



TITLE:

Gaussian mixture model for robust design optimization of planar steel frames

AUTHOR(S):

Do, Bach; Ohsaki, Makoto

CITATION:

Do, Bach ...[et al]. Gaussian mixture model for robust design optimization of planar steel frames. *Structural and Multidisciplinary Optimization* 2021, 63(1): 137-160

ISSUE DATE:

2021-01

URL:

<http://hdl.handle.net/2433/267868>

RIGHT:

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s00158-020-02676-3>; The full-text file will be made open to the public on 14 July 2021 in accordance with publisher's 'Terms and Conditions for Self-Archiving'; This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。

Submitted to Structural and Multidisciplinary Optimization

Gaussian mixture model for robust design optimization of planar steel frames

Bach Do^{1*} and Makoto Ohsaki¹

¹ Department of Architecture and Architectural Engineering, Graduate School of Engineering,
Kyoto University, Kyoto-Daigaku Katsura, Nishikyo, Kyoto 615-8540, Japan

(*Corresponding author email: do.bach.35a@st.kyoto-u.ac.jp)

Abstract

A new method is presented for an application of Gaussian mixture model (GMM) to a multi-objective robust design optimization (RDO) of planar steel frame structures under aleatory (stochastic) uncertainty in material properties, external loads, and discrete design variables. Uncertainty in the discrete design variables is modeled in the wide range between the smallest and largest values in the catalog of the cross-sectional areas. A weighted sum of Gaussians is statistically trained based on the sampled training data to capture an underlying joint probability distribution function (PDF) of random input variables and the corresponding structural response. A simple regression function for predicting the structural response can be found by extracting the information from a conditional PDF, which is directly derived from the captured joint PDF. A multi-objective RDO problem is formulated with three objective functions, namely, the total mass of the structure, and the mean and variance values of the maximum inter-story drift under some constraints on design strength and serviceability requirements. The optimization problem is solved using a multi-objective genetic algorithm utilizing the trained GMM for calculating the statistical values of objective and constraint functions to obtain Pareto-optimal solutions. Since the three objective functions are highly conflicting, the best trade-off solution is desired and found from the obtained Pareto-optimal solutions by performing fuzzy-based compromise programming. Robustness and feasibility of the proposed method for finding the RDO of planar steel frame structures with discrete variables is demonstrated through two design examples.

Keywords: Gaussian mixture model; Uncertainty; Robust design optimization; Steel frame; Genetic algorithm.

1 Introduction

Uncertainties in the structural engineering field include aleatory and epistemic uncertainties; the former comes from the natural randomness of geometric dimensions, material properties, external loads, and boundary conditions, and the latter is from insufficient knowledge of designers about simplified models used for analyzing and designing structures (Kiureghian and Ditlevsen 2009; Roy and Oberkampf 2011). They must be considered to estimate the variation in the structural responses, because the variation measures correspond to the safety of the structural system. In the current practical design procedures of building frames, these uncertainties have been treated by the use of partial factors that are strength reduction and load factors for considering random variations in mechanical resistances of structural members and in external loads, respectively (AISC 2016b; ASCE 2017). However, since the partial factors have been calibrated using some simple probability distributions of input variables with different structural failure scenarios, the safety measures of building frames may involve uncertainty. Therefore, using the partial factors without uncertainty may lead to a poor solution for an optimization problem that is usually on the boundary of the feasible region, and accordingly, is sensitive to even small fluctuations in the input variables (Chateaufneuf 2008). Thus, reliability-based design optimization (RBDO) and robust design optimization (RDO) have been extensively investigated for considering uncertainties in the structural optimization field (Qu and Haftka 2004; Schuëller and Jensen 2008; Ito *et al.* 2018).

Under uncertain input variables, RBDO searches for a balance between the reduction of a structural cost and a confidence level in the structural safety measure by controlling the probability of failure of the structure below an allowable value, which is specified by a design code. Meanwhile, RDO aims at finding a solution that minimizes an objective function considering variation in the structural performance with respect to the uncertain parameters (Ben-Tal *et al.* 2009). When the bound in the

constraints of the RDO problem cannot be specifically determined, it is natural to move some constraint functions to objective functions and formulate the RDO problem with several conflicting performance measures, e.g., the structural cost and statistical properties of the structural responses (Beyer and Sendhoff 2007). Therefore, the RDO problem is to be formulated as a multi-objective optimization problem so that the designers can carry out trade-off analysis among the conflicting performance measures by finding a set of nondominated solutions called Pareto-optimal solutions (Pareto front). Several methods have been found for solving multi-objective structural optimization problems (Marler and Arora 2004); however, most of them are limited in two objective functions, and it is very difficult to characterize the properties of the Pareto-optimal solutions for a problem with many objective functions. Ohsaki *et al.* (2019) formulated a multi-objective RDO problem with many objective functions based on the order statistics.

The first step of dealing with an RDO problem is to determine uncertainty sources. It is usually done by collecting experimental data and analyzing them statistically. Survey by Hess *et al.* (2002) provides detailed statistical data on the geometry of steel plates and the yield stress, ultimate strength, Young's modulus, and Poisson's ratio of steel materials. Aleatory uncertainty in external loads acting on building structures is described in ASCE (2017). Papadrakakis *et al.* (2005) modeled uncertainty in structural material and external loads by normal and log-normal distributions, respectively. Guo *et al.* (2013) considered variation in boundary conditions for robust structural topology optimization problems via level set approach. Jalalpour *et al.* (2013) investigated the effect of uncertainty in Young's modulus on the solution to reliability-based topology optimization problem of truss structures.

After the uncertainty sources are determined, the second step is to propagate uncertainty through a mechanical model for obtaining statistical parameters of the structural responses required for the optimization problem. The propagation process is repeated until the RDO solution is found. There exist two dominant approaches for this purpose: deterministic and probabilistic (Du and Chen 2000; Elishakoff and Ohsaki 2010). The former approach such as interval analysis is set-based wherein the uncertainty is bounded by a given set of the input variables. This approach is formulated based on rigorous nonlinear programming problem, thereby providing accurate solutions. However, it may lead

to a too conservative design since the bounds for the input variables are generally unknown in advance, and the response of the solution to the RDO problem corresponds to the worst-case variable set. Meanwhile, the latter approach is based on probabilistic properties of the input variables propagated through the mechanical model using, e.g., the Monte-Carlo simulation (MCS). Although this approach can produce a moderately conservative design, it is computationally expensive in general, especially when solving RDO problems with many variables. For both mentioned approaches, the uncertainty may also be propagated through surrogate models, which are used to approximate the structural responses (Jin *et al.* 2003). After these models are developed, statistical moments of the structural responses can be estimated using either Monte-Carlo integration method or Taylor series expansion at the mean values of input variables. The issue of the surrogate-based approach is that high-fidelity surrogate models themselves involve interpolation uncertainty (Arendt *et al.* 2013), which becomes significant when the number of input variables increases. Also, the accuracy of using Taylor series expansion to estimate the statistical moments strongly depends on accuracy in estimating high-order derivatives such as the Hessian of the response function. In the same manner, the polynomial chaos expansion method (Crestaux *et al.* 2009) can be used to fit the structural response by using an approximate model of a linear combination of polynomial basis functions and then utilizing the resulting model to estimate the statistical moments of the response. The multivariate basis functions for problems with multivariate objective or constraint functions are commonly generated by constructing every possible combination of univariate polynomials for each input variable. Therefore, the use of polynomial chaos expansion is subject to the curse of dimensionality, i.e., the number of multivariate basis functions increases with a polynomial order of the number of input variables, which quickly leads to a prohibitive numerical cost of constructing the models and estimating the statistical moments.

The problem of minimizing the total mass of steel frames under some constraints on structural strength and serviceability requirements has been well investigated using heuristic methods such as genetic algorithm (Pezeshk *et al.* 2000), simulated annealing (Ohsaki *et al.* 2007), particle swarm optimization (Doğan and Saka 2012), teaching-learning method (Toğan 2012), and harmony search (Maheri and Narimani 2014); however, there have been few studies on multi-objective RDO problems

with more than two objectives. Papadrakakis *et al.* (2005) minimized the total mass and standard deviation of the total drift of a steel truss structure using a two-phase analysis, i.e., a training phase with a neural network and an optimization phase with an evolutionary algorithm. Aleatory uncertainty was considered in the cross-sectional dimensions, material properties, and external loads. The standard deviation of the total drift and mean values of constraints were evaluated using the MCS sampling on the obtained neural network model. Lagaros and Papadrakakis (2007) minimized the total mass and standard deviation of the maximum inter-story drift of a steel frame structure against seismic loading with a log-normal distribution. They directly used the MCS and finite element analysis to estimate the standard deviation of the objective and mean values of the constraints required for the RDO problem. Liu *et al.* (2013) formulated a multi-objective RDO problem for minimizing the total mass, mean value of seismic demand, and variation of seismic demand for a steel moment-resisting frame against seismic loading with a log-normal distribution. The evaluation of mean and variation values of seismic demand was not explicitly shown. Recently, Ohsaki *et al.* (2019) combined the tolerance interval and confidence interval of order statistics and the multi-objective RDO to minimize twenty order statistics of the maximum inter-story drift of a shear frame against seismic motions. They considered uncertainty in the story stiffness, floor mass, and damping coefficient. However, most of the previous structural RDO studies use the MCS to assist in the iterative computations for the optimization process and do not aim at statistically understanding the cause-effect relationship between the random inputs and the corresponding structural response for application to the optimum design. Once the cause-effect relationship is established, the evaluation of the randomness, such as mean, variance, and also the probability distribution of the objective and constraint functions can reduce the computational effort. In this study, therefore, we start by constructing probabilistic models for the input variables and structural responses. Then, we utilize these models for solving a multi-objective RDO problem.

Gaussian mixture model (GMM) is a probabilistic model using a weighted sum of Gaussians for exploring an underlying joint probability distribution function (PDF) of random input variables and the corresponding output variables (responses) (Sung 2004; Hastie *et al.* 2009; McLachlan *et al.* 2019). Many powerful abilities of the GMM have been demonstrated in the field of statistical learning. It can

transform non-Gaussian random input variables to a weighted sum of Gaussians, thereby enabling an analytical solution for a moment-generating function, which is used to characterize probabilistic properties and employed in further reliability analyses of structural systems (Papadimitriou *et al.* 2018). Furthermore, the model has been demonstrated to provide simple regression functions that facilitate the calculations of the mean, variation, and higher-order central moments of the response of interest by reducing the computational effort in calculating the gradient and Hessian of the response. The GMM is characterized by a set of unknown parameters, namely, the mixing proportion vector, the number of Gaussians, the mean vectors, and the covariance matrices. A drawback of the GMM is its sensitivity to the amount of training data; therefore, an approach is required for determination of appropriate number of training data.

In this study, we formulate a multi-objective RDO problem for planar steel frames against static loads in the presence of aleatory uncertainty in material properties, external loads, and design variables. Let \mathbf{s} denote the vector of discrete design variables, which represents the set of pairs of nominal value of cross-sectional area and moment of inertia of structural members. Uncertainty in \mathbf{s} is modeled in the wide range between the smallest and largest values in the catalog of the cross-sectional areas, although the uncertainty for each selected discrete value is restricted in a small range. Let \mathbf{u} denote the vector of random parameters, i.e., material properties and external loads. Let f_1 , f_2 , and f_3 denote the total mass of the structure, and the mean and variance values of the limit state function regarding the maximum inter-story drift, respectively. The multi-objective RDO problem is formulated with three conflicting objectives f_1 , f_2 , and f_3 under some constraints on design strength and serviceability requirements, denoted by g_i , and on constructional requirements, denoted by h_j , as follows:

$$\min. f_1(\mathbf{s}), f_2(\mathbf{s}, \mathbf{u}), f_3(\mathbf{s}, \mathbf{u})$$

subject to

$$g_i(\mathbf{s}, \mathbf{u}) \leq 0, \quad i = 1, 2, \dots, n$$

$$h_j(\mathbf{s}) \leq 0, \quad j = 1, 2, \dots, m$$

(1)

As \mathbf{s} and \mathbf{u} involve randomness, the constraint functions are generic functions with some statistical metrics such as the mean and variance as well as the probability distribution. This study uses the mean metric to measure the randomness of the constraint functions. Detailed derivations of the objective and constraint functions are given in Section 4.1.

Since the set-based and probabilistic approaches (Beyer and Sendhoff 2007) may lead to either an excessively conservative design or a computationally expensive design process, we utilize the GMM in solving problem (1) to reduce the computational cost of estimating statistical values of the objective and constraint functions, while incorporating probabilistic propagation of random input variables. We first generate a dataset of input and output values for training the GMM using the MCS. Uncertain design variables and random parameters are equivalently considered in the formulation of the GMM. Before application to our design optimization problem, the accuracy of GMM performance is confirmed by three numerical experiments. The optimization problem is solved using a multi-objective genetic algorithm utilizing the trained GMM for calculating the statistical values of the objective and constraint functions to obtain Pareto-optimal solutions. Various methods have been proposed for selecting the best solution from the set of Pareto-optimal solutions (Greco *et al.* 2016), e.g., PROMETHEE-II method (Brans *et al.* 1986; Ma *et al.* 2019). However, the most appropriate approach depends on the problem type (Chen and Hwang 1992). The best trade-off solution is found by performing fuzzy-based compromise programming on the Pareto-optimal solutions since its effectiveness has been verified by many authors (Abido 2003; Vahdani *et al.* 2013). We demonstrate the robustness and feasibility of the proposed method for finding the robust optimal designs of two planar steel frame structures.

Our main contributions to the RDO of structures are as follows. (1) We present the first attempt at utilizing the GMM for multi-objective RDO problem of planar steel frames to facilitate the computation of the statistical values of objective and constraint functions, which is still a challenging task since the design optimization of steel frames is a combinatorial problem. Once the GMMs are constructed, the statistical moments efficiently arrive with analytical results, which yields a quick convergence to the RDO solution. (2) We introduce the approach to consider uncertainty in discrete design variables that can be easily extracted from the GMM trained only once before carrying out optimization for a wide

range of uncertainty under various conditions. (3) Since it equally treats the randomness of design variables and parameters during its construction, we demonstrate that the GMM can be effectively used to solve the optimization problems with different bounds of constraint functions and training should be carried out only once before solving the optimization problem with various bounds of constraints.

The remaining part of this paper is organized as follows. Section 2 provides a brief description of the concept of the GMM and the procedure to obtain the model for a particular problem. To verify the performance of the GMM and provide a proper number of training data points, we conduct three numerical experiments in Section 3, where a known bivariate function, a two-bar truss, and a planar steel frame are investigated. In Section 4, we formulate the RDO problem, describe the methods used for generating the training data and obtaining the best compromise solution, and introduce the optimization procedure. We apply the proposed method to the RDO of two planar steel frames in Section 5. Finally, Section 6 summarizes and concludes this paper.

2 Gaussian mixture model

Theoretical aspects of the GMM and the determination of its parameters are briefly described in this section for completeness of the paper. The readers may refer to Sung (2004) and McLachlan *et al.* (2019) for detailed derivations.

