ranging from -0.404 to 0, while the approximate joint CDF obtained from the Gaussian copula has a wider range of correlation

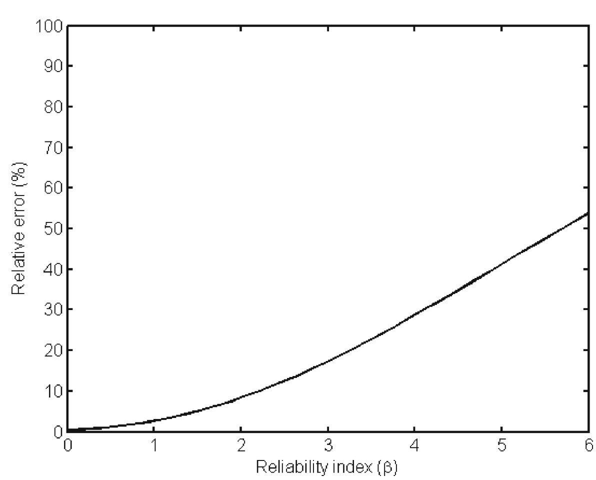

Fig. 10 Relative error in lognormal CDF versus reliability index

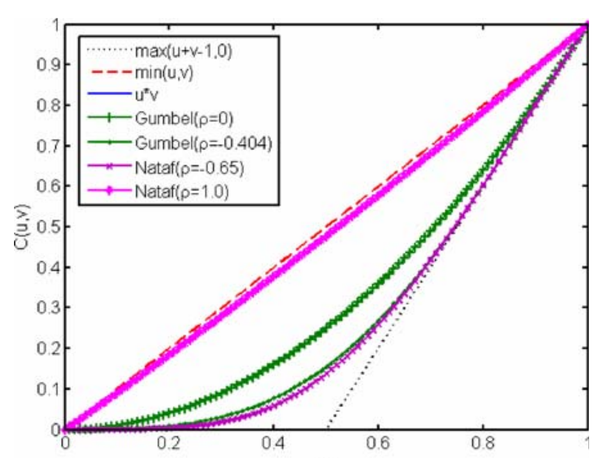

Fig. 12 Diagonal contour diagrams for different copulas

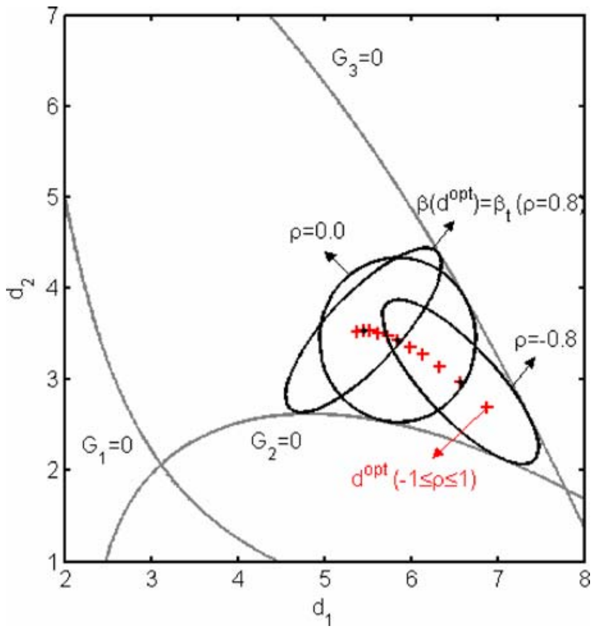

Fig. 13 Optimal design points and reliability target contours for different correlation coefficients.

coefficients, from −0.65 to 1.0, which is obtained from (15) for two exponential variables as

$$\rho'_{ij} = \rho_{ij} R_{ij} = \rho_{ij}\left(1.229 - 0.367\rho_{ij} + 0.153\rho_{ij}^2\right) \quad (44)$$

When the Gaussian copula is used to approximate joint exponential CDF, it provides a limited range of correlation between variables, but for other types of non-normal variables, it provides a much larger range of correlation coefficients (Liu and Der Kiureghian 1986) than for the exponential variables. Thus, the Nataf transformation provides a large admissible range of the correlation coefficients for most non-normal variables. In particular, for normal variables, since the Nataf transformation originates from the Gaussian copula, it can be used for the full range of the correlation coefficients, from −1.0 to 1.0, which corresponds to the lower and upper bound copulas, respectively. In addition, it can accurately estimate the joint lognormal CDF for a wide range of correlation coefficients if the coefficient of variation and the target reliability index are small, i.e., less than three.

