Analytical Robust Design Optimization for Hybrid Design Variables: An Active-learning Methodology Based on Polynomial Chaos Kriging

Chaolin Song[1], Abdollah Shafieezadeh[2*], Rucheng Xiao[1], Bin Sun[1]

[1]Department of Bridge Engineering, Tongji University, Shanghai, 200092, China
[2]Risk Assessment and Management of Structural and Infrastructure Systems (RAMSIS) Lab, Department of Civil, Environmental, and Geodetic Engineering, The Ohio State University, Columbus, OH, 43210, United States
[*]Email address of the corresponding author: shafieezadeh.1@osu.edu

ABSTRACT

In robust design optimization, statistical moments of performance are widely adopted in formulating robustness metrics. To address the high computational costs stemming from the many-query nature of such optimizations with respect to robustness metrics, analytical formulas of the statistical moments have been developed based on surrogate models. However, existing methods consider random variables as the sole model input, which excludes, from the application scope, problems that also involve deterministic design variables. To remedy this issue, this paper proposes a new Polynomial Chaos Kriging-based methodology for efficient and accurate analytical robust design optimization. The analytical solutions for the statistical moments of performance are developed considering that the Polynomial Chaos Kriging model is established in the augmented space of the deterministic design and random variables. This is achieved by systematically decoupling associations with deterministic input from random input, providing effective solutions even when the orthonormality of the basis function is not applicable in the augmented space. This work also presents an active-learning framework enabling seamless implementation of various numerical optimization methods. Several numerical examples and a practical application illustrate the performance and superiority of the proposed method.

Key words: *Robust design optimization, Polynomial Chaos Kriging, analytical formula, robustness index, hybrid random and deterministic design variables*

1. Introduction

Structural optimization has emerged as a powerful tool for developing optimal design solutions, taking into account various constraints within the design domain. Besides many successful implementations, it has been acknowledged that multiple sources of aleatory and epistemic uncertainties are able to influence the system performance [1]. Therefore, incorporating uncertainties into the design optimization has received much attention and become a vital branch of structural optimization. Reliability-based design optimization and robust design optimization are two allied but distinct approaches [2]. The former is concerned with providing the most desirable design while quantitatively setting constraints for the failure probability [3]. The latter, on the other hand, typically aims at optimizing the mean response of a system while reducing its sensitivity in the face of variations [4].

The earliest work for robust design was pioneered by Genichi Taguchi [5]. Subsequently, a number of studies extended Taguchi's method (see reviews by Beyer and Sendhoff [6], Zang et al. [7], and Park et al. [8]). With the developments of computational mechanics and numerical optimization methods, researchers have also made efforts to take a broader view on robust design [9]. Various formulations have been present as robustness metrics under optimization frameworks. In general, these measures can be grouped into the possibilistic type (e.g., fuzzy sets, convex models, information gap methods [10, 11]) and probabilistic type [4]. The latter, as the primary concern of this work, is well suited when sufficient information is provided to define the joint probability distribution of random parameters. In such cases, the mean and variance of structural response are typically consolidated to define the problem. Notably, compared with deterministic design optimization, a significantly higher computational challenge intrinsically exists due to the incorporation of uncertainty analysis into the optimization routine. Techniques grounded in surrogate modeling [12-14], which aim at creating a cheaper-to-evaluate mathematical model as a substitute for the original computational model, have garnered significant interest for mitigating the computational challenges in RDO [15, 16]. To fully take advantage of the information from surrogate models and eliminate the computational burden of uncertainty quantification, a key research direction is the development of analytical approaches for computing the robustness index. With such analytical estimation, solving a RDO problem can be as simple as handling a deterministic counterpart.

Surrogate models akin to Polynomial Chaos Expansion (PCE) and Polynomial Dimensional Decomposition (PDD) typically use a set of orthonormal polynomials to fit the trend of a generic output function. Taking advantage of the orthonormality feature, Polynomial basis-based models serve as powerful tools for analytically estimating various properties of the stochastic response. For example, Ren and Rahman [17] established analytical formulas for the first two statistical moments and first-order derivatives with respect to a variable using PDD and proposed a sequential gradient-based optimization framework. Lee and Rahman [18] further developed a generalized PCE model founded on a whitening transformation algorithm, and therefore random variables obeying dependent distributions can be better handled in their proposed framework. Chakraborty et al. [19] formulated an analytical expression for estimating robustness via hybrid Polynomial Correlated Function Expansion. Zhou et al. [20] developed the PC-GK-SBL surrogate model, which combined polynomial chaos expansion (PCE) and Gaussian kernel (GK) in the sparse Bayesian learning (SBL) framework. Liu et al. [21] investigated the analytical RDO method based on PC-GK-SBL, and proposed an active learning function to adaptively refine the training data combining the nearest distance to existing training samples and the distance to the located optimal solution.

Despite recent advancements in surrogate model-based analytical RDO, critical gaps within existing methodologies persist and warrant attention. In many practical engineering applications, despite facing a variety of internal and environmental uncertainties, some design variables should still be regarded as deterministic. This is either because the corresponding output is intrinsically certain (e.g., the design of the number of elements and other topological parameters) or because the associated uncertainty can be confidently eliminated through validation (e.g., the parameters of a damper can be accurately measured during the manufacturing process in some cases). However, most existing methods assume random variables as the sole input to the surrogate model and subsequently develop analytical formulas. These methods therefore struggle with problems that simultaneously incorporate deterministic design variables and random variables, because the primary assumption for derivation is not founded. Artificially assuming the deterministic variables as random inputs by introducing very small coefficients of variation is the simplest approach to transform the problem. However, this approach may not be appropriate since some design variables are inherently certain, thus introducing a conceptual inconsistency. Beyond that, treating such variables as random may considerably affect the efficiency of RDO by incurring additional computational burdens and introduce errors challenging the ability of RDO in minimizing the cost function and meeting constraint requirements.

To bridge the gap, this paper introduces a new method for efficient analytical robust design optimization considering hybrid random and deterministic design variables. The Polynomial Chaos Kriging (PCK) is used here as the surrogate model. PCK interprets the PCE in the universal Kriging framework to capture the global trend, making it a suitable choice due to its superior predictive capabilities as evidenced in [22]. The main contributions of this work are in the following two fronts. First, different from established methods [17-19, 21], this study proposes analytical solutions for the statistical moments based upon a global PCK surrogate model that is established on the augmented domain for all deterministic design and random variables. During the derivation process, this work systematically decouples the parts with deterministic input from those with random input, therefore yielding effective solutions even when the orthonormality of basis function is not applicable. Several classical numerical examples and a practical application of robust tuned mass damper (TMD) design optimization demonstrate the superiority of the proposed method compared to the state-of-the-art approaches. Additionally, compared with previous work [23], this paper symmetrically compares the method performance with and without the introduction of PCE trend for Kriging in RDO. Through investigations, it is evident that the PCK surrogate model generally perform slightly better than the Kriging surrogate model for solving RDO problems, because of the improvement of the ability to capture the global trend.

The paper is organized as follows. Section 2 recalls the fundamental theory about robust design optimization and the PCK surrogate model. Section 3 presents the proposed method, specifically active learning and PCK-assisted analytical RDO. Section 4 provides numerical and practical examples to illustrate the performance of the proposed method. Section 5 summarizes the conclusions of this work.

2. Preliminaries

This section presents an overview of robust design optimization including a typical formulation of RDO problems. This is followed by an introduction to Kriging and PCK surrogate models.

2.1 Robust Design Optimization

As an important branch of structural optimization, RDO aims at optimizing the design scheme while maintaining the objective or feasibility robustness. Let $\boldsymbol{d}$ denote the design variables of optimization, which consists of two parts, i.e., $\boldsymbol{d}_d$, the deterministic design variables, and $\boldsymbol{d}_\mu$, the mean value of some random variables. Let $\boldsymbol{\xi}$ denote the vector of random variables. The variance-based robust design optimization can be formulated as:

$$\min_{(\boldsymbol{d})} obj(\boldsymbol{d}_d, \boldsymbol{\xi})$$
$$s.t.: \begin{cases} con_i(\boldsymbol{d}_d, \boldsymbol{\xi}) \le 0, i = 1,2, \dots, n \\ \boldsymbol{d} = [\boldsymbol{d}_d, \boldsymbol{d}_\mu] \\ \boldsymbol{\xi} \sim f(\boldsymbol{\xi} | \boldsymbol{d}_\mu) \\ \boldsymbol{d}_l \le \boldsymbol{d} \le \boldsymbol{d}_u \end{cases} \tag{1}$$

where $obj(\cdot)$ denotes the objective function, $con_i(\cdot)$ denotes the constraint function ($i = 1,2, \dots, n$), $f(\boldsymbol{\xi}|\boldsymbol{d}_\mu)$ denotes the distribution of $\boldsymbol{\xi}$ given $\boldsymbol{d}_\mu$, and $\boldsymbol{d}_l$ and $\boldsymbol{d}_u$ denote the upper and lower bounds for the design variables, respectively. The weighted sum of the expected value and standard deviation of a performance function has been typically adopted to describe $obj(\cdot)$ and $con_i(\cdot)$, e.g., $\alpha_E \mathbb{E}[g(\boldsymbol{d}_d, \boldsymbol{\xi})] + \alpha_s \mathbb{S}[g(\boldsymbol{d}_d, \boldsymbol{\xi})]$ where $g(\cdot)$ denotes the performance function, α_E and α_s denote the weight factors. In some works, this problem is presented as a dual-objective optimization problem [24, 25]; however, due to space limit, the rest of this paper only focuses on the single-objective optimization as presented in Eq. (1).