2.1 Outline of GMM

Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ be a set of N training data, where $\mathbf{x}_i \in \mathbf{X} \in \mathbb{R}^D$ are D -dimensional vectors of input variables and $y_i \in Y \in \mathbb{R}$ are the corresponding observed output variables. We assume that the training data emerges from an underlying joint PDF of \mathbf{X} and Y defined by $p_{\mathbf{X},Y}(\mathbf{x}, y)$. We seek for a transformation from the space of the input variables to that of the output variable through a conditional PDF $p_{Y|\mathbf{X}}(y | \mathbf{x})$, which is derived from the joint PDF $p_{\mathbf{X},Y}(\mathbf{x}, y)$ and the marginal PDF $p_{\mathbf{X}}(\mathbf{x})$, as follows:

$$p_{Y|\mathbf{X}}(y|\mathbf{x}) = \frac{p_{\mathbf{X},Y}(\mathbf{x}, y)}{p_{\mathbf{X}}(\mathbf{x})} \quad (2)$$

The joint PDF of \mathbf{X} and Y is assumed to be a weighted sum of Gaussians as

$$p_{\mathbf{X},Y}(\mathbf{x}, y) = \sum_{k=1}^K \pi_k \phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (3)$$

$$\sum_{k=1}^K \pi_k = 1, \quad 0 \leq \pi_k \leq 1 \quad (4)$$

$$\boldsymbol{\mu}_k = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{X},k}^T & \mu_{Y,k} \end{bmatrix}^T \quad (5)$$

$$\boldsymbol{\Sigma}_k = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k} & \boldsymbol{\Sigma}_{\mathbf{X}Y,k} \\ \boldsymbol{\Sigma}_{Y\mathbf{X},k} & \boldsymbol{\Sigma}_{YY,k} \end{bmatrix} \quad (6)$$

where $\phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \square \mathcal{N}_{D+1}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the k th $(D+1)$ -variate Gaussian; K is the number of Gaussians; π_k , $\boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ are the mixing proportion, mean vector, and variance matrix of the k th Gaussian, respectively. They are unknown parameters of the GMM.

The k th Gaussian $\phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ in Eq. (3) can be decomposed as (Sung 2004)

$$\phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \phi(y|\mathbf{x}; m_k(\mathbf{x}), \sigma_k^2) \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}) \quad (7)$$

where m_k and σ_k^2 are the conditional mean and conditional variance defined by

$$m_k(\mathbf{x}) = \mu_{Y,k} + \boldsymbol{\Sigma}_{Y\mathbf{X},k} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X},k}) \quad (8)$$

$$\sigma_k^2 = \boldsymbol{\Sigma}_{YY,k} - \boldsymbol{\Sigma}_{Y\mathbf{X},k} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}^{-1} \boldsymbol{\Sigma}_{\mathbf{X}Y,k} \quad (9)$$

Substituting $\phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ in Eq. (7) into Eq. (3), the joint PDF can be rewritten as

$$p_{\mathbf{X},Y}(\mathbf{x}, y) = \sum_{k=1}^K \pi_k \phi(y|\mathbf{x}; m_k(\mathbf{x}), \sigma_k^2) \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}) \quad (10)$$

The marginal PDF is directly derived from the joint PDF as

$$\begin{aligned} p_{\mathbf{X}}(\mathbf{x}) &= \int_{-\infty}^{+\infty} p_{\mathbf{X},Y}(\mathbf{x}, y) dy = \int_{-\infty}^{+\infty} \sum_{k=1}^K \pi_k \phi(y|\mathbf{x}; m_k(\mathbf{x}), \sigma_k^2) \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}) dy \\ &= \sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X},k}) \end{aligned} \quad (11)$$

The conditional PDF in Eq. (2) now becomes

$$p_{Y|\mathbf{X}}(y|\mathbf{x}) = \frac{\sum_{k=1}^K \pi_k \phi(y|\mathbf{x}; m_k(\mathbf{x}), \sigma_k^2) \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X},k})}{\sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X},k})} = \sum_{k=1}^K w_k(\mathbf{x}) \phi(y|\mathbf{x}; m_k(\mathbf{x}), \sigma_k^2) \quad (12)$$

where

$$w_k(\mathbf{x}) = \frac{\pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X},k})}{\sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{X},k}, \boldsymbol{\Sigma}_{\mathbf{X},k})} \quad (13)$$

is the mixing weight.

The conditional mean function (regression function) of y , denoted by $g(\mathbf{x})$, is derived from Eq. (12) as

$$g(\mathbf{x}) = \mathbb{E}[y|\mathbf{x}] = \sum_{k=1}^K w_k(\mathbf{x}) m_k(\mathbf{x}) \quad (14)$$

where $m_k(\mathbf{x})$ and $w_k(\mathbf{x})$ are given in Eqs. (8) and (13), respectively.

To compute the mean and variance of $g(\mathbf{x})$ in Eq. (14) under the effect of the randomness in input variables $\mathbf{x} = [x_1, x_2, \dots, x_D]^T$, we utilize the second-order Taylor expansion (Beyer and Sendhoff 2007)

with respect to \mathbf{x} at the mean $\boldsymbol{\mu}_{\mathbf{X}} = [\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_D}]^T$ as

$$\mathbb{E}[g(\mathbf{x})] = g(\boldsymbol{\mu}_{\mathbf{X}}) + \frac{1}{2} \sum_{i=1}^D \sum_{j=1}^D \frac{\partial^2 g}{\partial x_i \partial x_j} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \sigma_{ij} \quad (15)$$

$$\begin{aligned} \text{var}[g(\mathbf{x})] &= \sum_{i=1}^D \sum_{j=1}^D \frac{\partial g}{\partial x_i} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \frac{\partial g}{\partial x_j} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \sigma_{ij} + \sum_{i=1}^D \sum_{j=1}^D \sum_{k=1}^D \frac{\partial g}{\partial x_i} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \frac{\partial^2 g}{\partial x_j \partial x_k} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \sigma_{ijk} \\ &\quad + \frac{1}{4} \sum_{i=1}^D \sum_{j=1}^D \sum_{k=1}^D \sum_{l=1}^D \frac{\partial^2 g}{\partial x_i \partial x_j} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} \frac{\partial^2 g}{\partial x_k \partial x_l} \Big|_{\boldsymbol{\mu}_{\mathbf{X}}} (\sigma_{ijkl} - \sigma_{ij} \sigma_{kl}) \end{aligned} \quad (16)$$

where σ_{ij} is the covariance of x_i and x_j ; σ_{ijk} is the third-order co-moment of x_i , x_j , and x_k ; and σ_{ijkl} is the fourth-order co-moment of x_i , x_j , x_k , and x_l . The calculations of the mean and variance of $g(\mathbf{x})$ are based on calculations of the gradient and Hessian of the GMM, which are derived in the Appendix.

2.2 Determination of GMM parameters

Unknown GMM parameters are determined by employing an expectation-maximization (EM) algorithm. It is an iterative solver and widely used to maximize a log-likelihood function of training data, thereby estimating unknown parameters of statistical models (Hastie *et al.* 2009; McLachlan *et al.* 2019). The log-likelihood function of a GMM, which measures how well a statistical model fits the observed training samples, is defined as

$$\mathcal{L} = \sum_{i=1}^N \log \left[\sum_{k=1}^K \pi_k \phi(\mathbf{x}_i, y_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right] \quad (17)$$

We maximize \mathcal{L} to find the best fit statistical model. However, since we cannot simply encapsulate the log function into the inner summation in Eq. (17), the maximization of \mathcal{L} becomes difficult. To overcome this difficulty, the latent variables $\mathbf{Z} = [z_1, z_2, \dots, z_N]^T$ are introduced, where $z_i \in \{1, 2, \dots, K\}$ (Hastie *et al.* 2009). If $z_i = k$, the i th data point belongs to the k th Gaussian. Therefore, the joint PDF of the complete set of N training data $\{(\mathbf{x}_i, y_i, z_i)\}_{i=1}^N$ is simplified as (Hastie *et al.* 2009)

$$p_{\mathbf{x}, y, \mathbf{Z}}(\mathbf{x}, y, \mathbf{z}) = \prod_{i=1}^N \prod_{k=1}^K [\pi_k \phi(\mathbf{x}_i, y_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)]^{z_{ik}} \quad (18)$$

where

$$z_{ik} = \begin{cases} 1, & \text{if } z_i = k \\ 0, & \text{if } z_i \neq k \end{cases} \quad (19)$$

The log-likelihood of the complete data becomes

$$\mathcal{L}_c = \sum_{i=1}^N \sum_{k=1}^K z_{ik} \log [\pi_k \phi(\mathbf{x}_i, y_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)] \quad (20)$$

E-step: Since z_{ik} is unknown in advance, the following conditional expected value \hat{z}_{ik} (Hastie *et al.* 2009) is substituted for z_{ik} in Eq. (20):

$$\hat{z}_{ik} = \frac{\pi_k \phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{k=1}^K \pi_k \phi(\mathbf{x}, y; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} \quad (21)$$

M-step: We maximize $\mathcal{L}_c = \sum_{i=1}^N \sum_{k=1}^K \hat{z}_{ik} \log[\pi_k \phi(\mathbf{x}_i, y_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)]$ under the constraint (4) to update the mixing proportion as

$$\pi_k = \frac{1}{N} \sum_{i=1}^N \hat{z}_{ik} \quad (22)$$

The mean vector $\boldsymbol{\mu}_k$ and the covariance matrix $\boldsymbol{\Sigma}_k$ are also updated as

$$\boldsymbol{\mu}_k = \frac{\sum_{i=1}^N \hat{z}_{ik} \mathbf{d}_i}{\sum_{i=1}^N \hat{z}_{ik}} \quad (23)$$

$$\boldsymbol{\Sigma}_k = \frac{\sum_{i=1}^N \hat{z}_{ik} (\mathbf{d}_i - \boldsymbol{\mu}_k)^T (\mathbf{d}_i - \boldsymbol{\mu}_k)}{\sum_{i=1}^N \hat{z}_{ik}} \quad (24)$$

where $\mathbf{d}_i = [\mathbf{x}_i^T \ y_i]^T$.

Thus, the EM algorithm can be summarized as follows:

Step 1: Fix number of Gaussians K with $1 \leq K \leq 50$ and the tolerance $\varepsilon = 10^{-6}$.

Give initial values $\pi_k^{(0)}$, $\boldsymbol{\mu}_k^{(0)}$, $\boldsymbol{\Sigma}_k^{(0)}$, and initialize the step number as $s = 1$.

Step 2: Compute $\hat{z}_{ik}^{(s)}$ with $\pi_k^{(s-1)}$, $\boldsymbol{\mu}_k^{(s-1)}$, and $\boldsymbol{\Sigma}_k^{(s-1)}$ using Eq. (21).

Step 3: Compute $\pi_k^{(s)}$ and $\boldsymbol{\mu}_k^{(s)}$ with $\hat{z}_{ik}^{(s)}$ using Eqs. (22) and (23), respectively.

Step 4: Compute $\boldsymbol{\Sigma}_k^{(s)}$ with $\boldsymbol{\mu}_k^{(s)}$ and $\hat{z}_{ik}^{(s)}$ using Eq. (24).

Step 5: Compare $\boldsymbol{\mu}_k^{(s)}$ and $\boldsymbol{\mu}_k^{(s-1)}$.

If $\|\boldsymbol{\mu}_k^{(s)} - \boldsymbol{\mu}_k^{(s-1)}\| \leq \varepsilon$, stop.

Else $s \leftarrow s + 1$ and return to Step 2.

For a specified value of K , the initial mixing proportions are uniform, the initial mean vectors are K vectors that are randomly selected from the training set, and the initial variance matrices for K Gaussians are diagonal, where the i th diagonal element is the variance of the i th input variable. The EM

algorithm is assured to terminate after a finite number of iterations since it never reduces the \mathcal{L} value at each generation (Hastie *et al.* 2009). However, to avoid the ill-conditioning of estimated covariance matrices of the GMM during the implementation of EM algorithm, we add a small positive regularizer of 0.001 to the diagonal components of Σ_k . To perform EM algorithm, the MATLAB R2018a Statistics and Machine Learning Toolbox is utilized (MathWorks 2019).

To determine the number of Gaussians K before implementation of the EM algorithm, we increase it step by step from 1 to 50, thereby producing a total of 50 GMMs for the model selection process, where 50 is a recommended value to limit the complexity of the GMM model (Chamroukhi 2016). The Bayesian information criterion (BIC) is chosen for the model selection process since its effectiveness in determining the number of components for GMMs has been verified by many authors (Campbell *et al.* 1997; Fraley and Raftery 1998). In the BIC context, the best GMM from among 50 GMM candidates minimizes BIC, which is equivalent to selecting the GMM with the largest posterior probability in the Bayesian consideration. The BIC is defined as follows (Hastie *et al.* 2009):

$$\text{BIC} = -\mathcal{L} + \frac{n_p}{2} \log N \quad (25)$$

where \mathcal{L} is the log-likelihood function in Eq. (17); n_p is the number of free parameters required for a total of K Gaussians; for a full covariance matrix, $n_p = (K - 1) + (D + 1)K + 0.5(D + 1)(D + 2)K$ with the three terms corresponding to the numbers of free parameters of K mixing proportions, K mean vectors, and K covariance matrices, respectively.

3 Computer experiments

Before application to the robust optimization problem of steel frames, the performance of the GMM is tested against three numerical experiments, which include a known bivariate conditional PDF, a two-bar truss, and a planar steel frame. In the first experiment, we approximate the known conditional PDF by using Eq. (12). In the second experiment, the GMM is used to solve the RDO problem of a two-bar truss with explicit objective and constraint functions to compare the results with those in existing

studies. In the third example, we again use Eq. (12) to capture a random response of the steel frame and use Eq. (14) to predict this response for different independent samples of test data. If the performance of the GMM does not meet a desired level of accuracy, more training data points should be added.

3.1 Bivariate model

Consider a conditional bivariate PDF taken from Fu *et al.* (2011), as

$$Y | \mathbf{X} \sim \mathcal{N}(x_1 + x_2, 0.15 + 0.05x_1x_2) \quad (26)$$

where x_1 and x_2 are uniformly distributed over the interval of $[-0.3, 0.3]$. To investigate the performance of the GMM in characterizing conditional PDFs, we approximate the PDF in Eq. (26) by using the conditional PDF in Eq. (12). To do so, a total of 4000 samples are generated, in which 2000 samples are randomly selected for training for constructing the GMM and the remaining 2000 samples for testing. Nine pairs of (x_1, x_2) are then selected within the given interval for obtaining the corresponding nine true conditional PDFs. By incorporating y values of the test data and each pair of (x_1, x_2) into Eq. (12), we obtain the estimated conditional PDF for each pair, which is plotted together with the true one in Fig. 1. As a result, the conditional PDF for each of the nine pairs of (x_1, x_2) is estimated based on the joint PDF $p_{\mathbf{x}, y}(\mathbf{x}, y)$ in Eq. (10), which is a mixture of seven Gaussians, i.e., $K = 7$. The mean vectors of seven Gaussians and the corresponding mixing proportions are also obtained as $\boldsymbol{\mu}_1 = [-0.0320, 0.0313, -0.0040]^T$, $\boldsymbol{\mu}_2 = [0.1869, 0.1266, 0.3064]^T$, $\boldsymbol{\mu}_3 = [0.0098, 0.2480, 0.2697]^T$, $\boldsymbol{\mu}_4 = [-0.1595, -0.2753, -0.4451]^T$, $\boldsymbol{\mu}_5 = [0.2399, -0.1456, 0.1066]^T$, $\boldsymbol{\mu}_6 = [-0.0162, -0.1982, -0.2287]^T$, $\boldsymbol{\mu}_7 = [-0.2472, 0.0077, -0.2293]^T$, $\pi_1 = 0.3390$, $\pi_2 = 0.1391$, $\pi_3 = 0.1183$, $\pi_4 = 0.0477$, $\pi_5 = 0.0805$, $\pi_6 = 0.1552$, and $\pi_7 = 0.1202$. It is demonstrated that the GMM can yield a very good estimation of the given conditional PDF.

