

Electronic supplementary materials

for <https://doi.org/10.1631/jzus.A2200300>

Reliability measure approach considering mixture uncertainties under insufficient input data

Zhenyu LIU^{1,4}, Yufeng LYU^{1,4}, Guodong SA^{2,3,4}✉, Jianrong TAN^{1,4}

¹State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058, China

²Ningbo Research Institute, Zhejiang University, Ningbo 315100, China

³Laboratory of Healthy & Intelligent Kitchen System Integration of Zhejiang Province, Ningbo 315336, China

⁴School of Mechanical Engineering, Zhejiang University, Hangzhou 310058, China

✉ Guodong SA, sgd@zju.edu.cn

S1 Finite Difference Method (FDM)

Finite Difference Method (FDM) is one of the methods used to solve differential equations that are difficult or impossible to solve analytically.

Three basic types are commonly considered: forward, backward, and central finite differences.

Forward difference:

$$\frac{\partial y}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{y(x + \Delta x) - y(x)}{\Delta x} \quad (\text{S9})$$

Backward difference:

$$\frac{\partial y}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{y(x) - y(x - \Delta x)}{\Delta x} \quad (\text{S10})$$

Central difference:

$$\frac{\partial y}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{y(x + \Delta x) - y(x - \Delta x)}{2\Delta x} \quad (\text{S11})$$

Through using FDM, a large number of sampling method are selected, and their corresponding performance functions are calculated, then the sensitivity index can be calculated.

The perturbation step is selected as $\Delta x \approx 0.1\%x$.

S2. The detail of representation method of sparse variables Y

The distribution types ζ and distribution parameters θ for dispersion part \tilde{y} of sparse variables Y are uncertainty variables due to insufficient input data. The available insufficient data of \tilde{y}_i contains α point data $[a_{1,\tilde{y}_i}, \dots, a_{\alpha,\tilde{y}_i}]$. Seven common distribution types (Normal

distribution, Lognormal distribution, Weibull distribution, Gumbel distribution, Gamma distribution, Extreme distribution, Extreme Type II distribution) are considered, let $f_{\tilde{y}_i}(\tilde{y}_i|\zeta_k, \theta)$ denote the PDF of \tilde{y}_i under k -th candidate distribution type ζ_k ($k=1,2,\dots,7$) and distribution parameters θ . For k -th candidate distribution type ζ_k , the likelihood estimation function $L_{\tilde{y}_i}(\zeta_k, \theta)$ (Sankararaman and Mahadevan 2011) based on the prescribed insufficient input data is constructed in Eq. (12).

$$L_{\tilde{y}_i}(\zeta_k, \theta) \sim \prod_{j=1}^{\alpha} f_{\tilde{y}_i}(\tilde{y}_i = a_{j,\tilde{y}_i} | \zeta_k, \theta) \quad (S12)$$

The optimum likelihood estimations of θ under distribution type ζ_k are calculated by maximizing $L_{\tilde{y}_i}(\zeta_k, \theta)$ in Eq. (S12). After calculating the maximum value of $L_{\tilde{y}_i}(\zeta_k, \theta)$, the optimum distribution type are determined using Akaike information criterion (AIC) method, the detailed procedure can be found in Peng's work (Peng et al. 2017). The AIC value of k -th distribution type ζ_k is calculated using Eq. (S13).

$$AIC_k = 2n_k - 2 \ln \left(\prod_{j=1}^{\alpha} f_{\tilde{y}_i}(\tilde{y}_i = a_{j,\tilde{y}_i} | \zeta_k, \theta_{opt,k}) \right) \quad (S13)$$

Where $n_k = 2$ is the amount of estimated distribution parameters for distribution type ζ_k ; $f_{\tilde{y}_i}(\tilde{y}_i = a_{j,\tilde{y}_i} | \zeta_k, \theta_{opt,k})$ is the PDF of j -th sampling points a_{j,\tilde{y}_i} under k -th distribution type ζ_k and corresponding optimum distribution parameters $\theta_{opt,k}$, $\theta_{opt,k}$ is calculated through maximizing $L_{\tilde{y}_i}(\zeta_k, \theta)$ in Eq. (S12).