2.2 Kriging

Kriging, developed by Krige [26] and Matheron [27], aims to predict the response of a function, based on the assumption that the function of interest is a realization of a stochastic Gaussian process. Let $\hat{g}$ denote the predictor of Kriging and $\boldsymbol{x}$ denote the vector of input; the predicted value of a point can be written as:

$$\hat{g}(\boldsymbol{x}) = F(\boldsymbol{x}) + z(\boldsymbol{x}) \tag{2}$$

where $F(\boldsymbol{x})$ is the regression part which can be expressed based on the regression model $\boldsymbol{f}(\boldsymbol{x})$ and regression parameters $\boldsymbol{\beta}$, and $z(\boldsymbol{x})$ is a stationary Gaussian random process.

When defining the covariance between the outputs of two points ($\boldsymbol{x}$ and $\boldsymbol{x}'$) based on the distance, for a dim dimensional problem, a typical Gaussian correlation model is:

$$R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x}') = \prod_{k=1}^{dim} r(\theta_k, x_k, x_k') = \prod_{k=1}^{dim} \exp\left[-\frac{1}{2}\left(\frac{x_k - x_k'}{\theta_k}\right)^2\right] \tag{3}$$

Subsequently, the mean and variance of $\hat{g}(\boldsymbol{x})$ can be expressed as [28]:

$$\mu_K(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{\beta}^* + \boldsymbol{r}^T\boldsymbol{R}^{-1}(\boldsymbol{Y}_m - \boldsymbol{f}(\boldsymbol{X}_m)\boldsymbol{\beta}^*) \tag{4}$$

$$\sigma_K^2(\boldsymbol{x}) = \sigma^2(1 + \boldsymbol{u}^{\boldsymbol{T}}(\boldsymbol{f}(\boldsymbol{X}_m)^T\boldsymbol{R}\boldsymbol{f}(\boldsymbol{X}_m))^{-1}\boldsymbol{u} - \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r}) \tag{5}$$

where $\boldsymbol{R}$ is the correlation matrix with $R_{ij} = R(\boldsymbol{\theta}, \boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)})$, and $\boldsymbol{r} = \{R(\boldsymbol{\theta}, \boldsymbol{x}^{(1)}, \boldsymbol{x}), \cdots, R(\boldsymbol{\theta}, \boldsymbol{x}^{(m)}, \boldsymbol{x})\}$. $\boldsymbol{\theta}$ is a vector of hyper-parameters that can be determined by the maximum likelihood estimation [29]. $\boldsymbol{\beta}^*$ denotes the generalized least-squares estimate.

2.3 PC-Kriging

Considering a vector $\boldsymbol{x}$ describing the input variable, Polynomial Chaos Expansions (PCE) [30] approximates the original performance function $g(\boldsymbol{x})$ by a sum of orthonormal polynomials:

$$\hat{g}(\boldsymbol{x}) = \sum_{\alpha \in \mathcal{A}} c_\alpha \Psi_\alpha(\boldsymbol{x}) \tag{6}$$

where $\Psi_\alpha(\cdot)$ denotes the multivariate orthonormal polynomials, c_α denotes the corresponding coefficient, α denotes the index of the order of polynomials, and $\mathcal{A}$ denotes the set of polynomials selected for approximating $g(\cdot)$ [31, 32]. As noticed, compared to traditional regression models such as linear or quadratic regressions, PCE shows better flexibility and performance in capturing the global trend of a function. Therefore, some researchers proposed the PC-Kriging (or called PCK in the rest of this paper) to take advantage of not only the global regression capability of PCE but also the interpolation and uncertainty quantification capabilities of Kriging [22].

Let $\boldsymbol{x}$ also denote the vector of input variables. The prediction of PCK can be formulated as:

$$\hat{g}(\boldsymbol{x}) = \sum_{\alpha \in \mathcal{A}} c_\alpha \Psi_\alpha(\boldsymbol{x}) + z(\boldsymbol{x}) \tag{7}$$

Compared with Eq. (2) the main difference lies in the introduction of the PCE as the global trend of the surrogate model. And therefore, the same correlation model, such as the Gaussian model of Eq. (3) and the corresponding hyperparameter $\boldsymbol{\theta}$ can be used to characterize the Gaussian random process $z(\boldsymbol{x})$.

3. Methodology: Active-learning PCK-assisted RDO

For a robust design optimization problem, as was explained in Section 2.1, the objective function ($obj(\cdot)$) or the constraint function ($con(\cdot)$) can be formulated as $\alpha_E\mathbb{E}[g(\boldsymbol{x})] + \alpha_s\mathbb{S}[g(\boldsymbol{x})]$, where $g(\cdot)$ denotes the performance function and α_E and α_s denote the weight factors. When both the deterministic design variables and random variables are considered as input, it is noticed that $\boldsymbol{x} = [\boldsymbol{d}_d, \boldsymbol{\xi}]$. Therefore, with the application of surrogate model, i.e., $\hat{g}(\cdot)$ as a substitute for the original function $g(\cdot)$, the statistical properties can be estimated by simulation methods, such as MCS. However, the nested structure of uncertainty quantification and optimization still challenges the solving procedure because numerous evaluations of the robustness metric can be required in the optimization routine. Motivated by using analytical solutions to fundamentally address this challenge and the genuine need to simultaneously incorporate random and deterministic design variables, this work proposes new analytical solutions based on the PCK surrogate model in the rest of this section. Several remarks such as the strategy for dealing with difference correction models or distributions are also provided at the end of this section.

3.1 Analytical solution of the expectation of PCK

Let us consider $g(\boldsymbol{x})$ as the objective of the RDO problem, where $\boldsymbol{x}$ denotes the input variables. As noticed, when the surrogate model is built on the augmented domain, the set of input variables can be presented as $\boldsymbol{x} = [\boldsymbol{d}_d, \boldsymbol{\xi}]$. Let dim_ξ denote the dimension of the input variable $\boldsymbol{\xi}$ with $\boldsymbol{\xi} = \{\xi_1, \cdots, \xi_{dim_\xi}\}$, and dim_d denote the dimension of the input variable $\boldsymbol{d}_d$ with $\boldsymbol{d}_d = \{d_{d,1}, \cdots, d_{d,dim_d}\}$. The dimension of the input of PCK is therefore $dim = dim_\xi + dim_d$.

With optimized hyperparameters $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_d^*, \boldsymbol{\theta}_\xi^*)$, $\boldsymbol{\theta}_d^* = \{\theta_{d,1}^* \cdots, \theta_{d,dim_d}^*\}$, $\boldsymbol{\theta}_\xi^* = \{\theta_{\xi,1}^* \cdots, \theta_{\xi,dim_\xi}^*\}$ and the optimized coefficients of PCE denoted by $\mathcal{A}^*$, the prediction of a PCK can be recast as:

$$\hat{g}(\boldsymbol{x}) = \sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x}) + \boldsymbol{r}^T \boldsymbol{R}^{-1}(\boldsymbol{Y}_m - F(\boldsymbol{X}_m)) \tag{8}$$

Let $\boldsymbol{\gamma}^*$ denote $\boldsymbol{R}^{-1}(\boldsymbol{Y} - F(\boldsymbol{X}_m))$. Note that $F(\boldsymbol{X}_m)$ is denoted by the PCE of Eq. (6), and $\boldsymbol{R}$ is the matrix with $R_{ij} = R(\boldsymbol{\theta}^*, \boldsymbol{x}, \boldsymbol{x}^{(i)})$. $\boldsymbol{R}^{-1}$, $\boldsymbol{Y}$ and $F(\boldsymbol{X}_m)$ are independent of the input $\boldsymbol{x}$. Considering that $\boldsymbol{\gamma}^* = [\gamma_i^*, \dots, \gamma_m^*]^{\mathrm{T}}$ where m denotes the number of training samples of PCK, Eq. (8) can be recast as:

$$\hat{g}(\boldsymbol{x}) = \sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x}) + \sum_{i=1}^{m} \gamma_i^* R(\boldsymbol{\theta}^*, \boldsymbol{x}, \boldsymbol{x}^{(i)}) \tag{9}$$

where $\boldsymbol{X}_m = \{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}\}$ denotes the database of the training samples. With input $\boldsymbol{x}$ consisting of $\boldsymbol{d}_d$ and $\boldsymbol{\xi}$, Eq. (9) can be further expanded to:

$$\hat{g}(\boldsymbol{x}) = \sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x}) + \sum_{i=1}^{m} \left[\gamma_i^* \prod_{j=1}^{dim_d} r(\theta_{d,j}^*, d_{d,j}, d_{d,j}^{(i)}) \prod_{k=1}^{dim_\xi} r(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)})\right] \tag{10}$$