Finally, since the normal CDF and lognormal CDFs with small coefficients of variation cover broader application areas, such as material properties (strength), fatigue life, chemical process, fatigue, crack propagation, and loads (Tobias and Trindade 1995; Hahn and Shapiro 1994), in terms of applicability to broad industrial applications, the Nataf transformation is valuable. For engineering applications with joint non-normal CDFs, an alternative copula needs to be used. Selection of a copula for given data and application of the selected copula to RBDO will be our next research topics.

7 Numerical examples

7.1 Mathematical example

Consider a mathematical problem with input random variables $X_i \sim N\left(5.0, 0.3^2\right)$, $i = 1, 2$. The RBDO formulation is defined as

$$\begin{aligned}
&\text{min. cost}(\mathbf{d}) = -d_1 + d_2 \\
&\text{st. } P\left(G_i(\mathbf{X}) > 0\right) \le \Phi(-\beta_{t_i}),\ i = 1, 2, 3 \\
&0 \le d_1 \le 10,\ 0 \le d_2 \le 10,\ \beta_{t_i} = 3.0 \\
&G_1(\mathbf{X}) = 1 - X_1^2 X_2/20 \\
&G_2(\mathbf{X}) = 1 - (X_1 + X_2 - 5)^2/30 - (X_1 - X_2 - 12)^2/120 \\
&G_3(\mathbf{X}) = 1 - 80/(X_1^2 + 8X_2 + 5).
\end{aligned} \quad (45)$$

Table 2 Optimum designs for the mathematical problem

ρ	X_1	X_2	Cost	No. of FE
−1.0	6.886	2.695	−4.191	38 + 38
−0.8	6.571	2.971	−3.601	51 + 51
−0.6	6.335	3.147	−3.188	43 + 43
−0.4	6.145	3.271	−2.874	44 + 44
−0.2	5.985	3.362	−2.623	44 + 44
0.0	5.846	3.433	−2.414	44 + 44
0.2	5.726	3.482	−2.244	44 + 44
0.4	5.622	3.516	−2.105	44 + 44
0.6	5.530	3.535	−1.995	44 + 44
0.8	5.452	3.540	−1.912	44 + 44
1.0	5.387	3.531	−1.856	34 + 34

No. of FE = function evaluation + sensitivity calculation

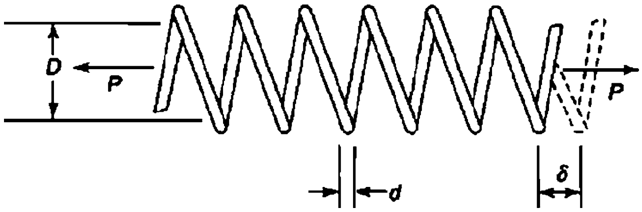

Fig. 14 Coil spring

Using the inverse reliability analysis method PMA+ (Youn et al. 2005a, b) for RBDO, the reliability-based optimum design can be obtained for the different correlation coefficients that range from −1.0 to 1.0 at 0.2 intervals. In Fig. 13, the circle indicates the norm of the standard random variables **U**, called the reliability target contour, where the radius is the target reliability index when the input variables are independent. However, when input variables are correlated, the circle becomes an ellipse, which has either positive or negative angle according to the sign of the correlation coefficient. For selected correlation coefficients that range from −1.0 to 1.0, reliability-based optimum designs are obtained as shown by the "+" sign in Fig. 13. The dots are the optimal design points where the correlation coefficients are −0.8, 0.0, and 0.8, respectively.

Table 2 shows the optimum designs, optimum costs, and number of function evaluations for the different correlation coefficients. In the table, all optimum designs have two active constraints G_2 and G_3, and the optimum designs and the corresponding optimum costs significantly depend on the correlation coefficients.

7.2 Coil spring problem

In next example, an engineering problem is used to show how the correlation in input variables affects RBDO results. The coil springs are widely used in practical applications. The design objective of the coil spring (Fig. 14) is to minimize the mass to carry a given axial load such that the design satisfies the minimum deflection and allowable shear stress requirement, and the surge wave frequency is above the lower limit (Arora 2004).