The AIC values of the seven candidate distribution types are calculated and denoted as $AIC_1, AIC_2, \dots, AIC_7$, respectively. Let AIC_{\min} be the minimum value of these seven values, the probability P_{ζ_k} that the k -th distribution type minimizes the estimated information loss is interpreted using Eq. (S14).

$$P_{\zeta_k} = \exp((AIC_{\min} - AIC_k)/2) \quad (S14)$$

The distribution types with $P_{\zeta_k} > 0.2$ are selected to represent the dispersion part \tilde{y} of sparse variables Y . The weight ratios w_k of these selected distribution types are proportional to the probability P_{ζ_k} , and the summation of these weight ratios is equal to 1.

$$w_k = \frac{P_{\zeta_k}}{\sum P_{\zeta_k}} \quad (S15)$$

After determining the distribution types and corresponding weight ratios, the distribution parameters are determined using the Bayesian model averaging method (Sankararaman and Mahadevan 2013; Nannapaneni et al. 2016). The combined likelihood function for insufficient

input data under multiple distribution types is calculated using Eq. (S16).

$$L_{\tilde{y}_i}(\boldsymbol{\theta}) = \prod_{j=1}^{\alpha} \sum w_k f_{\tilde{y}_i}(\tilde{y}_i = a_{j,\tilde{y}_i} | \zeta_k, \boldsymbol{\theta}) \quad (\text{S16})$$

Where ζ_k is the selected distribution type, w_k is the corresponding weight ratio, and $f_{\tilde{y}_i}(\tilde{y}_i = a_{j,\tilde{y}_i} | \zeta_k, \boldsymbol{\theta})$ is the PDF of j -th sampling points a_{j,\tilde{y}_i} under distribution type ζ_k .

The uncertainties of the distribution parameters $\boldsymbol{\theta}$ are calculated using Bayes' theorem. The PDF $f_{\boldsymbol{\theta},\tilde{y}_i}(\boldsymbol{\theta})$ of uncertain distribution parameters $\boldsymbol{\theta}$ for dispersion part \tilde{y}_i is expressed in Eq. (S17).

$$f_{\boldsymbol{\theta},\tilde{y}_i}(\boldsymbol{\theta}) = \frac{L_{\tilde{y}_i}(\boldsymbol{\theta})}{\int L_{\tilde{y}_i}(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (\text{S17})$$

The dispersion part \tilde{y} are represented using weight summation of multiple types of distribution function, the represented function of sparse variables Y under design points \bar{y} can be determined.

S3. The detail computation procedure of sensitivity analysis for reliability index

The failure probability p_G is an uncertainty variable due to insufficient input data. However, it is also a critical constraint in many reliability-based design optimization problems. Sensitivity analysis is an efficient technique to determinate the optimum individuals in the iteration steps of gradient-based reliability optimization methods. The design sensitivity of failure probability can be obtained using the first-order score function and chain rules in Eq. (S18) (Cho et al. 2016).

$$\frac{\partial}{\partial \bar{\mathbf{d}}} F_{p_G}(\hat{p}_G | \mathbf{a}) = \int_0^{\hat{p}_G} \int_{\Omega_\theta} \int_{\Omega_\psi} f(p_G, \boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{a}) \frac{\partial}{\partial \bar{\mathbf{d}}} \ln f_d(\mathbf{d} | \mathbf{a}) d\boldsymbol{\theta} d\boldsymbol{\psi} d\phi \quad (\text{S18})$$

Where the input variables \mathbf{d} contain X , Y and Z , the design points $\bar{\mathbf{d}}$ are their corresponding mean values \bar{x} , \bar{y} , \bar{z} , respectively. The representation types of X , Y and

Z are different, therefore, the detailed formulations of $\frac{\partial}{\partial \bar{x}_i} F_{p_G}(\hat{p}_G | \mathbf{a})$, $\frac{\partial}{\partial \bar{y}_i} F_{p_G}(\hat{p}_G | \mathbf{a})$, and

$\frac{\partial}{\partial \bar{z}_i} F_{p_G}(\hat{p}_G | \mathbf{a})$ are different.