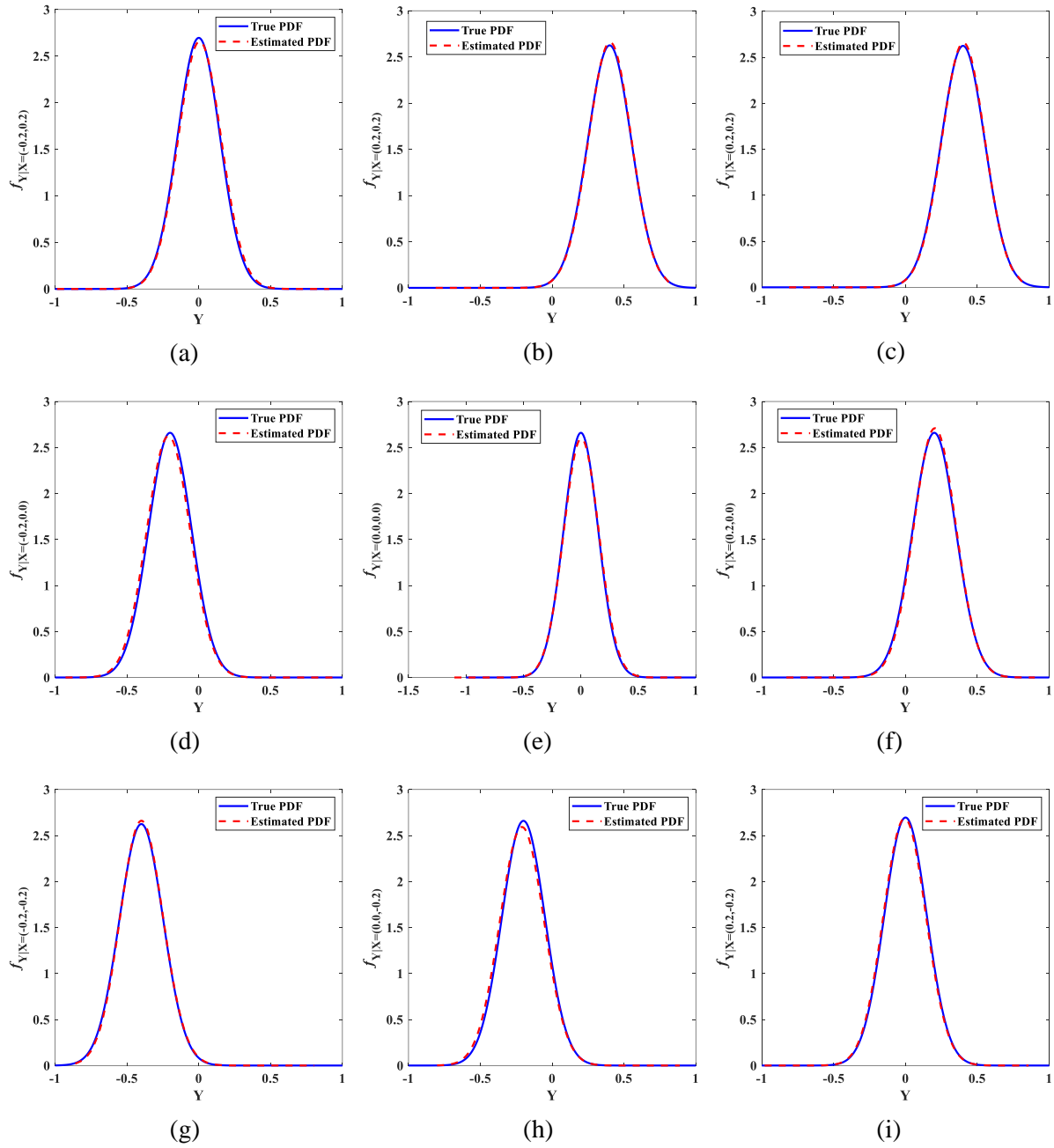


Fig. 1 True and estimated conditional PDFs at nine tested pairs of (x_1, x_2) : (a) $(-0.2, 0.2)$; (b) $(0.0, 0.2)$; (c) $(0.2, 0.2)$; (d) $(-0.2, 0.0)$; (e) $(0.0, 0.0)$; (f) $(0.2, 0.0)$; (g) $(-0.2, -0.2)$; (h) $(0.0, -0.2)$; and (i) $(0.2, -0.2)$.

3.2 Two-bar truss

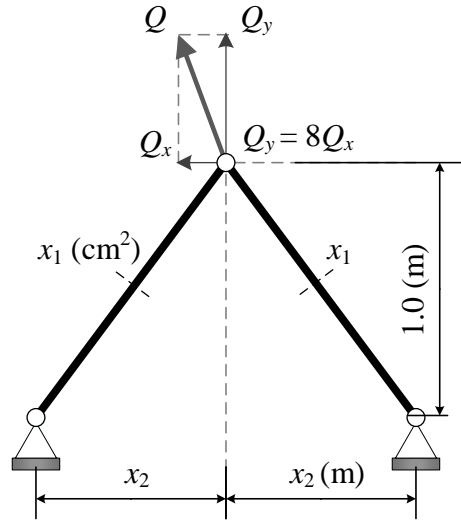


Fig. 2 Two-bar truss.

The GMM is used to solve the RDO problem of a two-bar truss taken from Lee *et al.* (2009) as shown in Fig. 2. The cross-sectional area of the truss members x_1 and the horizontal span of the truss x_2 are considered as the design variables. The density of bar material ρ , the external load Q , and the material's tensile strength F_y are the three random parameters. Statistical properties of the design variables and random parameters are given in Table 1. The deterministic optimization problem of the truss is formulated to minimize the total material volume of the truss subject to constraints on the axial strength of the truss members as follows (Lee *et al.* 2009):

$$\min. f(x_1, x_2) = \rho x_1 \sqrt{1 + x_2^2}$$

subject to

$$g_1(x_1, x_2) = 1 - \frac{5Q}{\sqrt{65}F_y} \sqrt{1 + x_2^2} \left(\frac{8}{x_1} + \frac{1}{x_1 x_2} \right) \geq 0$$

$$g_2(x_1, x_2) = 1 - \frac{5Q}{\sqrt{65}F_y} \sqrt{1 + x_2^2} \left(\frac{8}{x_1} - \frac{1}{x_1 x_2} \right) \geq 0$$

$$0.2 \leq x_1 \leq 20, \quad 0.1 \leq x_2 \leq 1.6 \tag{27}$$

Table 1 Probabilistic characteristics of random variables of the two-bar truss.

Variable	Description	Mean	COV	Distribution
x_1	Cross-sectional area [cm ²]	μ_{x_1}	0.02	Normal
x_2	Horizontal span [m]	μ_{x_2}	0.02	Normal
ρ	Material density [kg/m ³]	10^4	0.20	Beta*
Q	External load [kN]	800	0.25	Gumbel
F_y	Tensile strength [MPa]	1050	0.24	Lognormal

* Shape parameters of beta distribution are set as $a = b = 5$ for generating the training set.

The RDO problem is formulated as follows with a single function using a weighted sum of the mean and variance values of the objective function in (27) (Lee *et al.* 2009):

$$\min. F = w_1 \frac{\mu_f}{\mu_f^*} + w_2 \frac{\sigma_f}{\sigma_f^*}$$

subject to

$$G_1 = k\sigma_{g_1} - \mu_{g_1} \leq 0$$

$$G_2 = k\sigma_{g_2} - \mu_{g_2} \leq 0$$

$$0.2 \leq \mu_{x_1} \leq 20, \quad 0.1 \leq \mu_{x_2} \leq 1.6 \quad (28)$$

where μ and σ denote the mean and standard deviation, respectively; the weight factors $w_1 = w_2 = 0.5$; the scaling factors $\mu_f^* = 10$ and $\sigma_f^* = 2$; and the risk attitude factor $k = 3$.

We generate a total of 8000 training samples using the MCS for developing three GMMs corresponding to f , g_1 , and g_2 . As a result, the obtained GMMs for f , g_1 , and g_2 are mixtures of 45, 30, and 31 Gaussians, respectively. The corresponding computational times required for training three GMMs, with an Intel i7-7700HQ 2.80 GHz CPU and 8.0 GB memory, are 142, 115, and 132 sec, respectively. Eqs. (15) and (16) are used to estimate the mean and variance values of the objective and constraint functions in (27). The RDO solution is found by using a pattern search. In Table 2, the obtained RDO solution is reported and compared with that of previous studies by Lee *et al.* (2009) and Chatterjee *et al.* (2019), where a tensor product Gaussian quadrature (TPQ) and the MCS are used for estimation of statistical moments of the objective and constraint functions, respectively. All RDO

solutions are confirmed by using the MCS with 10^5 samples. It is observed that the GMM can quickly provide a reasonable RDO solution without using complex analyses as the Gaussian quadrature method and enormous computer resources as the MCS.

Table 2 Comparison of robust optimization results for two-bar truss.

Results	Gaussian quadrature (TPQ) (Lee <i>et al.</i> 2009)	MCS 10^5 samples (Chatterjee <i>et al.</i> 2019)	GMM (This study)
μ_{x_1} [cm ²]	11.5655	11.6795	11.5297
μ_{x_2} [m]	0.3771	0.3771	0.3586
F	1.2393 (1.2428)	1.2516	1.2463 (1.2292)
G_1	0 (0.0096)	(0.0012)	0 (0.0116)
G_2	-0.4978 (-0.4931)	(-0.4973)	-0.4922 (-0.5084)
Function calls	3672	$>10^5$	894

Numbers in parentheses are confirmed values using the MCS with 10^5 samples.

As seen from Table 2, the obtained RDO solutions still violate the constraint G_1 . To make the RDO solution more reliable, we replace the constraint on g_1 in problem (28) using a tighter bound as

$$G_1 = k\sigma_{g_1} - \mu_{g_1} - \varepsilon \leq 0 \quad (29)$$

where $\varepsilon = 0.1$.

Now, the solution to problem (28) is $\mu_{x_1} = 11.4146$ cm² and $\mu_{x_2} = 0.4063$ m. The corresponding objective and constraint values are $F = 1.2540$, $G_1 = 0$ and $G_2 = -0.4650$. It shows that the GMM can be effectively used to solve RDO problems with different bounds of constraints, which can be explained by the fact that the GMM equally handles the randomness of the design variables and parameters during its construction, and training should be carried out only once before solving the optimization problem with various bounds of constraints.

3.3 Steel frame

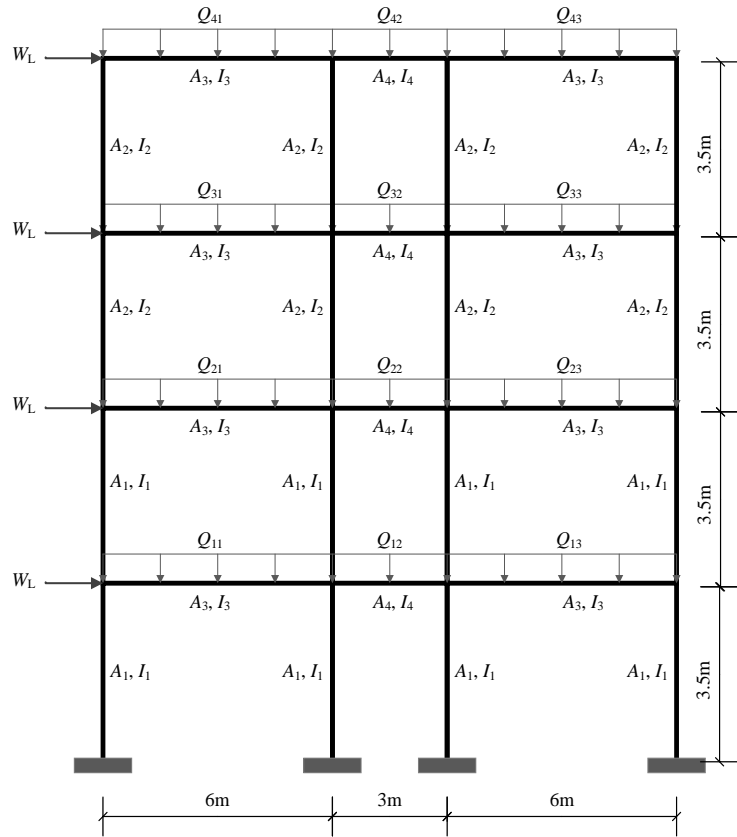


Fig. 3 Four-story three-bay steel frame.

To examine the regression and probabilistic modeling performances of the GMM for structural models, we investigate a four-story three-bay steel frame taken from Zhang and Pandey (2013), as shown in Fig. 3. External loads include dead loads D_L , short-term live load S , long-term live load L , snow load on the roof S_L , and wind load W_L . Vertical loads on each beam of the frame, indicated in Fig. 3, are described in Table 3. The function of interest is the limit state function of the frame, which is defined as

$$g(\mathbf{s}, \mathbf{u}) = \frac{\delta_{\max}(\mathbf{s}, \mathbf{u})}{\delta_a} - 1 \quad (30)$$

where \mathbf{s} is the vector of design variables, i.e., four pairs of the cross-sectional area and moment of inertia $\{(A_1, I_1), (A_2, I_2), (A_3, I_3), (A_4, I_4)\}$ of four groups of members described in Fig. 3; \mathbf{u} is the vector of random parameters, including Young's modulus and external loads $W_L, Q_{11}, Q_{12}, \dots, Q_{43}$, with associated

PDFs specified in Table 4, where COV is the coefficient of variation; δ_{\max} is the maximum inter-story drift; and δ_a is the allowable upper bound for δ_{\max} , which is equal to 11.7 mm.

Table 3 Vertical loads applied to each beam of the steel frame.

Load	Combination
$Q_{11}, Q_{13}, Q_{21}, Q_{23}, Q_{31}, Q_{33}$	$D_L + S_1 + L_1$
Q_{12}, Q_{22}, Q_{32}	$D_L + S_2 + L_2$
Q_{41}, Q_{42}, Q_{43}	$D_L + S_L$

Table 4 Probabilistic characteristics of random parameters of the steel frame.

Variable	Description	Mean	COV	Distribution
D_L	Dead load [kN/m]	20	0.10	Normal
S_1	Short term live load 1[kN/m]	10	0.10	Gumbel
S_2	Short term live load 2 [kN/m]	5	0.30	Gumbel
L_1	Long term live load 1[kN/m]	10	0.30	Gumbel
L_2	Long term live load 2 [kN/m]	5	0.30	Gumbel
S_L	Snow load [kN/m]	5	0.26	Gumbel
W_L	Wind load [kN]	8	0.37	Gumbel
E	Young's Modulus [GPa]	210	0.08	Normal

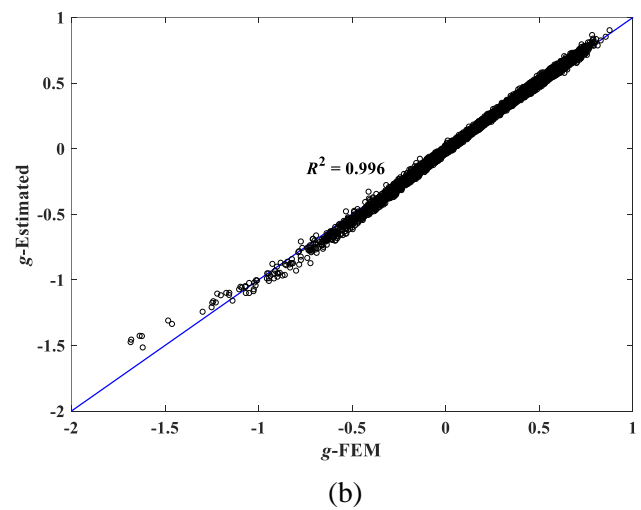
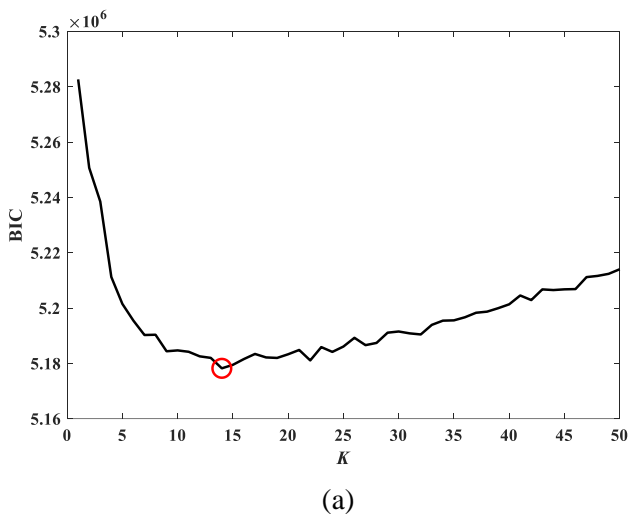
Table 5 Predefined intervals for the design variables in the steel frame.