In this example, five design parameters, which are the mean inner diameter of coil spring (D), wire diameter (d), number of active coils (N), shear modulus (G), and mass density of material (ρ), are selected. Other data are given as: weight density of spring material, $\gamma = 0.285$ lb/in.^{3}; shear modulus, $G = (1.15 \times 10^7)$ lb/in.^{2}; allowable shear stress, $\tau_a = 80{,}000$ lb/in.^{2}; number of inactive coils, $Q = 2$; applied load, $P = 10$ lb; minimum spring deflection, $\Delta = 0.5$ in.; and lower limit of surge wave frequency. The design and random variables such as number of active coils (X_1), coil inner diameter (X_2), wire diameter (X_3), mass density of material (X_4), and shear modulus (X_5) have normal CDFs and have the properties shown in Table 3.

As stated before, the constraints shown in (46) through (48) must be satisfied to carry out a given axial load without material failure. The first constraint is that the deflection δ under the load P should be at least Δ as

$$\delta = \frac{P}{K} = \frac{8P(D+d)^3N}{d^4G} \geq \Delta \tag{46}$$

The second constraint is that the shear stress in the wire should not be larger than τ_a, which is formulated as

$$\tau = \frac{8kP(D+d)}{\pi d^3} = \frac{8P(D+d)}{\pi d^3} \times \left(\frac{4(D+d)-d}{4D} + \frac{0.615d}{D+d}\right) \leq \tau_a \tag{47}$$

where k is Wahl stress concentration factor. The third constraint requires that the surge wave frequency of the spring should be higher than ω_0 as

$$\omega = \frac{d}{2\pi N(D+d)^2}\sqrt{\frac{G}{2\rho}} \geq \omega_0 \tag{48}$$

Table 3 Properties of design and random variables for the coil spring

Random variable	Standard deviation	d_L	d	d_U
X_1	0.1	7.0	30.0	100.0
X_2	0.1	0.8	1.0	1.5
X_3	0.01	0.1	0.15	0.3
X_4	0.10E − 4	0.85E − 4	7.38E − 4	10.0E − 4
X_5	0.10E + 7	0.10E + 7	1.15E + 7	3.00E + 7

Table 4 RBDO results for different correlation coefficients

ρ_{23}	0.0	0.2	0.4	0.6	0.7
X_1	0.107 E + 2	0.990 E + 1	0.821 E + 1	0.798 E + 1	0.784 E + 1
X_2	0.137 E + 1	0.134 E + 1	0.137 E + 1	0.127 E + 1	0.121 E + 1
X_3	0.111 E + 0	0.110 E + 0	0.110 E + 0	0.106 E + 0	0.104 E + 0
X_4	0.850 E − 4	0.850 E − 4	0.850 E − 4	0.850 E − 4	0.850 E − 4
X_5	0.115 E + 8	0.115 E + 8	0.115 E + 8	0.115 E + 8	0.115 E + 8
G_1	0.000	0.000	0.000	0.000	0.000
G_2	−0.055	−0.070	−0.085	−0.063	−0.051
G_3	−0.389	−0.621	−0.993	−1.430	−1.811
No. of FE	30 + 30	30 + 30	30 + 30	30 + 30	30 + 30
Cost	4.863	4.386	3.825	3.224	2.905

No. of FE = function evaluation + sensitivity calculation

Using the data and normalized constraints for the coil spring problem, the RBDO formulation is defined as

$$\text{min. mass}(\mathbf{d}) = 25000 \times (d_1 + Q)\pi^2 (d_2 + d_3) d_3^2 d_4$$

$$\text{st. } P\left(G_i(\mathbf{X}) > 0\right) \leq \Phi(-\beta_t),\ i = 1, 2, 3$$

$$G_1(\mathbf{X}) = 1.0 - \frac{8P(X_2 + X_3)^3 X_1}{X_3^4 X_5 \Delta}$$

$$G_2(\mathbf{X}) = -1.0 + \frac{8P(X_2 + X_3)}{\pi X_3^3 \tau_a}\left[\frac{(4X_2 + 3X_3)}{4X_2} + \frac{0.615X_3}{(X_2 + X_3)}\right]$$

$$G_3(\mathbf{X}) = 1.0 - \frac{X_3}{2\pi X_1 (X_2 + X_3)^2 w_0}\sqrt{\frac{X_5}{2X_4}}. \tag{49}$$

In the manufacturing process, it may be possible that the coil inner diameter and the wire diameter are correlated, and thus the correlation coefficient between those two variables is considered for RBDO.

Table 4 shows the optimal designs, constraints, function evaluation, and cost for the different correlation coefficients. As shown in the table, the optimum designs and costs significantly depend on the correlation coefficients. To minimize the mass of the spring, the mass density goes to the lower bound and the shear modulus does not change because the third constraint is always inactive and the shear modulus does not affect the cost. From these two examples, it is clear that the correlation should be considered in the RBDO of practical applications.