For the sparse variables Y , which are represented using the weight summation of multiple distribution types ζ with uncertain distribution parameters $\boldsymbol{\theta}$, the sensitivity index is calculated using Eq. (S19).

$$\begin{aligned} \frac{\partial}{\partial \bar{y}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) &= \int_0^{\hat{p}_G} \int_{\Omega_\theta} \int_{\Omega_\psi} f(p_G, \boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{a}) \frac{\partial}{\partial \bar{y}_i} \ln \left[\sum w_{k,i} \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a}) \right] d\boldsymbol{\theta} d\boldsymbol{\psi} d\phi \\ &= \int_0^{\hat{p}_G} \int_{\Omega_\theta} \int_{\Omega_\psi} f(p_G, \boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{a}) \frac{\sum w_{k,i} \frac{\partial}{\partial \bar{y}_i} [\zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})]}{\sum w_{k,i} \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})} d\boldsymbol{\theta} d\boldsymbol{\psi} d\phi \end{aligned} \quad (\text{S19})$$

$$SF(\bar{y}_i, \theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a}) = \frac{\sum w_{k,i} \frac{\partial}{\partial \bar{y}_i} [\zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})]}{\sum w_{k,i} \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})} \quad (S20)$$

$$\frac{\partial}{\partial \bar{y}_i} [\zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})] = \frac{\partial \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})}{\partial \theta_{k,i}^1} \times \frac{\partial \theta_{k,i}^1}{\partial \bar{y}_i} + \frac{\partial \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})}{\partial \theta_{k,i}^2} \times \frac{\partial \theta_{k,i}^2}{\partial \bar{y}_i} \quad (S21)$$

Where $\frac{\partial \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})}{\partial \theta_{k,i}^1}$ and $\frac{\partial \zeta_{k,i}(\theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})}{\partial \theta_{k,i}^2}$ are sensitivity indices between PDF of k -th

distribution type $\zeta_{k,i}$ for Y_i and their corresponding distribution parameter $\theta_{k,i}^1$ and $\theta_{k,i}^2$,

which can be calculated directly based on the PDF of specific distribution types. $\frac{\partial \theta_{k,i}^1}{\partial \bar{y}_i}$ and

$\frac{\partial \theta_{k,i}^2}{\partial \bar{y}_i}$ are the sensitivity indices of distribution parameters $\theta_{k,i}^1$ and $\theta_{k,i}^2$ with respect to mean

value \bar{y}_i , which can be calculated based on correlation relationship between distribution parameters and mean values of specific distribution types. The derivation results of some common distribution types are listed in Ref. (Cho et al. 2017). Therefore, $SF(\bar{y}_i, \theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})$ in Eq. (S20) can be calculated directly according to the uncertainty representation function of \mathbf{Y} at design point \bar{y}_i , which do not need the complicated computation of reliability performance function $G(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$.

The design sensitivity of sparse variables \mathbf{Y} can be calculated using the two-level MCS sampling method of probability of failure probability. In the sampling loop of θ and ψ , the distribution parameters θ for \mathbf{Y} are determinate, $SF(\bar{y}_i, \theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})$ can be calculated according to their weight ratios, distribution types and distribution parameters. The sensitivity index of sparse variables \mathbf{Y} is calculated using Eq. (S22).

$$\frac{\partial}{\partial \bar{y}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) = \frac{1}{N_1 N_2} \sum_{m=1}^{N_1} \sum_{l=1}^{N_2} I_{[0, p_G]} [p_G(\theta^{(m)}, \psi^{(m)}, \mathbf{X}^{(l)}, \mathbf{Y}^{(l)}, \mathbf{Z}^{(l)})] SF(\bar{y}_i, \theta_i^{(m)} | \mathbf{a}) \quad (S22)$$

For the statistical variables \mathbf{X} , the distribution type and distribution parameters are determinate, the PDF $f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)$ is also a determinate value, the sensitivity index

$\frac{\partial}{\partial \bar{x}_i} F_{p_G}(\hat{p}_G | \mathbf{a})$ can be calculated using Eq. (S23).

$$\frac{\partial}{\partial \bar{x}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) = \int_0^{\hat{p}_G} \int_{\Omega_\theta} \int_{\Omega_\psi} f(p_G, \theta, \psi | \mathbf{a}) \frac{\frac{\partial}{\partial \bar{x}_i} f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)}{f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)} d\theta d\psi d\phi \quad (S23)$$

$$SF(\bar{x}_i, \theta_{x,i}^1, \theta_{x,i}^2) = \frac{\frac{\partial}{\partial \bar{x}_i} f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)}{f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)} \quad (S24)$$