Design variable	Interval [cm ²]	Design variable	Interval [cm ⁴]
A_1	[46, 122]	I_1	[2118, 12949]
A_2	[29, 90]	I_2	[735, 6551]
A_3	[39, 97]	I_3	[4922, 26599]
A_4	[24, 62]	I_4	[1652, 9610]

To generate a training dataset, we uniformly sample the design variables \mathbf{s} , over eight predefined intervals, as provided in Table 5. We also generate the random parameters \mathbf{u} based on their associated PDFs given in Table 4 using the MCS. The training dataset of 3×10^4 samples of the design variables, the random parameters, and the limit state function value is created. Then, the GMM parameters are obtained using the EM algorithm on the training set.

To generate a test dataset, the mean value for each design variable is randomly picked within the corresponding interval as provided in Table 5. Then, we assume that the design variables \mathbf{s} are normally distributed around the picked means with a COV of 0.05 for both the cross-sectional area and the moment of inertia. Meanwhile, the random parameters \mathbf{u} are generated in the same way as creating the training set. A total of 2×10^4 samples of the test dataset are generated and used for obtaining a comparison between the performances of the GMM and MCS using finite element analyses.

Fig. 4 (a) shows the process of automatic selection of the number of Gaussians K using BIC. The optimal number of Gaussians found in this experiment is 14 when BIC has the minimum value of 5.1782×10^6 . Fig. 4 (b) shows a scatter of $g(\mathbf{s}, \mathbf{u})$ values computed by the use of Eq. (14) (estimated) over those from finite element analyses (true). A high coefficient of determination $R^2 = 0.996$ is observed, which is a reliable indicator for a good regression performance of the GMM. The probabilistic property of the limit state function $g(\mathbf{s}, \mathbf{u})$ under uncertainty is shown in Figs. 4(c) and (d), where the true versus estimated PDFs and cumulative distribution functions (CDFs) are compared, respectively. The mean values for the true and estimated $g(\mathbf{s}, \mathbf{u})$ are 0.2342 and 0.2335, respectively, and the corresponding variance values are 0.0822 and 0.0901. It is observed that the GMM has an ability to accurately extract the limit state function and characterize its probabilistic property for a frame structure in the presence of uncertainty in design variables and random parameters.



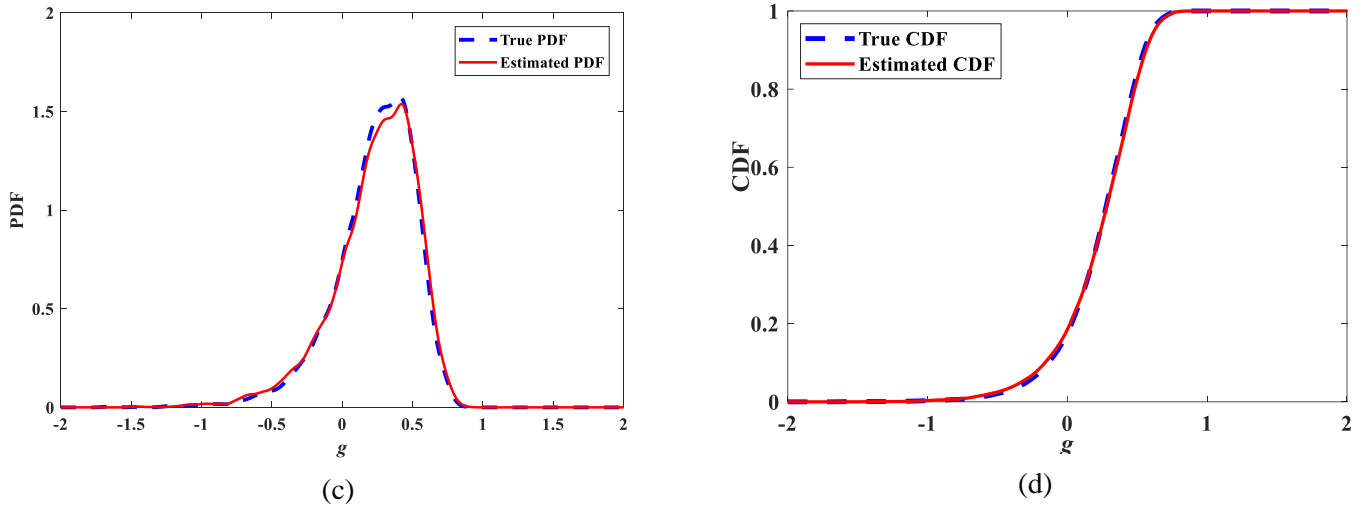


Fig. 4 Results of GMM performances test for four-story three-bay steel frame: (a) BIC vs K ;
(b) Scatter plot; (c) PDF; and (d) CDF.

4 Robust design optimization of steel frames

4.1 Formulation of robust design optimization problem

We extend the deterministic optimum design problem of steel frames subject to the constraints on design strength, serviceability, and constructional requirements specified by the load and resistance design approach (Pezeshk *et al.* 2000; Doğan and Saka 2012; Toğan 2012; Maheri and Narimani 2014) to incorporate uncertainty and state our multi-objective RDO problem. Since the inter-story drift is used to measure the expected damage level of building frames, we consider it as an additional objective function in our problem.

The design variable vector, denoted by \mathbf{s}^0 , represents the set of pairs of nominal value of cross-sectional area and moment of inertia of members. Because the sections are selected from the list of standard steel sections, the components of \mathbf{s}^0 are integer and deterministic, and the values of cross-sectional area and moment of inertia as well as the elastic and plastic sectional moduli are linked. The random vector \mathbf{s} is generated in the neighborhood of \mathbf{s}^0 based on the specified PDF of geometric parameters such as the overall depth, flange width, web thickness, and flange thickness; see Section

4.2. The vector of random parameters, denoted by \mathbf{u} , represents Young's modulus, the yield and tensile strengths, and the external loads.

The objective functions f_1 , f_2 , and f_3 are the total structural mass, and the mean and variance values of the limit state function related to the maximum inter-story drift, respectively. Let n_e , L_i , W_i , and δ_j and δ_{j-1} denote the number of members including beams and columns, the length of the i th member, the nominal mass (kg/m) of the section for the i th member, and the lateral drifts of two adjacent story levels j and $j-1$, respectively. The allowable value of the inter-story drift is assigned as $h_j / 300$ for the story right below the j th story level with the story height h_j . Hence, the objective functions f_1 , f_2 , and f_3 for a frame with m story levels are defined as

$$f_1(\mathbf{s}^0) = \sum_{i=1}^{n_e} W_i (s_i^0) L_i \quad (31)$$

$$f_2(\mathbf{s}^0) = E[f_2^*(\mathbf{s}, \mathbf{u})] - 1, \quad f_2^*(\mathbf{s}, \mathbf{u}) = \max\left(\frac{\delta_2 - \delta_1}{h_2 / 300}, \dots, \frac{\delta_m - \delta_{m-1}}{h_m / 300}\right) \quad (32)$$

$$f_3(\mathbf{s}^0) = \text{var}[f_2^*(\mathbf{s}, \mathbf{u}) - 1] \quad (33)$$

The constraint on the mean performance of the i th beam-column member is assigned based on the AISC-LRFD interaction formula (AISC 2016b) as

$$g_{1,i}(\mathbf{s}^0) = E[g_{1,i}^*(\mathbf{s}, \mathbf{u})] \leq 0, \quad g_{1,i}^*(\mathbf{s}, \mathbf{u}) = \begin{cases} \frac{P_{r,i}}{P_{c,i}} + \frac{8}{9} \left(\frac{M_{rx,i}}{M_{cx,i}} + \frac{M_{ry,i}}{M_{cy,i}} \right) - 1, & \text{if } \frac{P_{r,i}}{P_{c,i}} \geq 0.2 \\ \frac{P_{r,i}}{2P_{c,i}} + \left(\frac{M_{rx,i}}{M_{cx,i}} + \frac{M_{ry,i}}{M_{cy,i}} \right) - 1, & \text{if } \frac{P_{r,i}}{P_{c,i}} < 0.2 \end{cases} \quad (34)$$

where P_r and P_c are the required and available axial strengths of the structural members (either tension or compression), respectively; M_{rx} and M_{ry} are the required flexural strengths about the major axis, x , and the minor axis, y , respectively; and M_{cx} and M_{cy} are the available flexural strengths about the major axis, x , and the minor axis, y , respectively.

The available axial and flexural strengths in (34) are estimated in accordance with chapters E and F of ANSI/AISC 360-16 (AISC 2016b), respectively. To determine the effective length factors for the

estimation of the available strengths for columns of an unbraced frame, the French Rule (Dumonteil 1992) is utilized as

$$K_c = \sqrt{\frac{1.6G_A G_B + 4(G_A + G_B) + 7.5}{G_A + G_B}} \quad (35)$$

where

$$G = \frac{\sum I_c / L_c}{\sum I_b / L_b} \quad (36)$$

with the subscripts A and B refer to the joints at lower (A) and upper (B) ends of the column being considered; and $I_{c(b)}$ and $L_{c(b)}$ are the moment of inertia and length of the column (beam) members connected to that joint, respectively. For a column in the 1st story, $G_A = 10$ if the column is pin-supported, and $G_A = 1$ if it is fixed to a strong foundation.

The constraint on the mean value of the total drift of the structure Δ_t is given as

$$g_2(\mathbf{s}^0) = E[g_2^*(\mathbf{s}, \mathbf{u})] \leq 0, \quad g_2^*(\mathbf{s}, \mathbf{u}) = \frac{|\Delta_t|}{H/400} - 1 \quad (37)$$

where $H/400$ is the allowable value of Δ_t of the frame with the overall height H .

The constraint on the mean value of the long-term deflection of the k th beam members $\Delta_{b,k}$ is assigned as

$$g_{3,k}(\mathbf{s}^0) = E[g_{3,k}^*(\mathbf{s}, \mathbf{u})] \leq 0, \quad g_{3,k}^*(\mathbf{s}, \mathbf{u}) = \frac{|\Delta_{b,k}|}{L_{b,k}/360} - 1 \quad (38)$$

where $L_{b,k}$ and $L_{b,k}/360$ are the span and the allowable long-term deflection of the k th beam, respectively.

The strong column-weak beam approach ensures that the columns are strong enough so that the frame exhibits an yielding mechanism over multiple stories of the frame, rather than an inelastic action concentrated in column hinges at a single story (AISC 2016a). Thus, the strong column-weak beam requirements defined for planar steel frames are stated as

$$g_{4,l}(\mathbf{s}^0) = \mathbb{E}[g_{4,l}^*(\mathbf{s}, \mathbf{u})] \leq 0, \quad g_{4,l}^*(\mathbf{s}, \mathbf{u}) = 1 - \frac{\sum M_{pc,l}}{\sum M_{pb,l}} \quad (39)$$

where $\sum M_{pc,l}$ is sum of the flexural strengths of the columns that constitutes the l th joint, i.e., above and below columns; and $\sum M_{pb,l}$ is sum of the flexural strengths of the beams connected to the l th joint. The readers may refer to ANSI/AISC 341-16 (AISC 2016a) for detailed calculations of these strengths.

The geometric constraints, which ensure the flange width of beam members b_{bf} , connected to a column should not exceed than the flange width of the column b_{cf} , are given as

$$g_{5,m}(\mathbf{s}^0) = \frac{b_{bf,m}}{b_{cf,m}} - 1 \leq 0 \quad (40)$$

The geometric constraints, which ensure the depth of column section on the upper story $d_{c,n}$ of column-column joint should not exceed the depth of the column in the lower story $d_{c,n-1}$, are defined as

$$g_{6,n}(\mathbf{s}^0) = \frac{d_{c,n}}{d_{c,n-1}} - 1 \leq 0 \quad (41)$$

Let n_b , n_j , n_{bc} , and n_{cc} denote the number of beams, number of joints between beams and columns, number of beam-column connections, and number of column-column joints, respectively. Then, the multi-objective RDO problem with three objective functions is formulated as follows:

$$\min. f_1(\mathbf{s}^0), f_2(\mathbf{s}^0), f_3(\mathbf{s}^0)$$

subject to

$$g_{1,i}(\mathbf{s}^0) \leq 0, \quad i = 1, 2, \dots, n_e$$

$$g_2(\mathbf{s}^0) \leq 0$$

$$g_{3,k}(\mathbf{s}^0) \leq 0, \quad k = 1, 2, \dots, n_b$$

$$g_{4,l}(\mathbf{s}^0) \leq 0, \quad l = 1, 2, \dots, n_j$$

$$g_{5,m}(\mathbf{s}^0) \leq 0, \quad m = 1, 2, \dots, n_{bc}$$

$$g_{6,n}(\mathbf{s}^0) \leq 0, \quad n = 1, 2, \dots, n_{cc}$$

(42)

In our design optimization problem, column and beam members of a frame are classified into b_1 and b_2 groups, respectively. The design variables \mathbf{s}^0 represent the nominal values of the cross-sectional area and moment of inertia that are linked. Without compromising the structural integrity, we only estimate the GMMs corresponding to f_2^* value, g_1^* values of b_1 column groups (each is the maximum value among the values for columns in the same group), g_1^* values of b_2 beam groups, g_2^* value, and g_3^* values of b_2 beam groups. Since the values of g_4 , g_5 , and g_6 can be easily calculated from the list of standard steel sections, they are directly assigned as inequality constraints to the optimization problem. Nominal values are simply used for computing the total structural mass $f_1(\mathbf{s}^0)$, because the uncertainty of cross-sectional dimension is normally distributed around the nominal value.

4.2 Uncertainty characteristic

We limit the uncertainty in our optimization problems to aleatory uncertainty, which is associated with the randomness of the random vectors \mathbf{s} and \mathbf{u} incorporated in problem (42). Generally, these vectors are evaluated by data of either onsite or laboratory experiments that contain a degree of variation due to unavoidably inaccurate measurements. After collecting the data, statistics tools can be applied to

generate a mathematical model to characterize the uncertainty. Thus, aleatory uncertainty is irreducible and described by either a CDF or a PDF. In this study, we refer to some previous surveys and ASCE (2017) to describe the aleatory uncertainty in \mathbf{s} and \mathbf{u} .

The types of probability of the geometric properties related to the design variable vector \mathbf{s} , including the overall depth, flange width, web thickness, and flange thickness of standard steel sections, are provided in Table 6. Those of the random parameters \mathbf{u} , i.e., material properties and external loads are given in Table 7.

Table 6 Probabilistic characteristic of geometric properties.

Variable	Nominal value	Statistical property		
		Mean/ Nominal	COV	Distribution
Overall depth ¹	d	1.00	0.03	Normal
Flange width ¹	b_f	1.00	0.02	Normal
Web thickness ²	t_w	1.00	0.05	Normal
Flange thickness ²	t_f	1.00	0.05	Normal

¹ Hess *et al.* (2002); ² assumed

Table 7 Probabilistic characteristics of material and load random parameters.

Variable	Nominal value	Statistical property		
		Mean/ Nominal	COV	Distribution
Young's Modulus ¹	E	1.00	0.04	Normal
Yield stress ¹	F_y	1.10	0.06	Normal
Tensile strength ²	F_u	1.07	0.08	Normal
Dead load ³	D_L	1.05	0.10	Normal
Live load ³	L	1.00	0.10	Gumbel
Wind load ³	W_L	0.92 (Site-dependent)	0.37 (Site-dependent)	Gumbel

¹ Bartlett *et al.* (2003); ² Hess *et al.* (2002); ³ ASCE (2017)

4.3 Training data preparation

As a machine-learning approach, the GMM requires a training dataset of various random vectors \mathbf{s} , random parameters \mathbf{u} , and the corresponding output responses of a particular planar steel frame. The training dataset is collected using linear elastic frame analyses.