Thus, the expectation for $\hat{g}(\boldsymbol{x})$ can be formulated as:

$$\begin{aligned} \mathbb{E}[\hat{g}(\boldsymbol{x})] &= \mathbb{E}\left[\sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x}) + \sum_{i=1}^{m} \gamma_i^* R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x}^{(i)})\right] \\ &= \mathbb{E}\left[\sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x})\right] + \sum_{i=1}^{m} \left(\gamma_i^* \cdot \prod_{j=1}^{dim_d} r(\theta_{d,j}^*, d_{d,j}, d_{d,j}^{(i)}) \cdot \mathbb{E}\left[\prod_{k=1}^{dim_\xi} r(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)})\right]\right) \end{aligned} \tag{11}$$

For the first part of Eq. (11), i.e., $\mathbb{E}[\sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x})]$, it worth noticing that due to deterministic design variable $\boldsymbol{d}_d$ is also considered as input, $\mathbb{E}[\sum_{\alpha \in \mathcal{A}} c_\alpha \Psi_\alpha(\boldsymbol{x})] \neq c_0$ that is only correct when $\boldsymbol{x}$ consists of all random variables and corresponding polynomial basis functions are adopted. Herein, let $\Psi_\alpha(\boldsymbol{x}) = \prod_{j=1}^{dim_d} \psi_\alpha(d_{d,j}) \cdot \prod_{k=1}^{dim_\xi} \psi_\alpha(\xi_k)$ generally denote the multi-variant polynomial that is the product of polynomial basis in each dimension. With independent input variables, the above equation can be reformulated as:

$$\mathbb{E}\left[\sum_{\alpha \in \mathcal{A}^*} c_\alpha \Psi_\alpha(\boldsymbol{x})\right] = \sum_{\alpha \in \mathcal{A}^*} \left(c_\alpha \cdot \prod_{j=1}^{dim_d} \psi_\alpha(d_{d,j}) \cdot \prod_{k=1}^{dim_\xi} \mathbb{E}[\psi_\alpha(\xi_k)]\right) \tag{12}$$

Furthermore, it is noticed that:

$$\mathbb{E}[\psi_\alpha(\xi_k)] = \int f(\xi_k | \boldsymbol{d}_\mu) \cdot \psi_\alpha(\xi_k) \mathrm{d}\xi_k \tag{13}$$

For the second part of Eq. (11), when the separated Gaussian correlation model as shown by Eq. (3) is adopted, and random variables are independent, $\mathbb{E}\left[\prod_{k=1}^{dim_\xi} r(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)})\right]$ can be formulated as:

$$\mathbb{E}\left[\prod_{k=1}^{dim_\xi} r(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)})\right] = \prod_{k=1}^{dim_\xi} \mathbb{E}\left(\exp\left[-\frac{1}{2}\left(\frac{\xi_k - \xi_k^{(i)}}{\theta_{\xi,k}^*}\right)^2\right]\right) \tag{14}$$

For a single component of the above equation, the expectation can be calculated as:

$$\mathbb{E}\left(\exp\left[-\frac{1}{2}\left(\frac{\xi_k-\xi_k^{(i)}}{\theta_{\xi,k}^*}\right)^2\right]\right)=\int\sqrt{2\pi}\theta_{\xi,k}^*\cdot f\left(\xi_k|\boldsymbol{d}_\mu\right)\cdot\mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta_{\xi,k}^*\right)^2\right]\mathrm{d}\xi_k \tag{15}$$

Therefore, combing Eqs. (11)-(15), the expected value of the prediction of PCK can be efficiently determined. Notably, the analytical solution for some parts of the equations depends on the distribution of random variables. For a comprehensive understanding, the formulas for the expectation when random variables follow Gaussian distributions are provided in Appendix A, and those for uniform distributions provided in Appendix B.

3.2 Analytical solution of the variance of PCK

This section focuses on the analytical solution for the variance of the PCK prediction. The variance of the stochastic output based on a PCK surrogate model can be formulated as:

$$\mathrm{var}\left(\hat{g}(\boldsymbol{x})\right)=\mathbb{E}[\hat{g}(\boldsymbol{x})^2]-\mathbb{E}[\hat{g}(\boldsymbol{x})]^2 \tag{16}$$

As the formula for determining $\mathbb{E}[\hat{g}(\boldsymbol{x})]$ has been provided in the previous section, $\mathbb{E}[\hat{g}(\boldsymbol{x})]^2$ can be easily calculated. The rest of this section focuses on the computation of $\mathbb{E}[\hat{g}(\boldsymbol{x})^2]$.

$$\begin{aligned}\mathbb{E}[\hat{g}(\boldsymbol{x})^2]=&\ \mathbb{E}\left[\left(\sum_{\alpha\in\mathcal{A}^*}c_\alpha\Psi_\alpha(\boldsymbol{x})+\sum_{i=1}^{m}\gamma_i^*R(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)})\right)^2\right]\\=\mathbb{E}\left(\left[\sum_{\alpha\in\mathcal{A}^*}c_\alpha\Psi_\alpha(\boldsymbol{x})\right]^2\right)&+2\times\mathbb{E}\left[\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}\gamma_i^*c_\alpha\Psi_\alpha(\boldsymbol{x})\,R(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)})\right]+\mathbb{E}\left(\left[\sum_{i=1}^{m}\gamma_i^*R(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)})\right]^2\right)\end{aligned} \tag{17}$$

For the first part of Eq. (17), it is worth noticing that due to the introduction of deterministic design variables, $\mathbb{E}[(\sum_{\alpha\in\mathcal{A}}c_\alpha\Psi_\alpha(\boldsymbol{x}))^2]\neq\sum_{\alpha\in\mathcal{A}}{c_\alpha}^2$. With independent variables, it can be formulated as:

$$\begin{aligned}&\mathbb{E}\left(\left[\sum_{\alpha\in\mathcal{A}^*}c_\alpha\Psi_\alpha(\boldsymbol{x})\right]^2\right)\\&=\sum_{\alpha\in\mathcal{A}^*}\sum_{\vartheta\in\mathcal{A}^*}\left(c_\alpha c_\vartheta\cdot\prod_{j=1}^{dim_d}\psi_\alpha\left(d_{d,j}\right)\psi_\vartheta\left(d_{d,j}\right)\cdot\prod_{k=1}^{dim_\xi}\mathbb{E}[\psi_\alpha(\xi_k)\psi_\vartheta(\xi_k)]\right)\end{aligned} \tag{18}$$

Then, $\mathbb{E}\left[\psi_\alpha\left(\xi^{(k)}\right)\psi_\vartheta\left(\xi^{(k)}\right)\right]$ can be obtained as:

$$\mathbb{E}[\psi_\alpha(\xi_k)\psi_\vartheta(\xi_k)]=\int f\left(\xi_k|\boldsymbol{d}_\mu\right)\cdot\psi_\alpha(\xi_k)\cdot\psi_\vartheta(\xi_k)\mathrm{d}\xi_k \tag{19}$$

For the second part of Eq. (17), $\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}\gamma_i^*c_\alpha\Psi_\alpha(\boldsymbol{x})\,R(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)})$ can be formulated as:

$$\begin{aligned}&\mathbb{E}\left[\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}\gamma_i^*c_\alpha\Psi_\alpha(\boldsymbol{x})\,R\left(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)}\right)\right]\\&\qquad=\mathbb{E}\left(\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}c_\alpha\Psi_\alpha(\boldsymbol{x})\left[\gamma_i^*\prod_{j=1}^{dim_d}r\left(\theta_{d,j}^*,d_{d,j},d_{d,j}^{(i)}\right)\prod_{k=1}^{dim_\xi}r\left(\theta_{\xi,k}^*,\xi_k,\xi_k^{(i)}\right)\right]\right)\end{aligned} \tag{20}$$

Let $\Psi_\alpha(\boldsymbol{x})=\prod_{j=1}^{dim_d}\psi_\alpha\left(d_{d,j}\right)\cdot\prod_{k=1}^{dim_\xi}\psi_\alpha\left(\xi_k\right)$ generally denote the multi-variant polynomial that is the product of the polynomial basis in each dimension. Subsequently, the above equation can be reformulated as:

$$\begin{aligned}&\mathbb{E}\left[\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}\gamma_i^*c_\alpha\Psi_\alpha(\boldsymbol{x})\,R\left(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{x}^{(i)}\right)\right]\\&=\sum_{i=1}^{m}\sum_{\alpha\in\mathcal{A}^*}\left(\gamma_i^*c_\alpha\cdot\prod_{j=1}^{dim_d}\psi_\alpha\left(d_{d,j}\right)r\left(\theta_{d,j}^*,d_{d,j},d_{d,j}^{(i)}\right)\cdot\mathbb{E}\left[\prod_{k=1}^{dim_\xi}\psi_\alpha(\xi_k)r\left(\theta_{\xi,k}^*,\xi_k,\xi_k^{(i)}\right)\right]\right)\end{aligned} \tag{21}$$