8 Conclusions

In this paper, an RBDO method that deals with the correlation of input variables is proposed. The Rosenblatt transformation and the Nataf transformation are investigated for applicability to RBDO problems with correlated input variables. The Rosenblatt transformation is a mathematically exact transformation method, but it requires a joint CDF to transform the correlated random variables to the independent standard normal variables, so that it can be used when the joint CDF is available or when input variables are independent. On the other hand, the Nataf transformation approximates the joint CDF using the Gaussian copula. Since the copula only requires marginal CDFs and correlation parameters such as covariance to generate a joint CDF, the joint CDF can be easily constructed for real engineering applications. The Gaussian copula provides an exact joint CDF when the input joint CDF is normal or when the normal and lognormal variables are combined in a two-dimensional case. When the joint CDF is lognormal, the Gaussian copula can accurately construct a joint CDF for positive and some negative correlations with small target reliability index if the coefficient of variation is small. Another advantage of the Gaussian copula is that it covers a wide range of correlation coefficients. In this paper, using the Nataf transformation, RBDO is carried out to solve numerical examples with correlated input variables to demonstrate that the correlation in the input variables significantly influences the optimum results of RBDO.

This paper focused on how much error exists between the true joint CDF and the approximate joint CDF using the Gaussian copula. For future research, other types of copulas for correlated input variables with joint non-normal CDF will be investigated to generate joint CDF and will be applied to RBDO.

Acknowledgment This research is supported by the Automotive Research Center that is sponsored by the US Army TARDEC.

References

Arora JS (2004) Introduction to optimum design, 2nd edn. Elsevier, Amsterdam

Basu AP, Sun K (1997) Multivariate exponential distributions with constant failure rates. J Multivar Anal 61:159–169

Block HW, Basu AP (1974) A continuous multivariate exponential extension. Am Stat Assoc 69:1031–1037

Ditlevsen O, Madsen HO (1996) Structural reliability methods. Wiley, New York

Freund JE (1961) A bivariate extension of the exponential distribution. J Am Stat Assoc 56:971–977

Gembel EJ (1965) Two systems of bivariate extremal distributions. 35th Sessions of the International Statistical Institute no. 69, Beograd

Hahn GJ, Shapiro SS (1994) Statistical models in engineering. Wiley, New York

Hohenbichler M, Rackwitz R (1981) Nonnormal dependent vectors in structural reliability. J Eng Mech ASCE 107:1127–1238

Kotz S, Barlakrishnan N, Johnson NL (2000) Continuous multivariate distribution, vol 1, 2nd edn. Wiley, New York

Lee I, Choi KK, Du L, Gorsich D (2008) Dimension reduction method for reliability-based robust design optimization. Comput Struct 86:1550–1562

Liu P-L, Der Kiureghian A (1986) Multivariate distribution models with prescribed marginals and covariances. Probab Eng Mech 1(2):105–112

Madsen HO, Krenk S, Lind NC (1986) Methods of structural safety. Prentice-Hall, Englewood Cliffs, NJ

Marshall AW, Olkin I (1967) A multivariate exponential distribution. J Am Stat Assoc 62:30–44

Melchers RE (1999) Structural reliability analysis and prediction, 2nd edn. Wiley, New York

Nataf A (1962) Détermination des distributions de probabilités dont les marges sont données. C R Hebd SeÂances Acad Sci 255:42–43

Raftery AE (1984) A continuous multivariate exponential distribution. Commun Stat Theory Methods 13:947–965

Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23:470–472

Roser BN (1999) An introduction to copulas. Springer, New York

Socie DF (2003) Seminar notes: Probabilistic aspects of fatigue. http://www.fatiguecaculator.com

Tobias PA, Trindade DC (1995) Applied reliability, 2nd edn. CRC, Boca Raton, FL

Youn BD, Choi KK, Du L (2005a) Adaptive probability analysis using an enhanced hybrid mean value method. Struct Multidiscipl Optim 29(2):134–148

Youn BD, Choi KK, Du L (2005b) Enriched performance measure approach for reliability-based design optimization. AIAA J 43(4):874–884

Youn BD, Xi Z, Wells LJ (2007) Sensitivity-free probability analysis: eigenvector dimension reduction (EDR) method and its extensive applications. 7th world congress on structural and multidisciplinary optimization, Seoul, Korea, May 21–25, pp 2316–2325