As inputs, the training samples of the design variables including geometric properties for the column and beam groups are uniformly generated over associated intervals. The MCS is then performed to generate data around those particular samples to characterize the probabilistic property of the sectional geometric dimensions, as given in Table 6. In other words, the used sampling technique is a simple random sampling scheme of two stages: generating uniformly a random set of samples and generating data around those samples. Alternatively, since the variances of the sectional geometric dimensions are very small, as shown in Table 6, only the first stage may be used. Meanwhile, the MCS is used to generate the random samples of Young's modulus, yield strength, tensile strength, and external loads, based on their basis probabilistic distributions listed in Table 7.

To determine a proper value of the number of training samples for a particular problem, a total of five training sets of design variables and random parameters with different numbers of samples, i.e., 10^4 , 2×10^4 , 3×10^4 , 4×10^4 , and 5×10^4 , are used for training the GMM. Since problem (42) is formulated in a high dimensional space of \mathbf{s} and \mathbf{u} , a large number of training samples is required to effectively generalize the design space. After the GMMs are obtained, a total of 2×10^4 samples of an independent test dataset are generated and used to compute the coefficient of determination R^2 , which evaluates the goodness-of-fit of each GMM and the MCS using finite element analyses. It is observed that the obtained GMMs associated with 5×10^4 samples provide the highest R^2 ($R^2 > 0.90$ for all design examples), it is used for solving problem (42). In the case when the obtained GMM performance does not meet the desired accuracy, more training data points should be added.

4.4 Best compromise solution

The objective functions f_2 and f_3 of problem (42) are highly conflicting so that it is desirable to find a compromise solution (Marler and Arora 2004). The best compromise solution has been efficiently

selected from Pareto-optimal solutions by using a fuzzy-based method (Abido 2003). In this regard, a membership function for the i th objective function of each Pareto-optimal solution is defined as (Abido 2003)

$$\mu_{i,j} = \begin{cases} 1 & f_{i,j} \leq f_i^{\min} \\ \frac{f_i^{\max} - f_{i,j}}{f_i^{\max} - f_i^{\min}} & f_i^{\min} < f_{i,j} < f_i^{\max}, \quad i = 1, 2, 3 \\ 0 & f_{i,j} \geq f_i^{\max} \end{cases} \quad (43)$$

where the subscript j indicates the value of the j th Pareto-optimal solution, and f_i^{\max} and f_i^{\min} are the maximum and minimum values of the i th objective function, respectively. A normalized membership function for each Pareto-optimal solution is computed as (Abido 2003)

$$\mu_j = \frac{\sum_{i=1}^3 \mu_{i,j}}{\sum_{j=1}^M \sum_{i=1}^3 \mu_{i,j}} \quad (44)$$

where M is the number of Pareto-optimal solutions. The best compromise solution is the one with the highest value of the normalized membership function, μ_j .

4.5 Robust design optimization procedure

The proposed RDO is implemented using the procedure in Fig. 5 with the following steps:

Step 1: The random variable vector \mathbf{s} and random parameter vector \mathbf{u} for a particular steel frame are defined and characterized; see Section 4.2. The training data is generated for the GMM process; see Section 4.3. A total of 5×10^4 training data points are generated.

Step 2: The GMM starts at the number of Gaussian $K = 1$, which increases step by step to $K = 50$, thereby producing 50 GMMs. The best GMM is chosen based on the minimum BIC in Eq. (25). Independent samples of a test dataset are generated and used for obtaining the discrepancy (or R^2) between the performances of the obtained GMM and MCS using finite element analyses. The accuracy of the obtained GMM performance is verified by a desired value of R^2 to decide whether more training data points should be added.

Step 3: Pareto-optimal solutions of problem (42) are found by employing the nondominated-sorting genetic algorithm II (NSGA-II) utilizing the trained GMM on the calculations of the mean values f_2 , g_1 , g_2 , and g_3 , and the variance f_3 .

Step 4: The best compromise solution among Pareto-optimal solutions is obtained based on the fuzzy set theory using Eqs. (43) and (44); see Section 4.4.

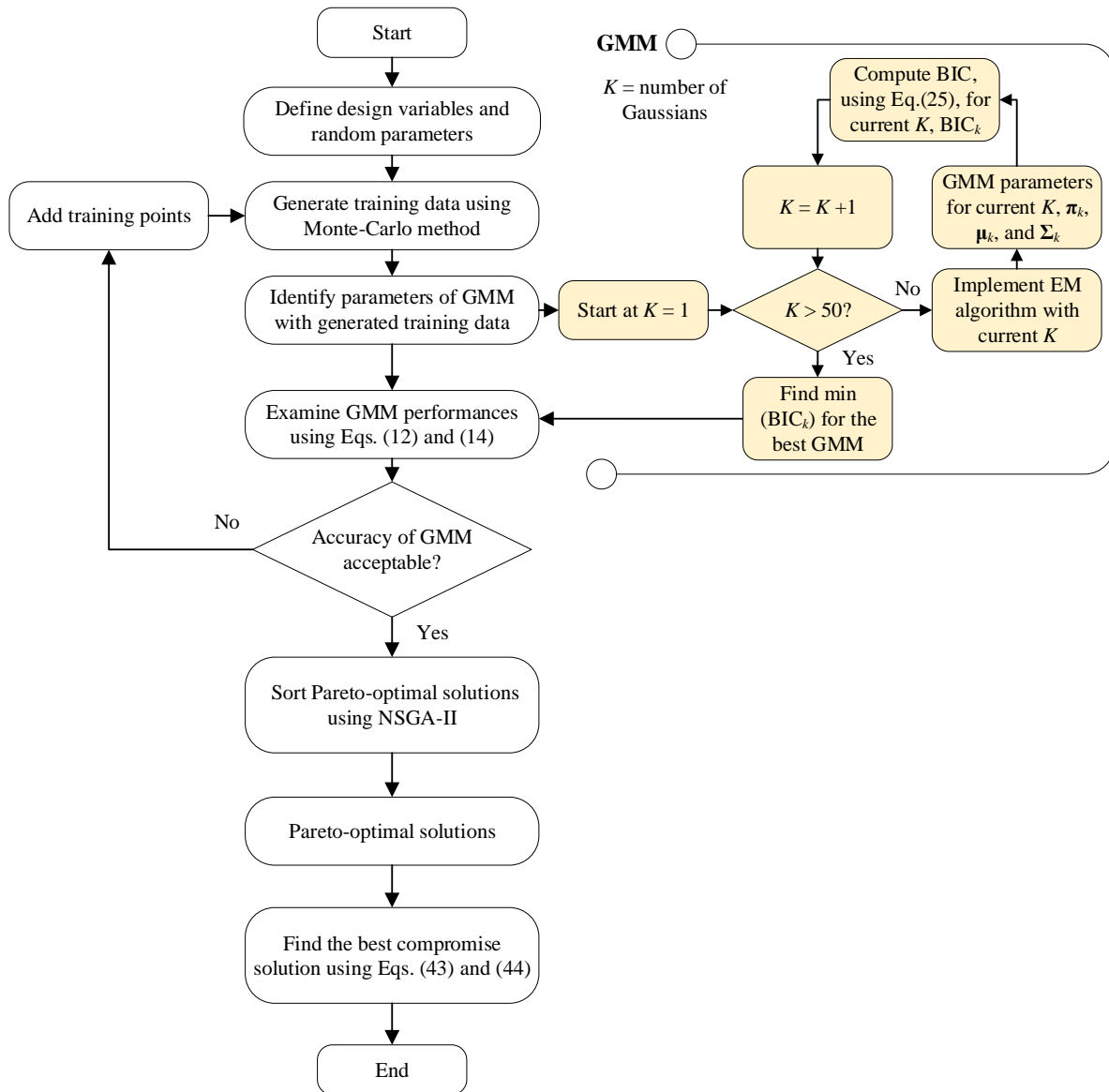


Fig. 5 Proposed RDO procedure.

Since design of steel frame structures is to select appropriate sections for structural members from a list of commercially available sections, the RDO problem in this study is a combinatorial optimization

problem. Therefore, an elitist-strategy NSGA-II (Deb *et al.* 2002), available in the MATLAB R2018a Global Optimization Toolbox, is utilized for obtaining the Pareto-optimal solutions in Step 3. The effectiveness of a genetic algorithm-based algorithm in solving multi-objective combinatorial optimization problems has been verified (Konak *et al.* 2006). Some control parameters characterizing the use of NSGA-II are listed in Table 8. To examine its stochastic property, NSGA-II is performed several times for each design example with an Intel Xeon E5-2643V4 3.40 GHz CPU and 63.9 GB memory. Different seeds of the random number generator are used to generate different initial populations for NSGA-II. Note that, to employ NSGA-II for the discrete optimization problem, an array of integers is randomly generated at the beginning of the algorithm for the initial population. We also define a mapping function to transform the integer decision vector \mathbf{s}^0 to the linked random vector \mathbf{s} , i.e., the cross-sectional dimensions, cross-sectional area, moment of inertia, elastic and plastic sectional moduli. Additionally, the crossover and mutation functions should be modified to guarantee that offsprings of the NSGA-II produced after each generation are integers.

Table 8 Parameters for NSGA-II.

Parameter	Value
Population size	1,000
Probability of mutation	20%
Maximum number of generations	100
Fitness function tolerance	10^{-6}
Constraint tolerance	10^{-6}
Parallel computing fitness and nonlinear constraint functions	true

5 Design examples

5.1 Three-story, two-bay steel frame

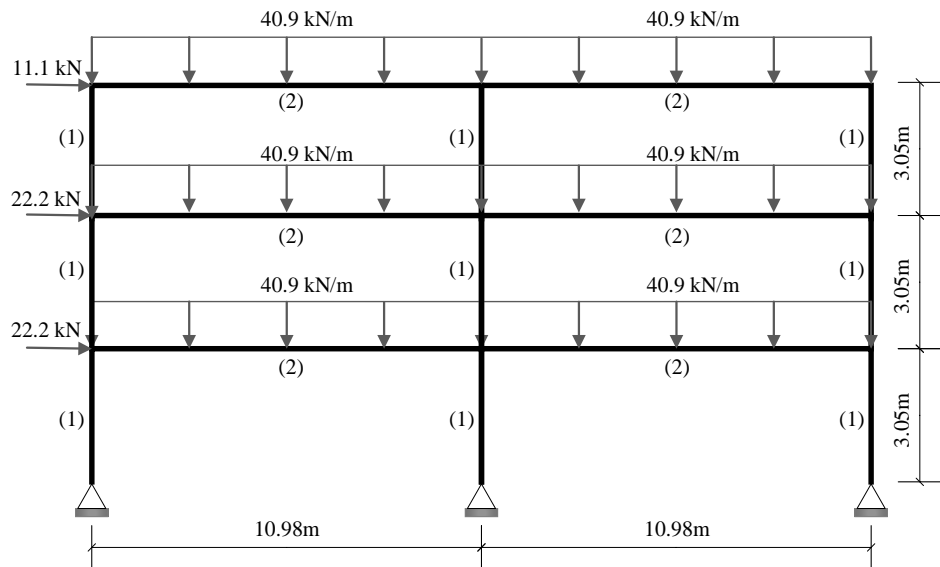


Fig. 6 Three-story two-bay steel frame.

Problem (42) is formulated for a benchmark problem of a three-story two-bay steel frame, as shown in Fig. 6. Although many deterministic optimal designs have been found for this frame on a minimum total mass using various optimization algorithms (Pezeshk *et al.* 2000; Toğan 2012; Maheri and Narimani 2014), there does not exist literature on the RDO problem considering uncertainty in material, geometry, and load parameters simultaneously.

The frame consists of 15 members classified into two groups; namely, group (1) for columns and group (2) for beams, as shown in Fig. 6. The cross-sectional dimensions, external loads, and material properties, i.e., yield strength, tensile strength, and Young's modulus, are sources of uncertainties. Nominal value of cross-sectional dimension of the members in each group is extracted from a list of American wide flange standard steel sections, as listed in Table 9. Nominal values of the yield stress, tensile strength, and Young's modulus of the steel material are 250 MPa, 400 MPa, and 200 GPa, respectively. The mean values of external loads are shown in Fig. 6. The probabilistic properties of these values are given in Tables 6 and 7.

Table 9 List of columns and beams for three-story two-bay steel frame.

Column		Beam	
W18×86	W12×72	W24×94	W18×65
W18×76	W12×65	W24×84	W18×60
W18×65	W12×58	W24×76	W16×77
W18×60	W12×53	W24×68	W16×67
W18×55	W12×50	W24×62	W16×57
W16×89	W12×45	W24×55	W16×50
W16×77	W12×40	W21×93	W16×45
W16×67	W10×77	W21×83	W16×40
W16×57	W10×68	W21×73	W14×61
W16×50	W10×60	W21×68	W14×53
W14×90	W10×54	W21×62	W14×48
W14×82	W10×49	W21×57	W14×38
W14×74	W10×45	W21×55	W14×34
W14×68	W8×40	W21×50	W14×30
W14×61	W8×35	W21×48	W14×26
W14×53	W8×31	W18×71	W14×22

To create the training data for this example, the intervals of the cross-sectional dimensions for the column group are d : [150, 500] mm, b_f : [140, 400] mm, t_w : [5.5, 16.0] mm, and t_f : [6, 24] mm, and those for the beam group are d : [350, 620] mm, b_f : [160, 250] mm, t_w : [5.5, 16.0] mm, and t_f : [9, 24] mm. A total of five GMMs corresponding to the responses f_2 (GMM-1), $g_{1,1}$ for the column group (GMM-2), $g_{1,2}$ for the beam group (GMM-3), g_2 (GMM-4), and g_3 (GMM-5) are constructed. Fig. 7 shows the process of selection of the number of Gaussians using BIC for five GMMs of the frame. The optimal number of Gaussians for GMM-1, GMM-2, GMM-3, GMM-4, and GMM-5 are 33, 28, 20, 26, and 27, respectively. The corresponding variation ranges of the mixing proportions for the obtained GMMs are [0.0077, 0.0599], [0.0098, 0.0954], [0.0266, 0.1083], [0.0074, 0.0963], and [0.0103, 0.0596], respectively. After the GMMs are obtained, we utilize Eq. (15) to compute the mean values of the responses f_2 , g_1 , g_2 , and g_3 when the steel section is fixed, in which the mean values of the design variables and random parameters are directly determined from the corresponding nominal values. The variance f_3 is computed using the Taylor series expansion in Eq. (16).

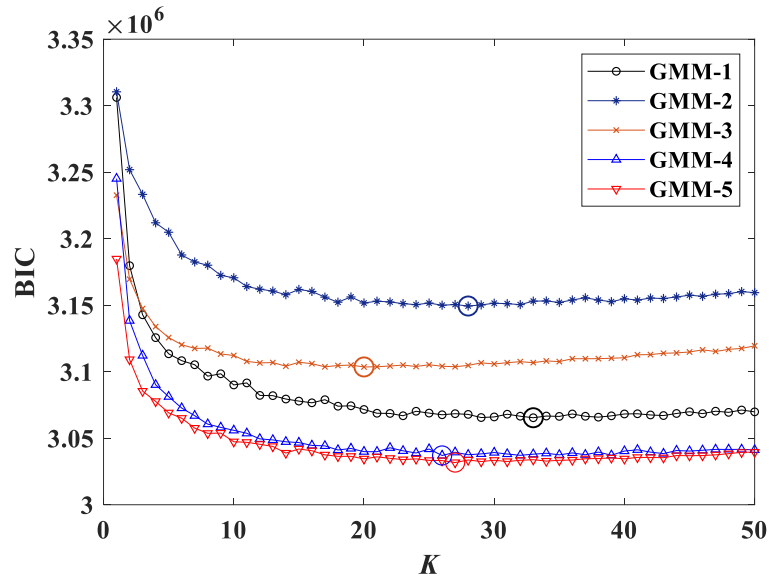


Fig. 7 Selection process of five GMMs for three-story two-bay steel frame.