Let us also consider that the separated Gaussian correlation model is applied, and the random variables are independent. It is therefore $\mathbb{E}\left[\prod_{k=1}^{dim_\xi}\psi_\alpha(\xi_k)r\left(\theta_{\xi,k}^*,\xi_k,\xi_k^{(i)}\right)\right]=\prod_{k=1}^{dim_\xi}\mathbb{E}\left[\psi_\alpha(\xi_k)r\left(\theta_{\xi,k}^*,\xi_k,\xi_k^{(i)}\right)\right]$ which can be recast as:

$$\mathbb{E}\left[\psi_\alpha(\xi_k) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right)\right] = \sqrt{2\pi}\theta^*_{\xi,k} \int f(\xi_k|\boldsymbol{d}_\mu) \cdot \psi_\alpha(\xi_k) \cdot \mathcal{N}\left[\xi_k | \xi_k^{(i)}, \left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \tag{22}$$

For the third part of Eq. (17), $\mathbb{E}\left(\left[\sum_{i=1}^m \gamma_i^* R(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x}^{(i)})\right]^2\right)$ can be recast as:

$$\begin{aligned} &\mathbb{E}\left(\left[\sum_{i=1}^m \gamma_i^* R\left(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x}^{(i)}\right)\right]^2\right) \\ &= \mathbb{E}\left[\sum_{i=1}^m \sum_{t=1}^m \left(\gamma_i^* \gamma_k^* \cdot \prod_{j=1}^{dim_d} \left[r\left(\theta^*_{d,j}, d_{d,j}, d_{d,j}^{(i)}\right) r\left(\theta^*_{d,j}, d_{d,j}, d_{d,j}^{(t)}\right)\right] \cdot \prod_{k=1}^{dim_\xi} \left[r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(t)}\right)\right]\right)\right] \end{aligned} \tag{23}$$

Let $\chi_{i,t}$ denote $\gamma_i^* \gamma_t^* \prod_{j=1}^{dim_d} \left[r\left(\theta^*_{d,j}, d_{d,j}, d_{d,j}^{(i)}\right) r\left(\theta^*_{d,j}, d_{d,j}, d_{d,j}^{(t)}\right)\right]$ that is deterministic given i, t. The above equation can be simplified as:

$$\mathbb{E}\left(\left[\sum_{i=1}^m \gamma_i^* R\left(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{x}^{(i)}\right)\right]^2\right) = \sum_{i=1}^m \sum_{t=1}^m \left(\chi_{i,t} \cdot \mathbb{E}\left[\prod_{k=1}^{dim_\xi} r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(t)}\right)\right]\right) \tag{24}$$

With the same assumption that the probability density functions for the random variables are independent and the Gaussian correlation model is adopted, $\mathbb{E}\left[\prod_{k=1}^{dim_\xi} r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(t)}\right)\right]$ $= \prod_{k=1}^{dim_\xi} \mathbb{E}\left[r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(t)}\right)\right]$, where:

$$\begin{aligned} &\mathbb{E}\left[r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(i)}\right) r\left(\theta^*_{\xi,k}, \xi_k, \xi_k^{(t)}\right)\right] \\ &= 2\pi\left(\theta^*_{\xi,k}\right)^2 \int f(\xi_k|\boldsymbol{d}_\mu) \cdot \mathcal{N}\left[\xi_k | \xi_k^{(i)}, \left(\theta^*_{\xi,k}\right)^2\right] \cdot \mathcal{N}\left[\xi_k | \xi_k^{(t)}, \left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \end{aligned} \tag{25}$$

Therefore, combining the above equations, $\mathrm{var}\left(\hat{g}(\boldsymbol{x})\right)$ can be efficiently obtained based on the distribution of random variables, together with the standard deviation of the performance function $\mathbb{S}[\hat{g}(\boldsymbol{x})] = \sqrt{\mathrm{var}\left(\hat{g}(\boldsymbol{x})\right)}$. For the full analytical solutions, readers are referred to Appendix A and B for Gaussian and uniform distributions, respectively. For the integrity of this work, the authors make some remarks as follows, to further discuss the advantages of this analytical robustness formula and the assumptions made during the derivation.

Remark 1*: *On the distribution of the random variables and the correlation model

This section discusses the main assumptions made for the derivation of the analytical formula of the robustness index, i.e., the expectation and standard deviation of $\hat{g}(\cdot)$, as well as the way to release those assumptions. A key assumption is that the random variables are independent and the random variables $\boldsymbol{\xi}$ follow certain types of distributions (Gaussian or uniform distributions). For other cases, probabilistic transformation strategies can be adopted to smooth the performance function and to facilitate the calculation of the statistics. Another assumption is that the separated Gaussian correlation model is adopted for the PCK surrogate model. Because the separated Gaussian correlation model is the most popular one and due to the limit on paper length, this work does not present the full analytical formulations for other types of correlation models. Future works can expand upon this by reformulating Eq. (13), Eq. (22) and Eq. (25).

Remark 2: On the difference between PCK and Kriging

The main difference between the PCK and the Kriging surrogate model lies in the establishment of the trend function, i.e., $F(\boldsymbol{x})$ in Eq. (1) designed as $\boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{\beta}$ or $\sum_{\alpha \in \mathcal{A}} c_\alpha \Psi_\alpha(\boldsymbol{x})$. Taking advantage of the ability to capture the global trend of the function, PCK provides a group of optimized basis functions. In comparison, Kriging requires that the user define the type of global trend function before the implementation of the surrogate model. Despite this difference, the above derivations in Section 3.1 and 3.2 can be extended to Kriging [23]. On the other hand, as PCK provides better ability to capture the global trend of the function, PCK is seen to provide slightly better performance when estimating the statistical moments than Kriging for the same training dataset. This is validated by the comparison in Appendix D for the performance of PCK and ordinary Kriging for five classical numerical examples.

Remark 3: On the objective of analytical robustness formula

To further highlight the motivation behind this work, Appendix E compares the computational cost of estimating the statistical moments using both MCS and the formulated analytical equations. It is evident that the proposed analytical formulas are significantly faster than repeated sampling, offering nearly 100 times improvement in cost for some cases. Considering that global search optimization methods, such as evolutionary algorithms, are applied for global searching or improved handling of problems with discrete design variables, the proposed analytical formula prove immensely beneficial as numerous evaluations of the robustness index might be needed during the optimization loop. Additionally, when viewed as a specific interpretation of Kriging, it is evident that estimating the statistical moments of PCK is more costly than Kriging, for both analytical estimation and MCS approaches. This is mainly because the trend function of a PCK surrogate model is inherently not simpler than that of an ordinary Kriging.

Remark 4*: *On the deterministic design variables

As one of the primary motivations of this work, this remark presents the main difference due to the introduction of deterministic design variables. Notably, for problems without deterministic design variables, the above equations can be simplified by removing all terms associated with d (see Eqs. (11), (18), (21) and (23)). The simplified equations are consistent with the previous work [20], despite the difference between the adopted surrogate models.

Remark 5*: *About the scaling

It is worthy to note that in order to guarantee the numerical stability, an auxiliary space can be introduced by scaling the input variables, i.e., $\boldsymbol{x} \rightarrow \boldsymbol{u}$ by:

$$u_j = \frac{x_j - \mu(x_j)}{\sigma(x_j)} \quad j = 1,2, \ldots, dim \tag{26}$$

In this case, the optimized correlation parameter $\boldsymbol{\theta}^*$ is based on the auxiliary space, and therefore scaling of the training data is needed when calculating the statistics. Let $\boldsymbol{X}_m = \{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}\}$ denote the training samples in the original space and $\boldsymbol{U}_m = \{\boldsymbol{u}^{(1)}, \cdots, \boldsymbol{u}^{(m)}\}$ denotes the training samples in the auxiliary space. It is $\boldsymbol{x}^{(i)} = \left[\boldsymbol{\xi}^{(i)}, \boldsymbol{d}_d^{(i)}\right]$ and $\boldsymbol{u}^{(i)} = \left[\boldsymbol{u}_\xi^{(i)}, \boldsymbol{u}_{d_d}^{(i)}\right]$ ($i = 1,2, \cdots, m$). The parameters and bounds for the distributions of random variables can also be scaled. The formulated analytical expressions are still feasible in the auxiliary space after scaling.

3.3 The proposed framework

It is worth noticing that when taking advantage of the proposed analytical formula to express robustness index, it is still vital to establish an accurate mapping from the input to output space, i.e., guaranteeing the accuracy of the surrogate model especially in the region of interest. To this end, following the well-known concept of active learning, a generalized framework for efficient PCK-assisted robust design optimization is presented in this section. Fig. 1 illustrates the flowchart of the proposed method.