The $SF(\bar{x}_i, \theta_{x,i}^1, \theta_{x,i}^2)$ in Eq. (S24) can be calculated directly according to the PDF $f_X(X, \theta_{x,i}^1, \theta_{x,i}^2)$ of statistical variables X , which is independent of distribution parameters θ for Y and auxiliary variables ψ for Z , therefore, $SF(\bar{x}_i, \theta_{x,i}^1, \theta_{x,i}^2)$ are changeable in the outer loop of the two-level sampling method, and the design sensitivity of statistical variables is calculated using Eq. (S25).

$$\frac{\partial}{\partial \bar{x}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) = \frac{1}{N_1 N_2} \sum_{l=1}^{N_2} SF(\bar{x}_i, \theta_{x,i}^1, \theta_{x,i}^2) \sum_{m=1}^{N_1} I_{[0, p_G]} \left[p_G(\theta^{(m)}, \psi^{(m)}, X^{(l)}, Y^{(l)}, Z^{(l)}) \right] \quad (S25)$$

For the interval variables Z , there is not determinate distribution types and distribution parameters, their probability density functions are random, it is difficult to integrate Z using probabilistic analysis method. Therefore, their PDFs are assumed to be weight summation of two distribution functions according to the auxiliary variables ψ in Eq. (S26). The two distribution types are chosen as two common distribution types, which are Normal distribution and Weibull distribution. To represent the randomness of interval variables Z , the auxiliary variables ψ are random variables, which is the same as the calculation of failure probability p_G . Therefore, the sensitivity index of Z can be calculated in Eq. (S27) using the similar method for sparse variables Y .

$$Z_i = \psi_z \zeta_{1,i}(\theta_{1,i}) + (1 - \psi_z) \zeta_{2,i}(\theta_{2,i}) \quad (S26)$$

$$\frac{\partial}{\partial \bar{z}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) = \int_0^{\hat{p}_G} \int_{\Omega_\theta} \int_{\Omega_\psi} f(p_G, \theta, \psi | \mathbf{a}) SF(\bar{z}_i, \psi_z) d\theta d\psi d\phi \quad (S27)$$

$$SF(\bar{z}_i, \psi_z) = \frac{\psi_z \frac{\partial}{\partial \bar{z}_i} [\zeta_{1,i}(\theta_{1,i}^1, \theta_{1,i}^2)] + (1 - \psi_z) \frac{\partial}{\partial \bar{z}_i} [\zeta_{2,i}(\theta_{2,i}^1, \theta_{2,i}^2)]}{\psi_z \zeta_{1,i}(\theta_{1,i}^1, \theta_{1,i}^2) + (1 - \psi_z) \zeta_{2,i}(\theta_{2,i}^1, \theta_{2,i}^2)} \quad (S28)$$

A two-level sampling method can be applied into the calculation of sensitivity index of interval variables Z , which is shown in Eq. (S29).

$$\frac{\partial}{\partial \bar{z}_i} F_{p_G}(\hat{p}_G | \mathbf{a}) = \frac{1}{N_1 N_2} \sum_{m=1}^{N_1} \sum_{l=1}^{N_2} I_{[0, p_G]} \left[p_G(\theta^{(m)}, \psi^{(m)}, X^{(l)}, Y^{(l)}, Z^{(l)}) \right] SF(\bar{z}_i, \psi_{z,i}^{(m)}) \quad (S29)$$

The sensitivity index of statistical variables X , sparse variables Y , and interval variables Z can be calculated along with the calculation of probability of p_G , the calculation consumption of additional terms $SF(\bar{x}_i, \theta_{x,i}^1, \theta_{x,i}^2)$, $SF(\bar{y}_i, \theta_{k,i}^1, \theta_{k,i}^2 | \mathbf{a})$, and $SF(\bar{z}_i, \psi_z)$ are little compared with the computation of performance function $G(X, Y, Z)$. Therefore, the reliability index and sensitivity index can be calculated simultaneously with little additional computational

burden.