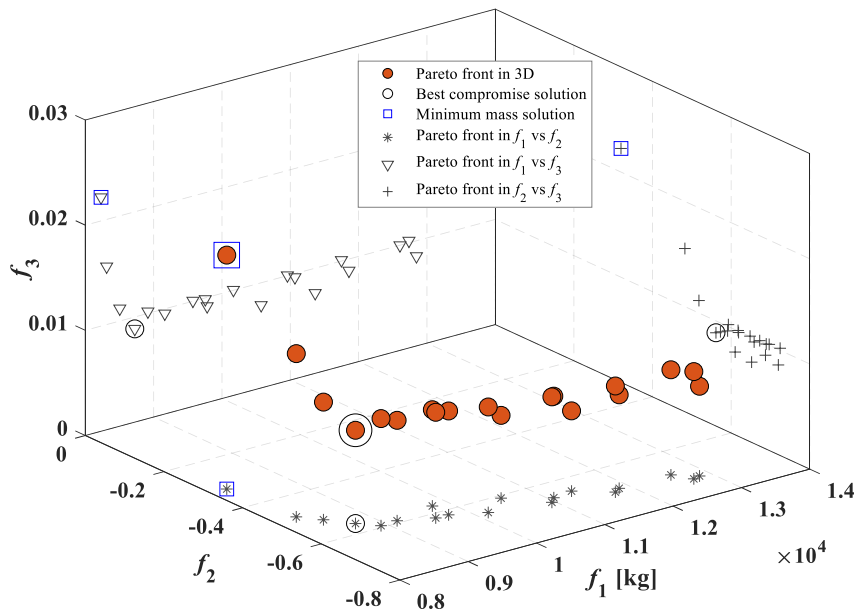


Fig. 8 Pareto front and the best compromise design for three-story two-bay steel frame.

NSGA-II is performed three times and a set of 19 Pareto-optimal solutions, the best compromise solution, and a solution with the minimum value of the structural mass are identically obtained, as shown in Fig. 8. Table 10 provides the selection of steel sections for the column and beam groups. This selection is compared with the previous deterministic optimal designs. It is observed that the minimum total mass of the frame is 8229 kg that outperforms the design by Pezeshk *et al.* (2000) and is 2% and 0.7% larger than the designs by Toğan (2012) and Maheri and Narimani (2014), respectively. The best

compromise design has the largest total mass of 8724 kg. However, it clearly shows a trade-off between the three objective functions thereby producing a very small variance $f_3 = 0.010$ in the ratio of inter-story drift to the allowable value, and reasonable structure design performance with $f_2 = -0.561$, far enough from the limit state of inter-story drift defined by $f_2 = 0$.

Table 10 Comparison of optimization results for three-story two-bay steel frame.

Group ID.	Pezeshk <i>et al.</i> (2000)	Toğan (2012)	Maheri and Narimani (2014)	NSGA-II	
				Min. mass	Robust design
Column (1)	W10×60	W10×49	W10×68	W14×53	W18×65
Beam (2)	W24×62	W24×62	W21×55	W24×62	W24×62
f_1 [kg]	8504	8069	8174	8229	8724
f_2	-	-	-	-0.319	-0.561
f_3	-	-	-	0.023	0.010
f_2 (MCS)	-	-	-	-0.317	-0.560
f_3 (MCS)	-	-	-	0.028	0.012
$g_{1,1}$ (MCS)	-	-	-	-0.128	-0.303
$g_{1,2}$ (MCS)	-	-	-	-0.108	-0.149
g_2 (MCS)	-	-	-	-0.617	-0.739
g_3 (MCS)	-	-	-	-0.317	-0.395

To confirm their actual robustness and feasibility, the obtained robust optimal designs are used to characterize the probabilistic property of the limit state function regarding the maximum inter-story drift in Eq. (32) under uncertainty. The mean (f_2) and variance (f_3) values of the maximum inter-story drift for the minimum mass and the best compromise designs are computed by using the MCS with 10^5 samples. The limit state functions of the constraints in Eqs. (34), (37), and (38) are also estimated and reported in Table 10. As a result, the obtained robust optimal designs are feasible as all the constraint values are negative. Also, the mean f_2 and variance f_3 are -0.560 and 0.012 for the best compromise design, and -0.317 and 0.028 for the minimum structural mass design, which are very close to those

obtained from the GMM-1. In Fig. 9, the PDFs and CDFs of the maximum inter-story drift limit state for the best compromise design generated by the MCS and the GMM-1 are compared. It shows that the PDF and CDF generated by the GMM-1 are in good agreement with those generated by the MCS.

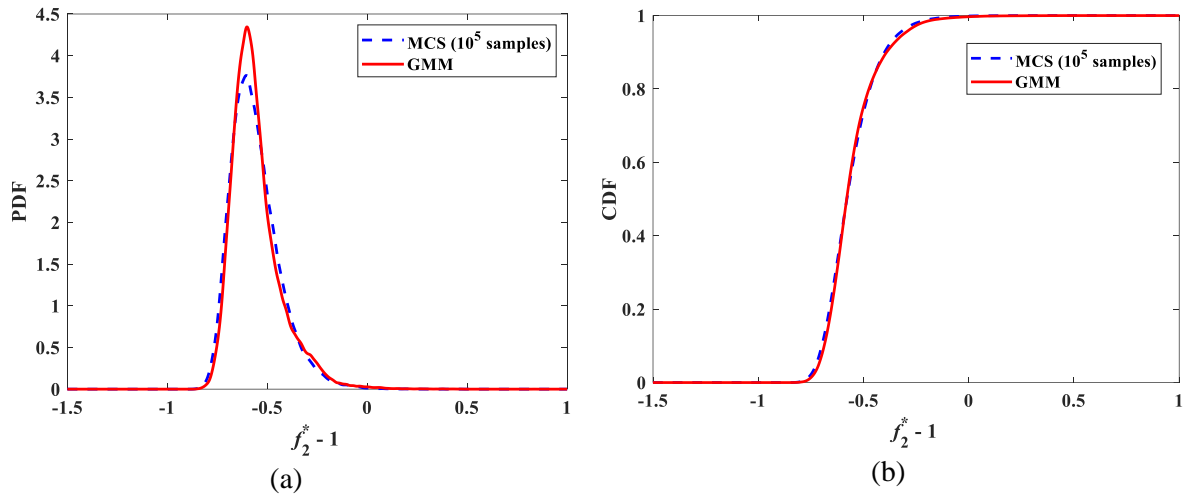


Fig. 9 Statistical property of the maximum inter-story drift for the robust design of three-story two-bay steel frame: (a) PDF and (b) CDF.

5.2 Six-story two-bay steel frame

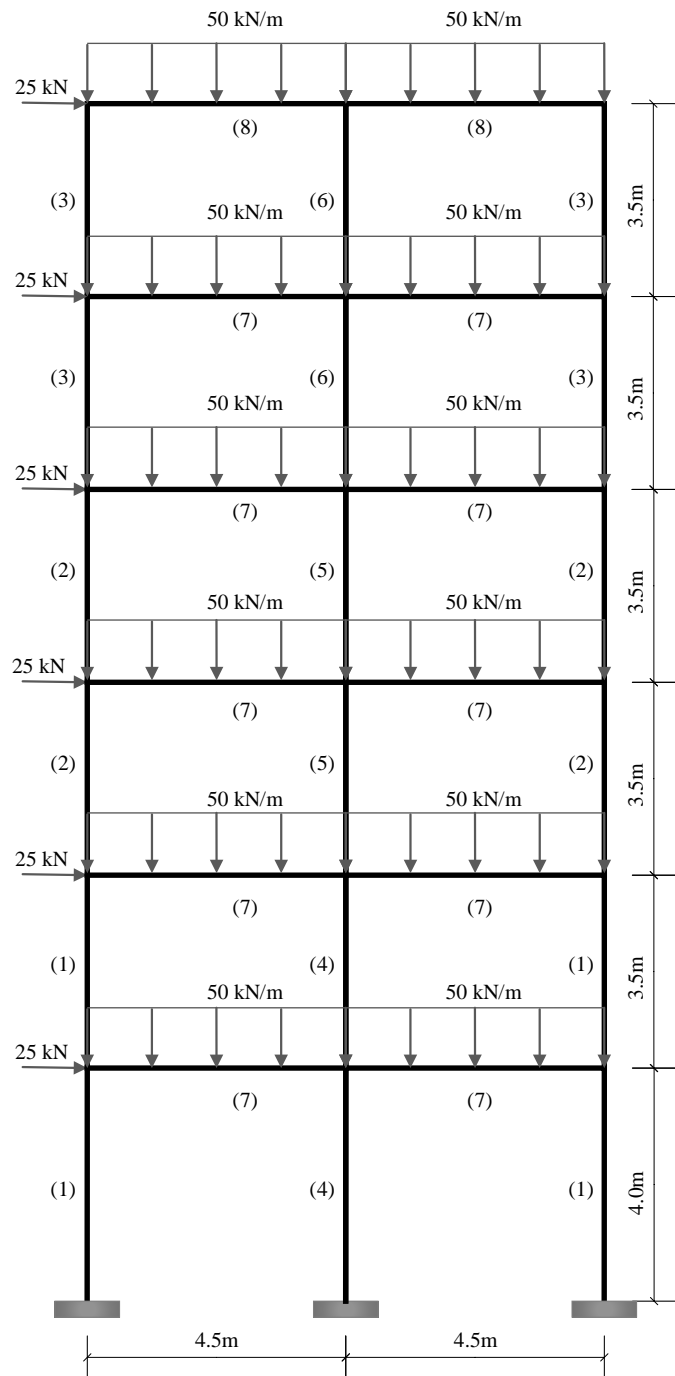


Fig. 10 Six-story two-bay steel frame.

In this design example, problem (42) is formulated for a six-story two-bay steel frame taken from Doğan and Saka (2012), as shown in Fig. 10. The frame includes 30 members classified into six column groups [(1) to (6)] and two beam groups [(7) and (8)]. The cross-sections of the members in beam and column

groups are listed in Table 11. Uncertainty and its characteristic for all random input variables are described in the same way as the three-story two-bay steel frame in the first example.

Table 11 List of columns and beams for six-story two-bay steel frame.

Column			Beam		
W24×84	W16×50	W12×50	W24×62	W16×50	W12×40
W24×76	W16×40	W12×45	W24×55	W16×45	W12×35
W24×68	W14×99	W12×40	W21×62	W16×40	W12×30
W24×62	W14×90	W10×68	W21×55	W16×36	W12×26
W21×93	W14×82	W10×60	W21×48	W16×31	W12×22
W21×83	W14×74	W10×54	W21×57	W16×26	W12×19
W21×73	W14×68	W10×49	W21×50	W14×43	W12×16
W21×68	W14×61	W10×45	W18×71	W14×38	W12×14
W18×86	W14×53	W10×39	W18×65	W14×34	W10×33
W18×76	W14×48	W8×40	W18×60	W14×30	W10×30
W18×65	W12×79	W8×35	W18×55	W14×26	W10×26
W18×60	W12×72	W8×31	W18×50	W14×22	W10×22
W18×55	W12×65	W6×25	W18×46	W12×53	W10×19
W16×77	W12×58	W6×20	W18×40	W12×50	W10×17
W16×67	W12×53	W6×15	W18×35	W12×45	W10×15
W16×57			W16×57		

Since there are many groups of columns and beams in this design example, NSGA-II is performed 25 times to examine the statistical characteristics of the robust optimal solutions. For each NSGA-II trial, the minimum mass and the best compromise solutions are obtained after 100 generations. Fig. 11 shows the variation in the total structural mass for the minimum mass and the best compromise solutions. The corresponding variation ranges are [7460, 7739] and [8520, 8694] kg, respectively. There exists a variation in the solution of 25 NSGA-II implementations. However, the differences in the minimum mass and in the best compromise solutions among the obtained solutions are not much higher than 3.7% and 2.0%, respectively. Figs. 12 (a), (b), and (c) show a set of 303, 218, and 284 Pareto-optimal solutions obtained from 1st, 2nd, and 3rd trials of NSGA-II, respectively. Although there are differences in the number of solutions and in the solutions themselves, the shapes of the Pareto-front obtained from the three trials tend to be similar. In Table 12, the minimum mass of the structure is 7460 kg (1st NSGA-II) which outperforms the previous deterministic optimal designs obtained using particle

swarm optimization (7533 kg) and harmony search (7829 kg). Meanwhile, the best compromise solution is 8568 kg (3rd NSGA-II) that is 14.9% heavier than that of the minimum mass design. The trade-off relation between the three objective functions is obviously observed. The minimum mass design although has a smaller mass, it is both un-safer, with $f_2 = -0.147$, and more sensitive, with $f_3 = 0.101$, than the best compromise design, with $f_2 = -0.333$ and $f_3 = 0.037$. Note that, due to a large number of design column and beam groups, we use larger number of candidates of standard steel sections for each group, as provided in Table 11, than those in Table 9. This yields a difficulty in finding the best compromise solution as many Pareto-optimal solutions have almost the same normalized membership functions defined in Eq. (44). For practical designs, depending on the function of the structure, practicing designers should use some rules of thumb for both column and beam members to shorten the decision list.

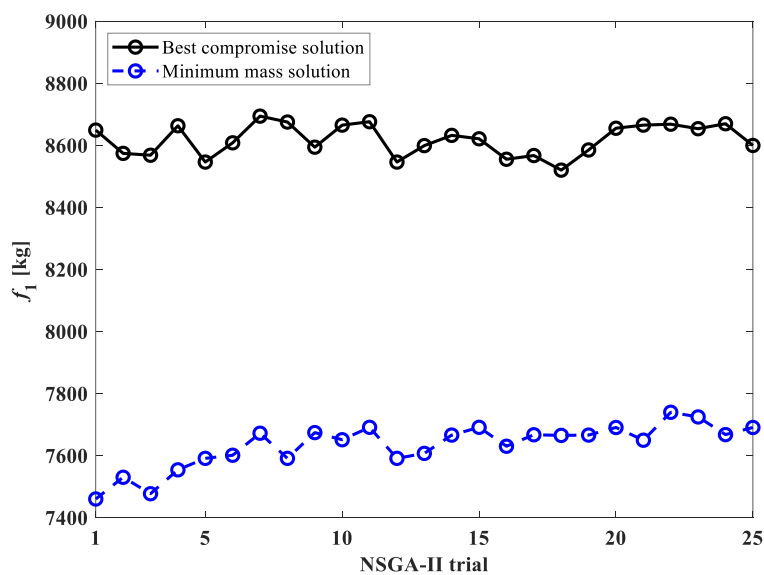
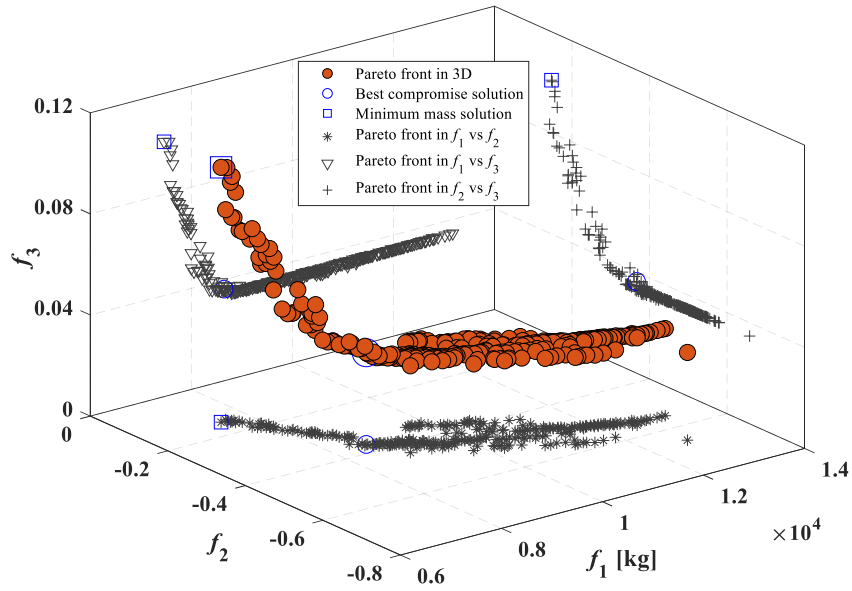


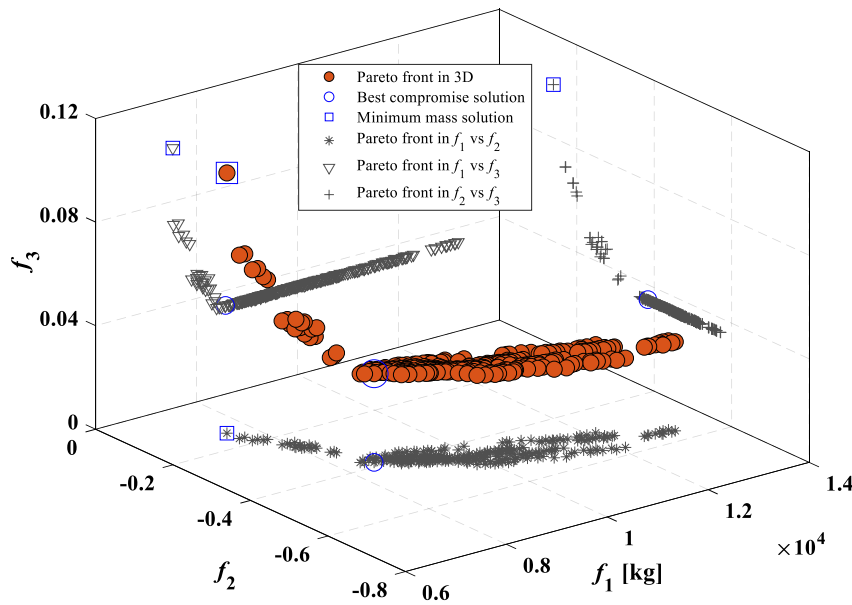
Fig. 11 Variation in the total structural mass due to different NSGA-II trials for six-story two-bay steel frame.