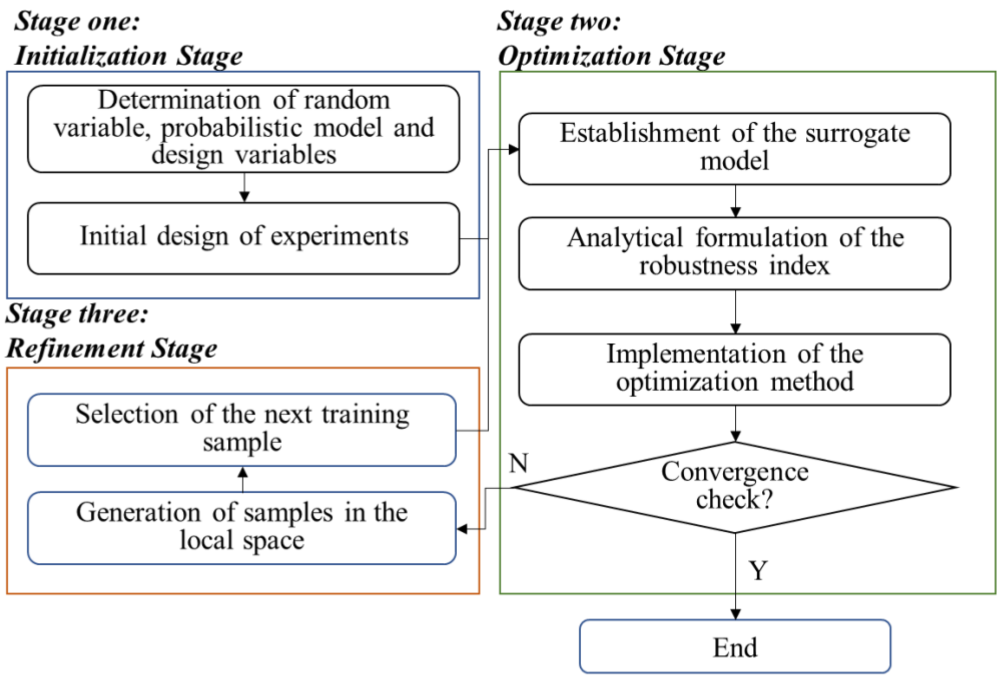


Fig. 1 Flowchart of the proposed RDO framework

As noticed, the proposed method mainly consists of three stages. The first step is named as ***Initialization Stage***. In this stage, based on the robust design optimization problem, the associated design variables, random variables and the probabilistic model should be established. The formulation of the RDO problem should be set up as indicated in Eq. (1). For performance functions involving complex computational models, an initial design of experiments should be determined, and the performance models are accordingly evaluated. Techniques such as Latin hypercube sampling can be used to generate initial training samples.

Subsequently, the framework enters the next stage: ***Optimization Stage***. The PCK surrogate model is firstly established in the augmented space, i.e., $\boldsymbol{d}_d \times \boldsymbol{\xi}$. Then, based on the proposed analytical formula of robustness in Section 3.2 and Section 3.3, the objective and constraint functions are analytically expressed. Based on the feature of the problem, such as discrete or continuous optimization, single-modal and multi-modal optimization, the numerical optimization method is selected. The RDO problem is subsequently solved with the current surrogate model. Three criteria are adopted in this paper to guarantee the accuracy of the final solution. First, the number of implementations of an optimization algorithm must be larger than Ψ_n. Furthermore, let Ψ_{mse} denote the threshold for the mean squared error. The mean value for the mean squared error (σ_K/μ_K) of local samples should be smaller than Ψ_{mse} to guarantee the accuracy of the surrogate model in the local region. Also, the change of the optimal solution in the current iteration and the last iteration should be smaller than the threshold Ψ_Δ to ensure the robustness of the solution. If all criteria are satisfied, the method ends and the optimal solution is provided. Otherwise, the method enters the third stage for the refinement of the surrogate model.

The third stage of the proposed method is the ***Refinement Stage*** of the surrogate model. A group of realizations are firstly sampled centered on the optimal solution from Stage Two. This work adopts the original distribution of random variables to general samples, i.e., implementing the MCS. Then, the probability density of the sample and the uncertainty level of the prediction are combined as the learning function to select the next training sample, e.g., $LF(\boldsymbol{x}) = PDF(\boldsymbol{x}) \cdot \sigma_K(\boldsymbol{x})$. The sample that maximizes the learning function is selected and evaluated on the original function. After refinement of the training database, the method goes back to the second stage for the next implementation of optimization.

4. Numerical Examples

In this section, the proposed method is carefully investigated on several numerical examples and a practical engineering application about TMD design optimization. Multiple other surrogate-based methods are compared to indicate the superiority of the proposed method. For a fair comparison, these methods are implemented 10 times independently. The results are then compared, either by using their average or by focusing on those with median performance in discrete problems. The UQLab toolbox (version 2.0.0) is adopted for the establishment of the surrogate models, e.g., PCK, Kriging and etc. Either a Gaussian or uniform distribution is defined to cover the design space of a design or random variable, thus facilitating the establishment of a PCK model in the toolbox. Hybrid Genetic Algorithm (HGA) is selected as the method for training PCK and Kriging [28]. All numerical experiments are implemented based on a computer with AMD Ryzen 5900HX CPU, RAM 32 GB.

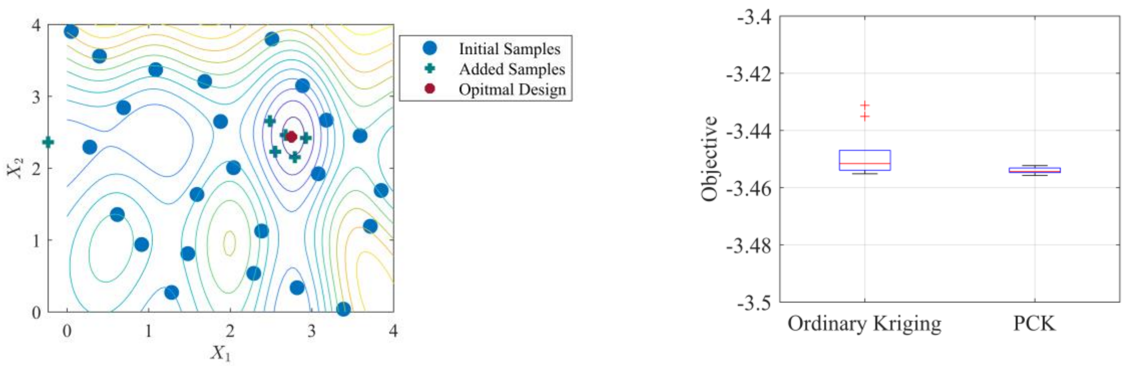
4.1 Example one

The first example is a two-dimensional numerical problem, named 2D Haupt function [21]. Two independent random variables are considered in this example. The mean values of the random variables are taken as the design variables with standard deviation of 0.2.

$$\min_{(\boldsymbol{d}_\mu)} \mu_g + \sigma_g$$
$$where: \begin{cases} g(\boldsymbol{\xi}) = \xi_1 \sin(4\xi_1) + 1.1\xi_2 \sin(2\xi_2) \\ \xi_i \sim \mathcal{N}\left(d_{\mu,i}, 0.2^2\right), i = 1,2 \\ 0 \le d_{\mu,i} \le 4, i = 1,2 \\ \boldsymbol{d}_\mu = \left[d_{\mu,1}, d_{\mu,2}\right] \end{cases} \tag{27}$$

As this example is a multimodal optimization problem, a heuristics optimization method, called the improved (μ+λ) differential evolution (IDE) [33], is applied to search for the best solution. Table 1 compares the results from the proposed method and several other techniques, including relevance vector machine (RVM), radial basis function (RBF), artificial neural network (ANN), sparse polynomial chaos expansion based on the least angle regression (PCE-LAR), Kriging with passive learning, Kriging with Expected Improvement (EI) learning function and Kriging with Mean Squared Error (MSE) learning function [21]. As the PCK can be considered as a specific interpretation of Kriging, in this context, the ordinary Kriging-based proposed framework is also compared here to illustrate the influence of the surrogate model. From the presented results, it can be observed that:

- The use of surrogate models, such as ANN, RBF or Kriging, can significantly reduce the calls to the performance function when solving this RDO problem; however, the direct application of these techniques may not guarantee that a true global optimal solution is found. The maximum relative error incurred by the surrogate model, such as RBF, can be as large as 122%. On the other hand, after adaptive refining the surrogate model, such as Kriging-EI and Kriging-MSE, the methods can yield much better solutions, and the proposed method generally achieves the best performance.
- The 'No' or 'Yes' is indicated in the last column of Table 1 to distinguish the methods with analytical solutions from the methods with Monte Carlo Simulation to assess the robustness index. Because the evaluations of robustness index are required many times in the optimization routine, surrogate model-based methods still suffer from the nested loop of optimization and uncertainty quantification with the lack of analytical solutions. For instance, taking into account 950 evaluations of robustness index as counted by the optimization with crude MCS, considerable computational time is required to solve this two-dimensional problem even with surrogate models. In comparison, the proposed surrogate model-based analytical method consumes only tens of seconds to accurately finish the whole computation; the computational efficiency can be boosted over 100 times. The advantage of developing analytical robust optimization can be therefore clearly shown by the avoidance of such tremendous computational cost.



(a) the establishment of the PCK model

(b) the performance of the proposed method with PCK and Ordinary Kriging

Fig. 2 The establishment of the surrogate model and comparison of results

Table 1 Comparison of the results from the proposed and other methods for Example One.