S4. Engineering Example : A Planar Ten-Bar Structure

The planar ten-bar structure, as shown in Fig. S1, is used to demonstrate the effectiveness of proposed approach under hybrid uncertainties. The three point loads P_1 , P_2 , and P_3 are uncertain variables. P_1 is a statistical variable with normal distribution type, whose mean value is $\bar{P}_1 = 80\text{kN}$ and standard deviation is 0.6kN . P_2 is a sparse variable, the available information for dispersion part is the 10 random points of normal distribution function $N(0,0.3)\text{kN}$. P_3 is an interval variable which can be represented with $[\bar{P}_3 - 2, \bar{P}_3 + 2]\text{kN}$. The length L of all horizontal and vertical bars is 1 m, the elastic modulus of all bars E is 100 GPa, the section area of every bar is fixed at $A_i = 10^{-3}\text{m}^2$. The performance function is the vertical displacement Δ of node 3 in Eq. (S30).

$$\Delta = \left(\sum_{i=1}^6 \frac{N_i^0 N_i}{A_i} + \sqrt{2} \sum_{i=7}^{10} \frac{N_i^0 N_i}{A_i} \right) \frac{L}{E} \quad (\text{S30})$$

Where N_i is the axial internal force of the bar with number i , and N_i^0 is the axial internal force of the bar with number i when $P_1 = P_3 = 0$ and $P_2 = 1$, the calculation details of Eq. (S30) can be found in Ref. (Wei et al. 2014).

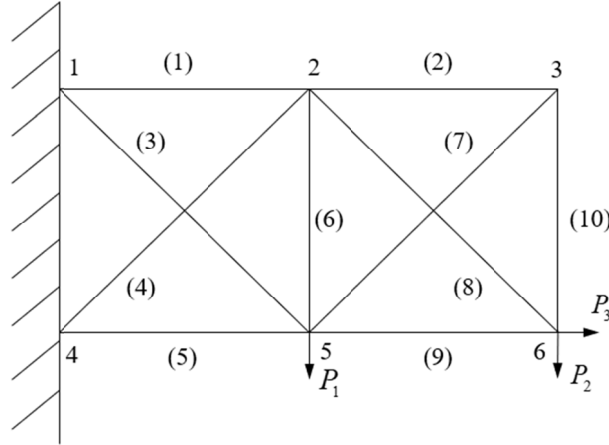


Fig. S1 A planar ten-bar structure

The limit state function is given as $g = 0.003 - \Delta$. At the design point $[\bar{P}_1, \bar{P}_2, \bar{P}_3] = [80, 10, 10]$, the uncertainty of sparse variable P_2 is represented firstly. The selected distribution types for P_2 are Normal and Extreme Value types, and the corresponding weight ratios are 0.6891 and 0.3109, respectively. The PDF of corresponding distribution parameters θ for P_2 are shown in Fig. S2.

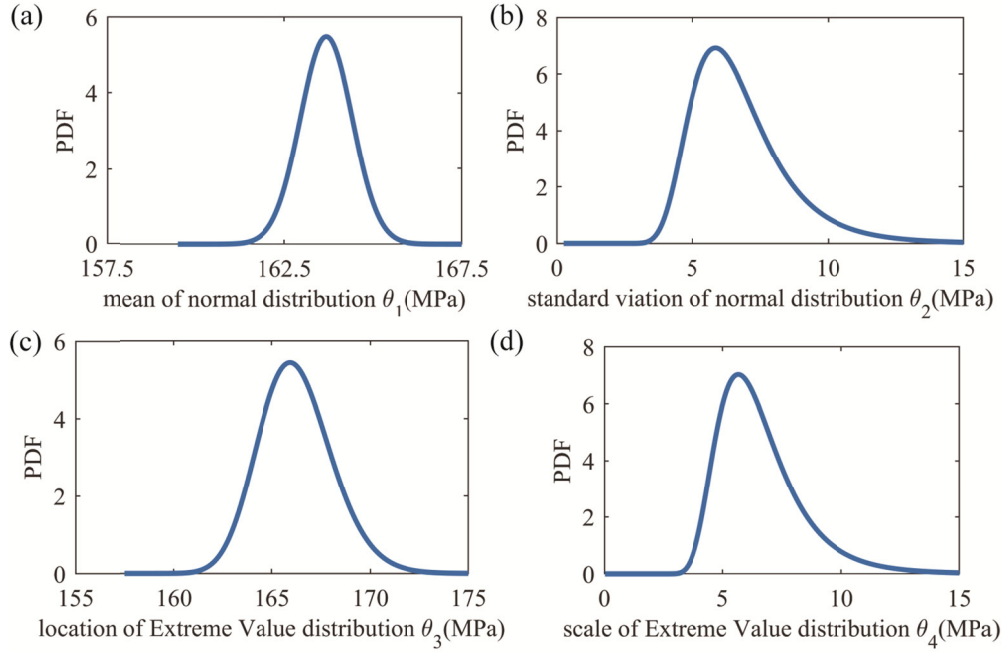


Fig. S2 Probability density function of distribution parameters for sparse variable P_2 : (a) θ_1 ; (b) θ_2 ; (c) θ_3 ; (d) θ_4