To validate their robustness and feasibility, the minimum mass and the best compromise designs obtained from the first three NSGA-II trials are used to characterize the probabilistic property of the limit state function related to the maximum inter-story drift in Eq. (32) using the MCS with 10^5 samples. The constraint function values in Eqs. (34), (37), and (38) are also computed, as provided in Table 12. It is observed that all constraint function values are negative, which demonstrates the feasibility of the

obtained robust optimal designs. Fig. 13 shows the PDFs and CDFs of the maximum inter-story drift limit state in Eq. (32) for the best compromise designs obtained from the first three NSGA-II trials. An agreement between the PDF and CDF generated by the GMM and those generated by the MCS is observed, although the errors are larger than those for the small examples.



(a)



(b)

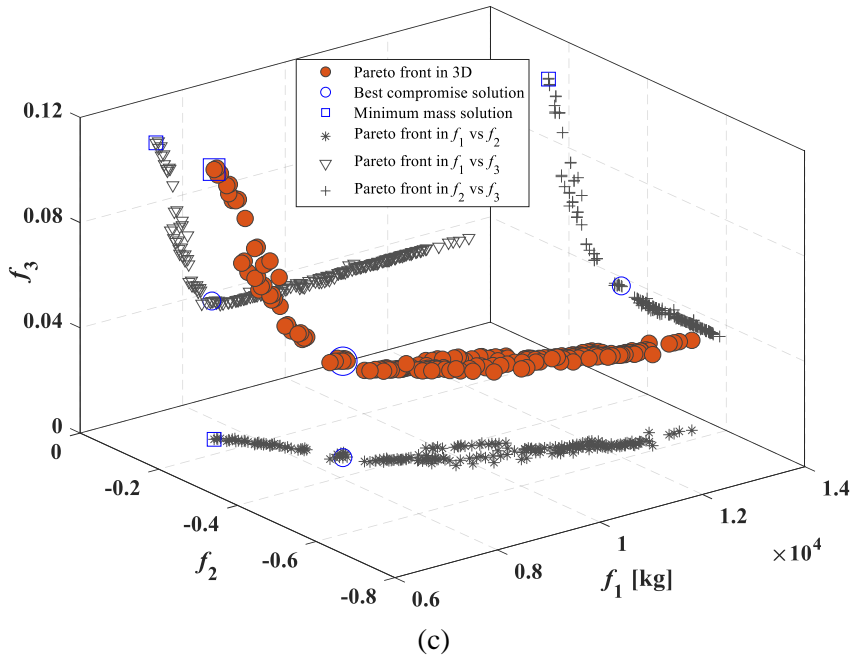


Fig. 12 Pareto front and the best compromise design for six-story two-bay steel frame:

(a) 1st NSGA-II; (b) 2nd NSGA-II; and (c) 3rd NSGA-II.

Table 12 Comparison of optimization results for six-story two-bay steel frame.

Group ID.	Particle swarm Doğan and Saka (2012)	Harmony search Doğan and Saka (2012)	1st NSGA-II		2nd NSGA-II		3rd NSGA-II	
			Min. mass	Robust design	Min. mass	Robust design	Min. mass	Robust design
Column (1)	W16×57	W18×55	W18×60	W18×60	W18×60	W24×62	W18×55	W24×62
Column (2)	W16×40	W12×50	W16×40	W16×40	W14×48	W16×40	W16×40	W14×53
Column (3)	W10×39	W8×31	W12×45	W12×50	W12×40	W12×40	W14×48	W12×50
Column (4)	W24×62	W21×73	W18×55	W18×60	W18×55	W18×55	W18×60	W24×62
Column (5)	W24×62	W18×65	W16×40	W16×40	W16×40	W16×50	W16×40	W18×60
Column (6)	W8×40	W12×40	W14×48	W14×53	W14×48	W14×48	W14×48	W18×55
Beam (7)	W14×30	W16×40	W18×35	W21×50	W18×35	W21×50	W18×35	W18×40
Beam (8)	W18×65	W14×22	W12×35	W12×35	W12×35	W12×40	W12×35	W14×34
f_1 [kg]	7533	7829	7460	8649	7530	8574	7477	8568
f_2	-	-	-0.147	-0.367	-0.140	-0.384	-0.148	-0.333
f_3	-	-	0.101	0.036	0.101	0.034	0.103	0.037
f_2 (MCS)	-	-	-0.172	-0.410	-0.168	-0.391	-0.175	-0.337
f_3 (MCS)	-	-	0.093	0.042	0.095	0.037	0.087	0.041
max. $g_{1,1-6}$ (MCS)	-	-	-0.119	-0.223	-0.118	-0.216	-0.163	-0.197
max. $g_{1,7-8}$ (MCS)	-	-	-0.040	-0.009	-0.040	-0.289	-0.004	-0.079
g_2 (MCS)	-	-	-0.243	-0.457	-0.239	-0.482	-0.245	-0.423
max. $g_{3,1-2}$ (MCS)	-	-	-0.135	-0.177	-0.134	-0.181	-0.162	-0.277

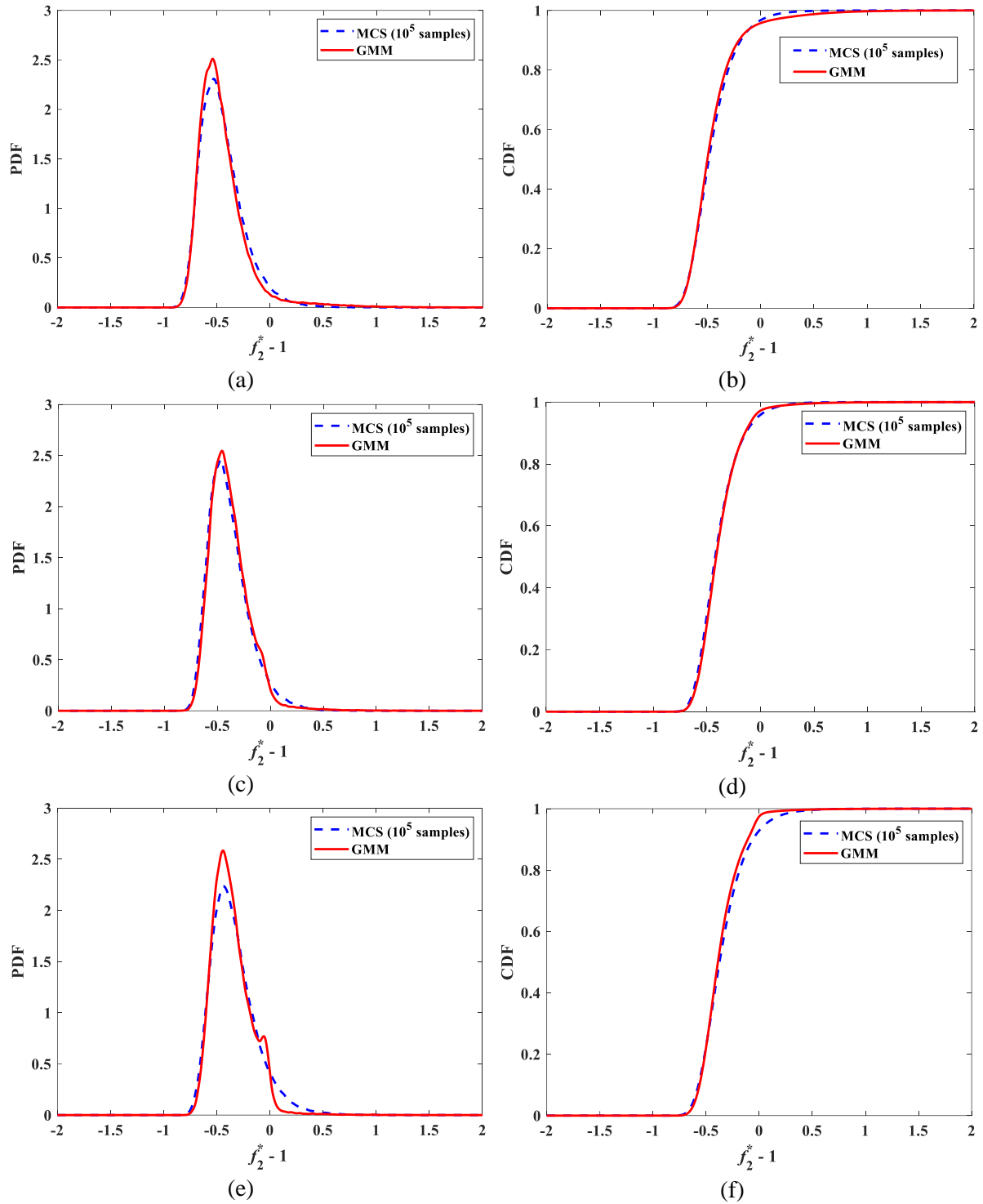


Fig. 13 Statistical property of the maximum inter-story drift for the robust design of six-story two-bay steel frame: (a) PDF-1st NSGA-II; (b) CDF-1st NSGA-II; (c) PDF-2nd NSGA-II; (d) CDF-2nd NSGA-II; (e) PDF-3rd NSGA-II; and (f) CDF-3rd NSGA-II.

6 Conclusions

A new method has been presented for an application of GMM to the RDO of planar steel frames. Uncertainty in the discrete design variables is modeled in the wide range between the smallest and largest values in the catalog of the cross-sectional areas. The main idea of the application is to let random input variables and the corresponding structural response behave according to a joint distribution under the presence of uncertainty. The GMM aims at statistically approximating this distribution based on sampled training data for both characterizing the probabilistic property of the structural response and providing a simple regression function to facilitate the calculation of this response and its statistical parameters required for the RDO problem. The uncertainty in the small ranges of the selected discrete variables can also be easily extracted from the GMM trained only once before carrying out optimization under various conditions. Before application to the RDO problems, the GMM demonstrates its ability in accurately capturing of a given conditional PDF, quickly providing a reasonable RDO solution for a two-bar truss, and accurately characterizing the probabilistic property of a limit state function for a steel frame structure. The multi-objective RDO problem is formulated for the steel frame to minimize three conflicting objective functions, i.e., the total mass, and the mean and variance values of the maximum inter-story drift under some constraints on the design strength and serviceability requirements. Pareto-optimal solutions of the optimization problem are found using a multi-objective genetic algorithm with supports from the trained GMM on the calculation of the objective and constraint functions. The best trade-off solution is also found to relieve the conflicting objective functions from the perspective of fuzzy-based compromise programming.

Two numerical design examples have demonstrated the robustness and feasibility of the proposed method. It is shown that the GMM-based RDO can yield optimum solutions that well represent different levels of the trade-off between the conflicting objective functions. Based on these solutions, structural engineers can specify their design and the best compromise solution is one reasonable choice.

Future researches may focus on the development of the GMM with multi-objective functions so that the performances of all structural members can be fully taken into account instead of using maximum response values for particular member groups. That means, y in Eq. (7) becomes a vector of

structural responses, rather than a scalar. In this study, only the regression function among many powerful abilities of the GMM is utilized. The regression function can also combine with the transformation ability of the GMM (Papadimitriou *et al.* 2018) to accurately estimate failure probabilities for reliability-based optimization problems. In addition, the clustering function of the GMM, i.e., grouping data into similar objects (Hastie *et al.* 2009), may have a potential to be applied for sorting the feasible space and reducing the solution space of the RDO problem.

7 Replication of results

The main steps for applying the proposed method to the RDO problem of planar steel frames are presented in detail in Section 4.5. Training data generated for constructing GMMs in Sections 5.1 and 5.2 are available online at https://bit.ly/gmm_rdosteelframes. Models or source codes used in this study are available from the corresponding author by request for non-commercial purposes only.

Acknowledgments

The authors are thankful to fruitful comments from Prof. Makoto Yamakawa at Tokyo University of Science. Financial support from the Japan International Cooperation Agency (JICA) and JSPS KAKENHI No. JP19H02286 is fully acknowledged.

Compliance with ethical standards

Conflict of interests The authors declare that they have no conflict of interest.