Method	Optimal Solution [a]	Objective	$\Delta_{obj}(\%)$	Sample Size	Analytical Solution
MCS	(2.7489, 2.4369)	-3.4548	-	10^7×950 [b]	No
RVM [c]	(2.4225, 2.4100)	-1.5399	55.42	32	No
ANN [c]	(1.7615, 2.2991)	-0.4049	88.27	32	No
RBF [c]	(1.4696, 1.5311)	0.7726	122.36	32	No
PCE-LAR [c]	(2.4215, 2.4338)	-1.5369	55.51	32	No
Kriging [c]	(2.2714, 2.4123)	-0.6384	81.51	32	No
Kriging-EI [c]	(2.7632, 2.4255)	-3.4474	0.23	24+4.4	No
Kriging-MSE [c]	(2.7563, 2.4119)	-3.4483	0.18	24+23.6	No
PC-GK-SBL [c]	(2.5921, 2.4374)	-2.8215	18.29	32	Yes
PC-GK-SBL-RLGE [c]	(2.7491, 2.4373)	-3.4543	0.01	31.1	Yes
Proposed method-Ordinary Kriging	(2.7473, 2.4273)	-3.4547	≈0	30.3	Yes
Proposed method-PCK	(2.7486, 2.4360)	-3.4544	0.01	30.3	Yes

- Compared with the existing surrogate model-based analytical RDO method [21], i.e., PC-GK-SBL with passive learning or PC-GK-SBL with RLGE learning function, the proposed method can achieve slightly better performance regarding the accuracy and efficiency of analysis. The PCK-based proposed method requires only 30.3 calls of the performance function on average, showing a relative error of nearly zero. Fig. 2 (a) shows the sequential sampling for the training database of PCK. The proposed method can accurately locate the true global optimum and converge to the optimum through adaptive refinements around the global optimal solution. When the Ordinary Kriging is adopted with the proposed analytical formula of robustness and active-learning framework, the performance is still good on average, however as Fig. 2 (b) shows, the PCK-based proposed method can perform slightly more stable compared with the Kriging-based approach. This is consistent with the aforementioned investigation in Appendix D.

[a] the results are averaged over 10 independent runs.

[b] the sample size for MCS indicates (the number of realizations to estimate the statistics) × (the number of evaluations of the robustness index during the optimization loop).

[c] [21] is referred to obtain the performance for the other surrogate model-based methods.

4.2 Example two

The second example consists of a revised robust design optimization problem for a truss structure, shown by Fig. 3 [17]. Three independent random variables are involved in this problem, i.e., the mass density ξ_1, the applied external loading ξ_2, and the material yield (tensile) strength ξ_3. Two deterministic design variables are considered, i.e., the cross-section area d_1 and the half of the distance between the two bottom nodes d_2. To highlight the features of the proposed method, d_1 is modeled as a discrete design variable with the candidate set of $\{10, 11, \cdots, 20\}$(unit: cm^2). The heuristics constrained optimization method, called the improved (μ+λ) constrained differential evolution (ICDE) [33], is applied in this example. Table 2 presents the distribution for the above random variables. Two constraints regarding the maximum stresses for the bars are considered. The robust design optimization problem is formulated as:

$$
\begin{gathered}
\min_{(\boldsymbol{d}_d)} 0.05\mu_{g_0} + 0.25\sigma_{g_0} \\
s.t.: \begin{cases} c_1(\boldsymbol{d}_d, \boldsymbol{\xi}) = 3\sigma_{g_1} - \mu_{g_1} \le 0 \\ c_2(\boldsymbol{d}_d, \boldsymbol{\xi}) = 3\sigma_{g_2} - \mu_{g_2} \le 0 \end{cases} \\
where: \begin{cases} g_0(\boldsymbol{d}_d, \boldsymbol{\xi}) = \xi_1 d_1 \sqrt{1 + d_2^2} \\ g_1(\boldsymbol{d}_d, \boldsymbol{\xi}) = 1 - \dfrac{5\xi_2\sqrt{1 + d_2^2}}{\sqrt{65}\xi_3}\left(\dfrac{8}{d_1} + \dfrac{1}{d_1 d_2}\right) \\ g_2(\boldsymbol{d}_d, \boldsymbol{\xi}) = 1 - \dfrac{5\xi_2\sqrt{1 + d_2^2}}{\sqrt{65}\xi_3}\left(\dfrac{8}{d_1} - \dfrac{1}{d_1 d_2}\right) \\ d_1 \in \{10, 11, \cdots, 20\} \text{ (unit: cm}^2\text{)} \\ 0.1 \le d_2 \le 1.6 \text{ (unit: m)} \\ \boldsymbol{d}_d = [d_1, d_2] \end{cases}
\end{gathered}
\tag{28}
$$

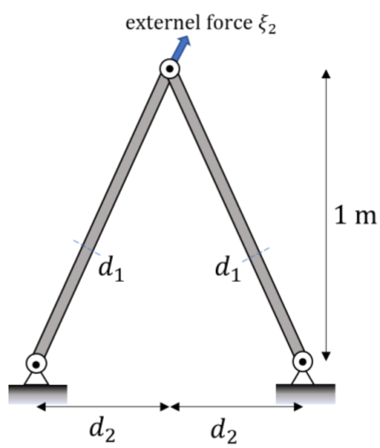


Fig. 3 The truss structure of Example two

Table 2 The design variables and distribution of random variables of Example Two.

Variables	Distribution	Mean	SD	Unit
Cross-sectional area (d_1)	Deterministic	d_1	—	cm^2
Half-horizontal span (d_2)	Deterministic	d_2	—	m
Mass density (ξ_1)	Beta	10,000	2,000	kg/m^3
Load (ξ_2)	Gumbel	800	200	kN
Yield strength (ξ_3)	Lognormal	1,050	250	MPa

Table 3 Comparison of the results from the proposed and other methods for Example Two.

Method	Optimal Solution [a]	Objective	$\Delta_{obj}(\%)$	Constraint1 $-c_1$	Constraint2 $-c_2$	Sample Size	Analytical Solution
MCS	(12.00, 0.312)	1.257	—	≈0	-0.572	$10^5 \times 785$ [b]	No
SVR	(13.00, 0.334)	1.368	8.83%	-0.086	-0.584	128	No
RBFNN	(10.00, 0.104)	1.003	20.21%	0.819	-0.996	128	No
PCE-LAR	(16.00, 0.102)	1.606	27.76%	0.144	-0.997	128	No
Ordinary Kriging	(13.00, 0.429)	1.303	3.66%	-0.040	-0.473	128	No
Proposed method-Ordinary Kriging	(12.00, 0.332)	1.265	0.64%	-0.012	-0.552	88.5	Yes
Proposed method-PCK	(12.00, 0.329)	1.261	0.32%	-0.007	-0.554	87.5	Yes

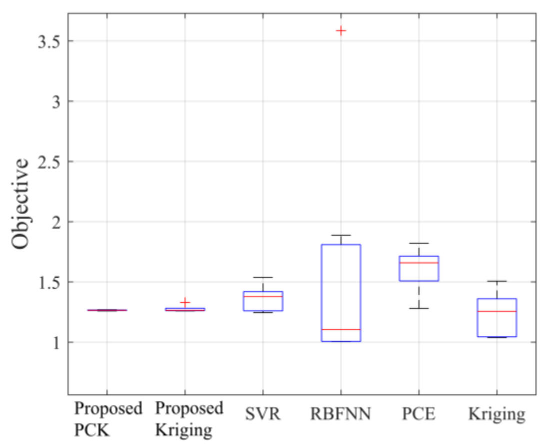


Fig. 4 The comparison of results for Example two

Table 3 compares the results from the proposed method, and from multiple other surrogate models, including support vector regression (SVR), radial basis function neural network (RBFNN), sparse polynomial chaos expansion based on the least angle regression (PCE-LAR), and ordinary Kriging [28]. Fig. 4 depicts the boxplots that compares the performance of these methods, and it was noticed that:

- As illustrated in Fig. 4, among the techniques used for comparison, SVR generally delivers the best performance with all solutions falling into the feasible domain. RNFNN exhibits the worst performance, characterized by considerable fluctuations in the results obtained, and the median performance solutions from both RBFNN and PCE are infeasible designs with respect to the first constraint function. The

[a] the median performance solution is compared, and the averaged required sample is compared.
[b] the sample size for MCS indicates (the number of realizations to estimate the statistics) × (the number of evaluations of the robustness index during the optimization loop).

performance of the proposed method is the best among all approaches. The required calls for the proposed methods with ordinary Kriging and PCK are very similar (88.5 verse 87.5). Compared with MCS, the relative error from the proposed method with PCK is only 0.32%, while the one for the proposed method with ordinary Kriging is relatively larger (0.6%). For the discrete design variable, i.e., the cross section of the trusses, the PCK-based proposed method yields an accurate result of 12 cm^2 for all independent runs, while that for Kriging-based proposed method can result in a section design of 11 cm^2 occasionally. Therefore, as also illustrated in Fig. 4, the performance of the PCK-based proposed method is more stable for the investigated truss problem. The results also confirm the enhancement achieved through the incorporation of the PCE trend into the Kriging surrogate model.