The conservativeness level of failure probability considering hybrid uncertainties is calculated using the proposed method and MCS method, and the results are shown in Fig. S3. Results indicate that the proposed method can obtain accurate conservativeness level of failure probability under hybrid uncertainties. However, compared to MCS method, the number of sampling points of distribution parameters and design variables is greatly decreased due to the determination of distribution types of sparse variables and probabilistic transformation of interval variables.

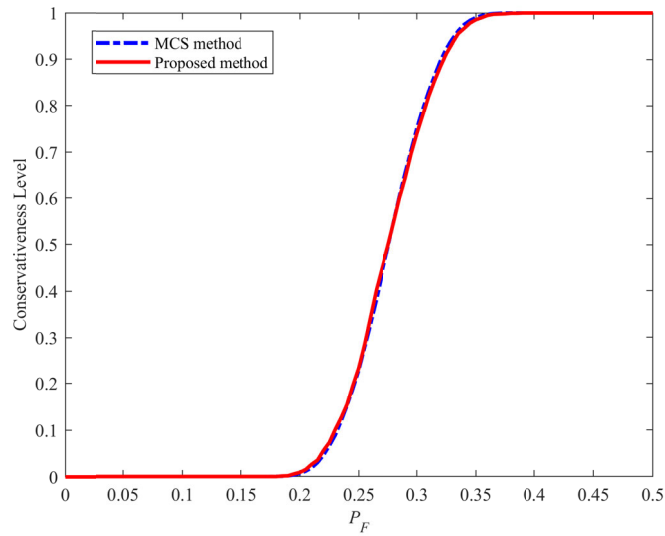


Fig. S3 Reliability measure result of planar ten-bar structure

The design sensitivity for $p_F = 30\%$ at $[\bar{P}_1, \bar{P}_2, \bar{P}_3] = [80, 10, 10]$ are computed using the

proposed method and the finite difference method (FDM), and the results are listed in Table 1. The agreement between the developed sensitivity indices and results of FDM method varies from 100.56% to 105.73%, which indicates that the proposed method can obtain accurate sensitivity results in the engineering problem of planar ten-bar structure.

Table S1 Design sensitivity of conservativeness level for ten-bar structure

	P_1	P_2	P_3	Time (h)
Proposed method	3.304	-5.866	-9.421	3.87
FDM	3.125	-5.833	-9.167	13.38
Agreement	105.73%	100.56%	102.77%	—

Reference

- Cho H, Choi KK, Lamb D (2017) Sensitivity Developments for RBDO With Dependent Input Variable and Varying Input Standard Deviation. *Journal of Mechanical Design* 139(7):071402. <https://doi.org/10.1115/1.4036568>
- Cho H, Choi KK, Lee I, Lamb D (2016) Design Sensitivity Method for Sampling-Based RBDO With Varying Standard Deviation. *Journal of Mechanical Design* 138(1):011405. <https://doi.org/10.1115/1.4031829>
- Nannapaneni S, Hu Z, Mahadevan S (2016) Uncertainty quantification in reliability estimation with limit state surrogates. *Struct Multidisc Optim* 54(6):1509–1526. <https://doi.org/10.1007/s00158-016-1487-1>
- Peng X, Li J, Jiang S (2017) Unified uncertainty representation and quantification based on insufficient input data. *Struct Multidiscip O* 56(6):1305–1317. <https://doi.org/10.1007/s00158-017-1722-4>
- Sankararaman S, Mahadevan S (2013) Distribution type uncertainty due to sparse and imprecise data. *Mech Syst Signal Pr* 37(1–2):182–198. <https://doi.org/10.1016/j.ymsp.2012.07.008>
- Sankararaman S, Mahadevan S (2011) Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. *Reliability Engineering & System Safety* 96(7):814–824. <https://doi.org/10.1016/j.ress.2011.02.003>