Appendix: Gradient and Hessian of GMMs

Let $\mathbf{x} = [x_1, x_2, \dots, x_D]^T \in \mathbb{R}^D$ be D -dimensional vectors generated from a D -variate Gaussian

$p_{\mathbf{x}}(\mathbf{x}) \sim \mathcal{N}_D(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, such that

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right] \quad (\text{A.1})$$

Let $\Sigma = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ be the singular value decomposition of the covariance matrix Σ , where \mathbf{U} is an orthogonal matrix and $\mathbf{\Lambda}$ is a non-singular diagonal matrix with λ_d at the d th diagonal component. Let $\mathbf{y} = \mathbf{U}^T(\mathbf{x} - \boldsymbol{\mu}) \in \mathbb{R}^D$ so that $\partial y_e / \partial x_d = u_{de}$, where $e \in \{1, 2, \dots, D\}$, and u_{de} is the (d, e) th element of \mathbf{U} . Taking the first derivatives of $p_{\mathbf{x}}(\mathbf{x})$, we have

$$\begin{aligned} \frac{\partial p}{\partial x_d} &= p_{\mathbf{x}}(\mathbf{x}) \frac{\partial}{\partial x_d} \left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right] = p_{\mathbf{x}}(\mathbf{x}) \frac{\partial}{\partial x_d} \left(-\frac{1}{2} \mathbf{y}^T \mathbf{\Lambda}^{-1} \mathbf{y} \right) \\ &= p_{\mathbf{x}}(\mathbf{x}) \sum_{e=1}^D \frac{\partial}{\partial y_e} \left(-\frac{1}{2} \sum_{d=1}^D \frac{y_d^2}{\lambda_d} \right) \frac{\partial y_e}{\partial x_d} = -p_{\mathbf{x}}(\mathbf{x}) \sum_{e=1}^D \frac{y_e}{\lambda_e} u_{de} \end{aligned} \quad (\text{A.2})$$

which is the d th element of vector $-p_{\mathbf{x}}(\mathbf{x})\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{y}$, which yields the gradient of $p_{\mathbf{x}}(\mathbf{x})$ as

$$\mathbf{g} = \nabla p_{\mathbf{x}}(\mathbf{x}) = p_{\mathbf{x}}(\mathbf{x})\Sigma^{-1}(\boldsymbol{\mu} - \mathbf{x}) \quad (\text{A.3})$$

Let $c \in \{1, 2, \dots, D\}$, and taking the second derivatives of $p_{\mathbf{x}}(\mathbf{x})$, we obtain

$$\begin{aligned} \frac{\partial}{\partial x_c} \left(\frac{\partial p}{\partial x_d} \right) &= \frac{\partial}{\partial x_c} \left(-p_{\mathbf{x}}(\mathbf{x}) \sum_{e=1}^D \frac{y_e}{\lambda_e} u_{de} \right) = -p_{\mathbf{x}}(\mathbf{x}) \frac{\partial}{\partial x_c} \left(\sum_{e=1}^D \frac{y_e}{\lambda_e} u_{de} \right) - \frac{\partial p_{\mathbf{x}}(\mathbf{x})}{\partial x_c} \sum_{e=1}^D \frac{y_e}{\lambda_e} u_{de} \\ &= -p_{\mathbf{x}}(\mathbf{x}) \left(\sum_{e=1}^D \frac{u_{de}}{\lambda_d} \frac{\partial y_e}{\partial x_c} \right) + p_{\mathbf{x}}(\mathbf{x}) \left(\sum_{e=1}^D \frac{y_e}{\lambda_e} u_{ce} \right) \left(\sum_{e=1}^D \frac{y_e}{\lambda_e} u_{de} \right) \\ &= -p_{\mathbf{x}}(\mathbf{x}) \left(\sum_{e=1}^D \frac{u_{de}}{\lambda_d} u_{ce} \right) + p_{\mathbf{x}}(\mathbf{x}) \left(\sum_{e=1}^D \frac{y_d}{\lambda_e} u_{ce} \right) \left(\sum_{e=1}^D \frac{y_d}{\lambda_e} u_{de} \right) \end{aligned} \quad (\text{A.4})$$

which is the (c, d) th element of the matrix $-p_{\mathbf{x}}(\mathbf{x})\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T + [p_{\mathbf{x}}(\mathbf{x})]^{-1} \mathbf{g}\mathbf{g}^T$ (Carreira-Perpinan 2000). Hence, the Hessian of $p_{\mathbf{x}}(\mathbf{x})$ is

$$\mathbf{H} = (\nabla \nabla^T) p_{\mathbf{x}}(\mathbf{x}) = -p_{\mathbf{x}}(\mathbf{x})\Sigma^{-1} + [p_{\mathbf{x}}(\mathbf{x})]^{-1} \mathbf{g}\mathbf{g}^T = p_{\mathbf{x}}(\mathbf{x})\Sigma^{-1} \left[-\Sigma + (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T \right] \Sigma^{-1} \quad (\text{A.5})$$

Results in (A.3) and (A.5) lead to the gradient and Hessian of the GMM $\sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x},k}, \Sigma_{\mathbf{x},k})$ as

$$\mathbf{g}_m = \nabla \left[\sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x},k}, \Sigma_{\mathbf{x},k}) \right] = \sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x},k}, \Sigma_{\mathbf{x},k}) \Sigma_{\mathbf{x},k}^{-1} (\boldsymbol{\mu}_{\mathbf{x},k} - \mathbf{x}) \quad (\text{A.6})$$

$$\begin{aligned} \mathbf{H}_m &= (\nabla \nabla^T) \left[\sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x},k}, \boldsymbol{\Sigma}_{\mathbf{x},k}) \right] \\ &= \sum_{k=1}^K \pi_k \phi(\mathbf{x}; \boldsymbol{\mu}_{\mathbf{x},k}, \boldsymbol{\Sigma}_{\mathbf{x},k}) \boldsymbol{\Sigma}_{\mathbf{x},k}^{-1} \left[-\boldsymbol{\Sigma}_{\mathbf{x},k} + (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x},k})(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x},k})^T \right] \boldsymbol{\Sigma}_{\mathbf{x},k}^{-1} \end{aligned} \quad (\text{A.7})$$

Since the mixing weight $w_k(\mathbf{x})$ of the regression function in Eq. (13) is a fraction of two GMMs, the following quotient rules, which is applied for calculating the derivatives of $f(x) = g(x)/h(x)$, can be utilized to obtain the gradient and Hessian of the regression function in Eqs. (15) and (16):

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{h \frac{\partial g}{\partial \mathbf{x}} - g \frac{\partial h}{\partial \mathbf{x}}}{h^2} \quad (\text{A.8})$$

$$\frac{\partial^2 f}{\partial \mathbf{x}^2} = \frac{\partial^2}{\partial \mathbf{x}^2} \left(\frac{g}{h} \right) = \frac{\partial^2 g / \partial \mathbf{x}^2 - 2 \left[\frac{\partial f}{\partial \mathbf{x}} \right] \left[\frac{\partial f}{\partial h} \right]^T - f \left[\partial^2 h / \partial \mathbf{x}^2 \right]}{h} \quad (\text{A.9})$$

References

- Abido MA (2003) A novel multiobjective evolutionary algorithm for environmental/economic power dispatch. *Electric Power Systems Research* 65:71-81. [https://doi.org/10.1016/S0378-7796\(02\)00221-3](https://doi.org/10.1016/S0378-7796(02)00221-3)
- AISC (2016a) Seismic Provisions for Structural Steel Buildings. ANSI/AISC 341-16, Chicago-Illinois
- AISC (2016b) Specification for structural steel buildings. ANSI/AISC 360-16, Chicago-Illinois
- Arendt PD, Apley DW, Chen W (2013) Objective-Oriented Sequential Sampling for Simulation Based Robust Design Considering Multiple Sources of Uncertainty. *Journal of Mechanical Design* 135. <https://doi.org/10.1115/1.4023922>
- ASCE (2017) Minimum design loads and associated criteria for buildings and other structures. ASCE 7-16, Reston, VA
- Bartlett FM, Dexter RJ, Graeser MD, Jelinek JJ, Schmidt BJ, Galambos TV (2003) Updating standard shape material properties database for design and reliability. *Engineering Journal-American Institute of Steel Construction* 40:2-14. <https://www.aisc.org/Updating-Standard-Shape-Material-Properties-Database-for-Design-and-Reliability>
- Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. Princeton University Press. <https://doi.org/10.1515/9781400831050>
- Beyer H-G, Sendhoff B (2007) Robust optimization – A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering* 196:3190-3218. <https://doi.org/10.1016/j.cma.2007.03.003>
- Brans JP, Vincke P, Mareschal B (1986) How to select and how to rank projects: The Promethee method. *European Journal of Operational Research* 24:228-238. [https://doi.org/10.1016/0377-2217\(86\)90044-5](https://doi.org/10.1016/0377-2217(86)90044-5)
- Campbell JG, Fraley C, Murtagh F, Raftery AE (1997) Linear flaw detection in woven textiles using model-based clustering. *Pattern Recognition Letters* 18:1539-1548. [https://doi.org/10.1016/S0167-8655\(97\)00148-7](https://doi.org/10.1016/S0167-8655(97)00148-7)
- Carreira-Perpinan MA (2000) Mode-finding for mixtures of Gaussian distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22:1318-1323. <https://doi.org/10.1109/34.888716>

- Chamroukhi F (2016) Unsupervised learning of regression mixture models with unknown number of components. *Journal of Statistical Computation and Simulation* 86:2308-2334.
<http://dx.doi.org/10.1080/00949655.2015.1109096>
- Chateaneuf A (2008) Principles of reliability-based design optimization. Structural design optimization considering uncertainties Taylor & Francis. <https://doi.org/10.1201/b10995-3>
- Chatterjee T, Chakraborty S, Chowdhury R (2019) A Critical Review of Surrogate Assisted Robust Design Optimization. *Archives of Computational Methods in Engineering* 26:245-274.
<https://doi.org/10.1007/s11831-017-9240-5>
- Chen S-J, Hwang C-L (1992) Fuzzy multiple attribute decision making methods. In: Fuzzy multiple attribute decision making. Springer, Berlin, Heidelberg, pp 289-486.
<https://doi.org/10.1007/978-3-642-46768-4>
- Crestaux T, Le Maître O, Martinez J-M (2009) Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety* 94:1161-1172.
<https://doi.org/10.1016/j.ress.2008.10.008>
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation* 6:182-197.
<https://doi.org/10.1109/4235.996017>
- Doğan E, Saka MP (2012) Optimum design of unbraced steel frames to LRFD–AISC using particle swarm optimization. *Advances in Engineering Software* 46:27-34.
<https://doi.org/10.1016/j.advengsoft.2011.05.008>
- Du X, Chen W (2000) Towards a better understanding of modeling feasibility robustness in engineering design. *J Mech Des* 122:385-394. <https://doi.org/10.1115/1.1290247>
- Dumonteil P (1992) Simple equations for effective length factors. *Engineering Journal-American Institute of Steel Construction* 29:111-115. <https://www.aisc.org/Simple-Equations-for-Effective-Length-Factors>
- Elishakoff I, Ohsaki M (2010) Optimization and Anti-Optimization of Structures Under Uncertainty. Imperial College Press. <https://doi.org/10.1142/p678>
- Fraley C, Raftery AE (1998) How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal* 41:578-588.
<https://doi.org/10.1093/comjnl/41.8.578>
- Fu G, Shih FY, Wang H (2011) A kernel-based parametric method for conditional density estimation. *Pattern Recognition* 44:284-294. <https://doi.org/10.1016/j.patcog.2010.08.027>
- Greco S, Matthias E, José F (2016) Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, New York, NY. <https://doi.org/10.1007/978-1-4939-3094-4>
- Guo X, Zhang W, Zhang L (2013) Robust structural topology optimization considering boundary uncertainties. *Computer Methods in Applied Mechanics and Engineering* 253:356-368.
<https://doi.org/10.1016/j.cma.2012.09.005>
- Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media. <https://doi.org/10.1007/b94608>
- Hess PE, Bruchman D, Assakkaf IA, Ayyub BM (2002) Uncertainties in material and geometric strength and load variables. *Naval engineers journal* 114:139-166.
<https://doi.org/10.1111/j.1559-3584.2002.tb00128.x>
- Ito M, Kim NH, Kogiso N (2018) Conservative reliability index for epistemic uncertainty in reliability-based design optimization. *Structural and Multidisciplinary Optimization* 57:1919-1935. <https://doi.org/10.1007/s00158-018-1903-9>
- Jalalpour M, Guest JK, Igusa T (2013) Reliability-based topology optimization of trusses with stochastic stiffness. *Structural Safety* 43:41-49. <https://doi.org/10.1016/j.strusafe.2013.02.003>
- Jin R, Du X, Chen W (2003) The use of metamodeling techniques for optimization under uncertainty. *Structural and Multidisciplinary Optimization* 25:99-116. <https://doi.org/10.1007/s00158-002-0277-0>
- Kiureghian AD, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? *Structural Safety* 31:105-112. <https://doi.org/10.1016/j.strusafe.2008.06.020>

- Konak A, Coit DW, Smith AE (2006) Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety* 91:992-1007.
<https://doi.org/10.1016/j.ress.2005.11.018>
- Lagaros ND, Papadrakakis M (2007) Robust seismic design optimization of steel structures. *Structural and Multidisciplinary Optimization* 33:457-469. <https://doi.org/10.1007/s00158-006-0047-5>
- Lee SH, Chen W, Kwak BM (2009) Robust design with arbitrary distributions using Gauss-type quadrature formula. *Structural and Multidisciplinary Optimization* 39:227-243.
<https://doi.org/10.1007/s00158-008-0328-2>
- Liu Z, Atamturktur S, Juang CH (2013) Performance based robust design optimization of steel moment resisting frames. *Journal of Constructional Steel Research* 89:165-174.
<https://doi.org/10.1016/j.jcsr.2013.07.011>
- Ma Q, Ohsaki M, Chen Z, Yan X (2019) Multi-objective optimization for prestress design of cable-strut structures. *International Journal of Solids and Structures* 165:137-147.
<https://doi.org/10.1016/j.ijsolstr.2019.01.035>
- Maheri MR, Narimani MM (2014) An enhanced harmony search algorithm for optimum design of side sway steel frames. *Computers & Structures* 136:78-89.
<https://doi.org/10.1016/j.compstruc.2014.02.001>
- Marler RT, Arora JS (2004) Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization* 26:369-395. <https://doi.org/10.1007/s00158-003-0368-6>
- MathWorks (2019) *Statistics and Machine Learning Toolbox User's Guide (r2019b)*.
- McLachlan GJ, Lee SX, Rathnayake SI (2019) Finite mixture models. *Annual review of statistics and its application* 6:355-378. <https://doi.org/10.1146/annurev-statistics-031017-100325>
- Ohsaki M, Kinoshita T, Pan P (2007) Multiobjective heuristic approaches to seismic design of steel frames with standard sections. *Earthquake Engineering & Structural Dynamics* 36:1481-1495.
<https://doi.org/10.1002/eqe.690>
- Ohsaki M, Yamakawa M, Fan W, Li Z (2019) An order statistics approach to multiobjective structural optimization considering robustness and confidence of responses. *Mechanics Research Communications* 97:33-38. <https://doi.org/10.1016/j.mechrescom.2019.04.003>
- Papadimitriou DI, Mourelatos ZP, Hu Z (2018) Reliability Analysis Using Second-Order Saddlepoint Approximation and Mixture Distributions. *Journal of Mechanical Design* 141(2):021401.
<https://doi.org/10.1115/1.4041370>
- Papadrakakis M, Lagaros ND, Plevris V (2005) Design optimization of steel structures considering uncertainties. *Engineering Structures* 27:1408-1418.
<https://doi.org/10.1016/j.engstruct.2005.04.002>
- Pezeshk S, Camp CV, Chen D (2000) Design of Nonlinear Framed Structures Using Genetic Optimization. *Journal of Structural Engineering* 126:382-388.
[https://doi.org/10.1061/\(ASCE\)0733-9445\(2000\)126:3\(382\)](https://doi.org/10.1061/(ASCE)0733-9445(2000)126:3(382))
- Qu X, Haftka RT (2004) Reliability-based design optimization using probabilistic sufficiency factor. *Structural and Multidisciplinary Optimization* 27:314-325. <https://doi.org/10.1007/s00158-004-0390-3>
- Roy CJ, Oberkampf WL (2011) A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering* 200:2131-2144. <https://doi.org/10.1016/j.cma.2011.03.016>
- Schuëller GI, Jensen HA (2008) Computational methods in optimization considering uncertainties – An overview. *Computer Methods in Applied Mechanics and Engineering* 198:2-13.
<https://doi.org/10.1016/j.cma.2008.05.004>
- Sung HG (2004) Gaussian mixture regression and classification. Dissertation, Rice University
- Toğan V (2012) Design of planar steel frames using Teaching–Learning Based Optimization. *Engineering Structures* 34:225-232. <https://doi.org/10.1016/j.engstruct.2011.08.035>
- Vahdani B, Mousavi SM, Hashemi H, Mousakhani M, Tavakkoli-Moghaddam R (2013) A new compromise solution method for fuzzy group decision-making problems with an application

to the contractor selection. *Engineering Applications of Artificial Intelligence* 26:779-788.
<https://doi.org/10.1016/j.engappai.2012.11.005>

Zhang X, Pandey MD (2013) Structural reliability analysis based on the concepts of entropy, fractional moment and dimensional reduction method. *Structural Safety* 43:28-40.
<https://doi.org/10.1016/j.strusafe.2013.03.001>