- The proposed method remains the capability of providing analytical solution of robustness index for the example with both deterministic design variables and random variables. The existing method, i.e., the PC-GK-SBL investigated in the last example, fails to provide the analytical solutions, because this method does not consider the involvement of deterministic design variables during the derivation, and therefore, the PC-GK-SBL is not compared in this example. This demonstrates the contribution of the proposed method with respect to extending the applicable scope of the analytical RDO methodology. Artificially introducing uncertainties into deterministic design variables can address challenges associated with hybrid random and deterministic variables; however, this approach may also introduce additional errors. For example, if considering that both the cross-sectional area and half-horizontal span follow lognormal distributions with a coefficient of variation of 0.05, the optimal design changes to (12, 0.331), representing approximately a 6% deviation in the span arrangement. Thus, it is vital to properly define the distribution parameters when treating deterministic variables as random.

4.3 Example three

The third example involves the robust topology design of a frame structure with element members considering the variations in the external loading [34, 35]. The topological design involves the definition of the number of active members and the corresponding sizes. The RDO problem is formulated as:

$$\min_{(\boldsymbol{d})} \mathbb{E}\left[g_y(\boldsymbol{d},\boldsymbol{\xi})\right] + 6\mathbb{S}\left[g_y(\boldsymbol{d},\boldsymbol{\xi})\right]$$
$$where: \begin{cases} \sum_{i=1}^{10} l_i d_i \leq 5.43 \text{ (unit: cm}^2) \\ \boldsymbol{\xi} = [\xi_{F,v}, \xi_{F,h}] \\ \boldsymbol{d} = [d_1, d_2, \cdots, d_6] \end{cases} \tag{29}$$

where $g_y(\cdot)$ is defined as the weighted sum of displacements along the external forces, with weight factors of five for the vertical displacement and one for the horizontal displacement. Fig. 5 depicts the layout of the frame structure and the numbering of its elements. $\xi_{F,h}$ and $\xi_{F,v}$ denote random variables for the load magnitudes applied in the horizontal and vertical directions, respectively. $\xi_{F,v}$ and $\xi_{F,h}$ are assumed to follow Gumbel distributions; the mean values for both are 100 kN, while their coefficients of variation are 0.2 and 0.02, respectively. The elastic modulus is considered as 100 GPa. Therefore, the objective is to optimize the frame design to design to ensure structural compliance, taking into account variations in the external loading and constraint on the material consumption. To find the optimum topology, each design variable d_i, representing the section areas for the corresponding frame members, can be selected from a discrete set. Zero is adopted to encode the inactivity of a member, i.e., $d_i \in \{0,1,2,\cdots,10\}$ ($unit$: cm^2). With introduction of the structural symmetry, total six design variables are considered in this problem.

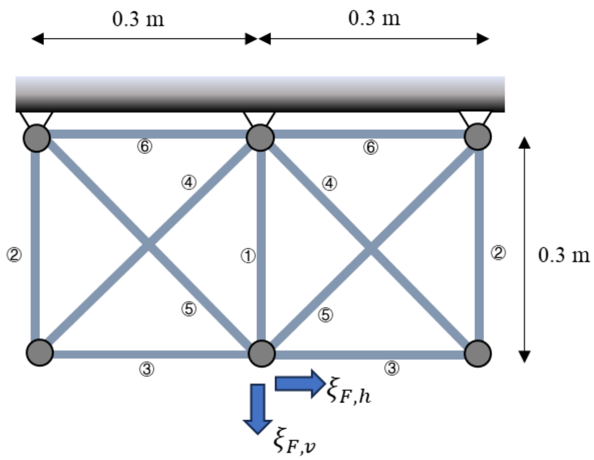


Fig. 5 The frame structure of Example three

The ICDE optimization method is also applied in this example to find the global optimal solution with discrete design variables. Table 4 compares the results from the proposed method, and from multiple other techniques, including SVR, RBFNN, PCE-LAR, and ordinary Kriging. Fig. 6 (a) depicts the optimum structural topology by the proposed method, and Fig. 6 (b) shows the boxplot for the performance of these methods. From the results, it can be noticed that:

- For multiple surrogate models employing passive learning with 128 training samples, the SVR method can generally facilitate the determination of the optimal structural topology, namely, for elements #1 and #5. However, the selection of element sizes is typically inconsistent with the solution determined by MCS. As shown in Fig. 6(b), the results from SVR exhibit certain discrepancies when compared with the true optimal solution. Moreover, with methods like ordinary Kriging, PCE and RBFNN, the determined topology can occasionally significantly differ from the optimal one. For example, RBFNN determined a result of (3,0,1,2,0,1) in one analysis. Thus, these methods are prone to incurring larger errors in this topology design problem.

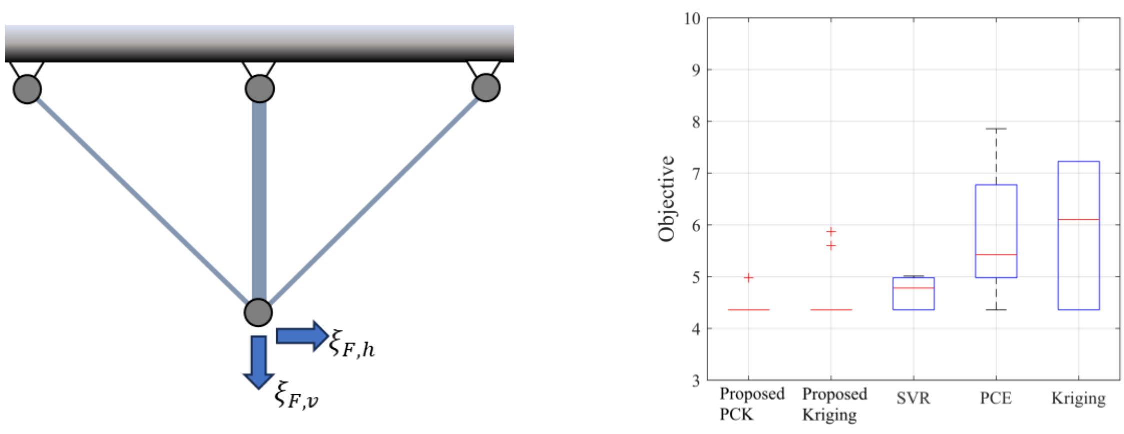
Table 4 Comparison of the results from the proposed and other methods for Example Three.

Method	Optimal Solution [a]	Objective (mm)	$\Delta_{obj}(\%)$	Sample Size	Analytical Solution
MCS	(9, 0, 0, 0, 3, 0)	4.34	—	$10^4 \times 5580$ [b]	No
SVR	(6, 0, 0, 0, 4, 0)	4.75	9.45	128	No
RBFNN	(3, 0, 1, 2, 0, 1)	792.49	>100	128	No
PCE-LAR	(10, 1, 0, 0, 2, 0)	4.95	14.06	128	No
Ordinary Kriging	(10, 1, 0, 1, 1, 0)	7.20	65.90	128	No
Proposed method-Ordinary Kriging	(9, 0, 0, 0, 3, 0)	4.34	0	82.7	Yes
Proposed method-PCK	(9, 0, 0, 0, 3, 0)	4.34	0	85.4	Yes

- On the other hand, the proposed method is the only technique that retains the capability to provide an analytical solution for robustness. It is worth noting that due to the involvement of the topological design variable, i.e., the inactivity of a member encoded by '0', artificially introducing uncertainty around

[a] the median performance solution is compared, and the averaged required sample is compared.
[b] the sample size for MCS indicates (the number of realizations to estimate the statistics) × (the number of evaluations of the robustness index during the optimization loop).

a deterministic design is not often feasible, given the convergence issues with FEM analysis and the discrete nature of topological design. Thus, this application further emphasizes the primary concern and motivation of this work. The proposed method, employing a PCK surrogate model, achieves the best performance among all investigated methods. As depicted in Fig. 6(b), with an average of 85.4 runs of the performance function, the proposed method with PCK can typically accurately find the true optimal design. In comparison, the proposed method using ordinary Kriging requires 82.7 runs on average and occasionally results in a sub-optimal design. Therefore, as demonstrated in this example, the integration of the PCE trend in the metamodel contributes to the accuracy and robustness of the method's performance, even in RDO problem involving topological design.

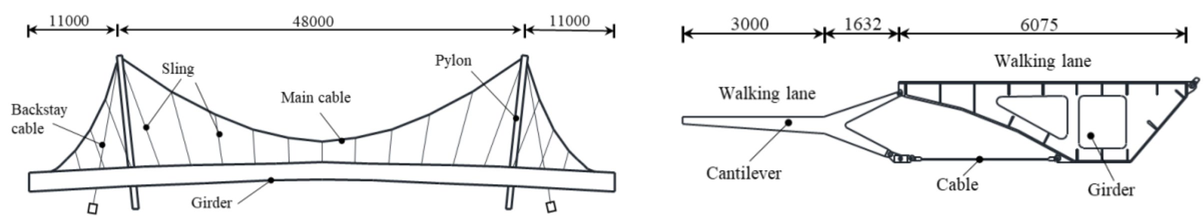


(a) the optimum topology (b) the comparison of results from different methods

Fig. 6 The optimum design for Example three

4.4 Example four

The last example of this paper focuses on the application of the proposed methodology on the robust design optimization of a tuned mass damper (TMD) for a footbridge, which is a cable-supported bridge with a span arrangement of 11 m+48 m+11 m. The steel girder has an asymmetric section design with a height of 2 m, and the pylon is designed as a concrete-filled steel tube column with a 0.9 m diameter. The backstay-cable is fixed on the top of the pylon. The details about the bridge design can refer to [36]. Fig. 7 (a) illustrates the layout the bridge, and Fig. 7 (b) presents the design scheme of the section.



(a) the layout of the bridge (unit: mm) (b) the design of the cross section (unit: mm)

Fig. 7 The cable-supported footbridge

Pedestrian-induced vibrations have caused serious challenges for bridge safety in many applications. The walking load incurred by a single pedestrian is modeled by the sum of multiple harmonic components written as a Fourier series. Moreover, referring to the existing work [37], a uniform distribution is adopted to model the walking load incurred by a group of pedestrians on the bridge. The equivalent uniformly distributed walking load can be therefore formulated as:

$$q_{eq} = \frac{N_{eq}}{S} \alpha_{e,h} G \psi(f_h) \tag{30}$$

where S denotes the surface of the bridge deck, N_{eq} is the equivalent number of perfectly synchronized pedestrians (N_{eq} can be calculated as $1.85\sqrt{N}$, where N denotes the number of pedestrians on the bridge based on the given pedestrian density [38, 39]), $\psi(f_h)$ denotes the reduction factor considering the possibility of the step frequency equaling the natural frequency f_h of the bridge, and $\alpha_{e,h}$ is the dynamic loading factor of the h^{th} harmonic of the load in direction of e (vertical direction or horizontal direction).

Imposing the equivalent uniformly distributed load on the bridge and assuming that the mode h dominates the response of the bridge, the projection of the load on the mode h can be calculated as:

$$F_h = q_{eq} \sum_{k=1}^{n_{eff}} a_{eff,k} |\phi_{h,k}| \tag{31}$$

where n_{eff} denotes the number of nodes of the bridge deck area, a_{eff} denotes the vector of the bridge deck area ($\sum_{k=1}^{n_{eff}} a_{eff} = S$), and ϕ_h denotes the vector for the mass-normalized modal displacement of mode h. Then, the maximum acceleration of the bridge can be calculated through dynamic analysis. Considering the design scheme of the background bridge, a finite element model is established to determine the dynamic parameters of the bridge, as Fig. 8 shows. The natural frequency for the first bending mode of the bridge is 2.0 Hz which falls in the sensitive range for the pedestrian-induced excitations. Therefore, it is considered that this mode dominates the vibration of the footbridge. Based on the dynamic analysis, it is noticed that the maximum acceleration can be larger than 6 m/s^2 that is considered as unacceptable for the serviceability of footbridges [38, 39]. Thus, this example considers the robust design of a TMD as a strategy to mitigate bridge responses. In this problem, the TMD is designed to be placed at the location of maximum vertical displacement in the dominant bending mode, as illustrated in Fig. 8.

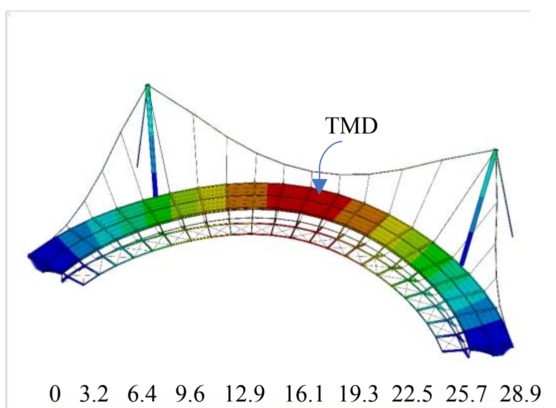


Fig. 8 The dynamics parameter and analysis of the bridge: the first bending mode (unit: cm) and the location the TMD

When a TMD is introduced as an energy absorber in the bridge system, the following coupling equation can be formulated to describe the response of the system, considering also the mode h of the footbridge:

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) + \boldsymbol{C}\dot{\boldsymbol{u}}(t) + \boldsymbol{K}\boldsymbol{u}(t) = \boldsymbol{F}(t) \tag{32}$$

where $\boldsymbol{M} = \begin{bmatrix} m_h & 0 \\ 0 & m_d \end{bmatrix}$, $\boldsymbol{C} = \begin{bmatrix} c_h + c_d & -c_d \\ -c_d & c_d \end{bmatrix}$, $\boldsymbol{K} = \begin{bmatrix} k_h + k_d & -k_d \\ -k_d & k_d \end{bmatrix}$, $\boldsymbol{F}(t) = \begin{bmatrix} p(t) \\ 0 \end{bmatrix}$ and $\boldsymbol{u}(t) = \begin{bmatrix} u_h \\ u_d \end{bmatrix}$. m_h, k_h and c_h denote the equivalent modal mass, stiffness and damping to unitize $\max|\phi_h|$, respectively. m_d, c_d and k_d represent the mass, damping and stiffness of the TMD device, respectively. $p(t)$ denotes the force generated by the pedestrian walk, which is formulated as:

$$p(t) = p_s \cos(2\pi f_s) \tag{33}$$

where f_s denotes the step frequency that is considered as f_h in the above equations, p_s represents the ground reaction force that is considered as $\frac{N_{eq}\alpha_{e,h}G}{S}\sum_{k=1}^{n_{eff}}\frac{a_{eff,k}|\phi_{h,k}|}{\max|\phi_h|}$.

Considering various sources of uncertainties, such as the bridge natural frequency, damping ratio and the frequency of pedestrian excitations, that could influence the bridge state and the external loading, this paper formulates the TMD design as a robust optimization problem, to minimize the weight of the TMD while maintaining the feasibility robustness in an uncertain environment. The problem for robust TMD design of the cable-supported bridge is formulated as:

$$\min_{(\boldsymbol{d})} m_d$$
$$s.t.: \begin{cases} \mathbb{E}[\ddot{u}_{h,max}(\boldsymbol{d}_d,\boldsymbol{\xi})] + 3\mathbb{S}[\ddot{u}_{h,max}(\boldsymbol{d}_d,\boldsymbol{\xi})] \le 0.75 \\ \boldsymbol{d}_d = (m_d, c_d, k_d) \\ \boldsymbol{\xi} = [\xi_c, \xi_f, \xi_w] \\ 0.01 \le m_d/m_h \le 0.11 \end{cases} \quad (34)$$

where $\boldsymbol{d}_d$ denotes the vector of deterministic design variables, i.e., the mass m_d, stiffness c_d and damping k_d of the TMD (the cost of the TMD typically depends on its mass and therefore m_d is taken as the objective of optimization). $\boldsymbol{\xi}$ denotes the vector of random variables, including the modal damping for the first bending mode of the bridge (ξ_c), the natural frequency for the first bending mode of the bridge (ξ_f) and the frequency for the walking loading (ξ_w). Table 5 lists the assumed distribution of random variables.

Table 5 The distribution of random variables of Example Four.

Variables	Distribution	Mean	SD	Unit
Modal damping (ξ_c)	Uniform	0.4	0.115	%
Natural frequency for the first bending mode (ξ_f)	Uniform	2.0	0.115	Hz
Frequency of the walking loading (ξ_w)	Normal	2.0	0.2	Hz

The proposed method is applied to solve this robust TMD design optimization problem. Some other surrogate model-based approaches, such as the PCE-LAR, SVR [28] and RBFNN [40], are compared to illustrate the advantages of the proposed method. To facilitate the determination of the optimal TMD parameters, based on the suggestion from [41], the optimal stiffness and damping are determined based on the design of mass, therefore the dimension of the design variables can be reduced. Table 6 illustrates the results from different methodologies. It is noticed that the PCE-LAR, SVR and RBFNN all fail to provide acceptable solutions in this example of TMD design optimization. An obvious discrepancy for the feasibility robustness index, defined by $\mathbb{E}(\ddot{u}_{h,max}) + 3\mathbb{S}(\ddot{u}_{h,max})$, can be observed for these methods. For instance, the SVR-based RDO can yield a relative error of nearly 59%. In comparison, the proposed methods with PCK and ordinary Kriging both yield much more accurate results for the TMD design. The relative error for the feasibility robustness from the PCK-based analytical RDO method is only 1% with 75.2 calls of the performance function on average, while the relative error from the ordinary Kriging-based approach is slightly larger (2%) with 68.5 calls of the performance function. Therefore, the proposed analytical robust design optimization method provides the best performance among all considered approaches. Furthermore, because the characteristics, e.g., mass, stiffness and damping, of a TMD are typically very preciously validated by experiments before the product launch, this example considers these parameters as deterministic design variables while taking into account some other essential uncertainty sources, such as the bridge modal frequency, damping and the external loading. The proposed method is the only approach that can provide analytical solution of robustness in such application background. This further validates the contribution of proposed method for extending the application scope of surrogate model-based analytical RDO to various practical engineering problems.

Table 6 Comparison of the results from the proposed and other methods for Example Four.

Method	Optimal m_t (kg)[a]	$\mathbb{E}(\ddot{u}_{h,max}) + 3\mathbb{S}(\ddot{u}_{h,max})$ [a]	$\lvert\Delta_{con}\rvert(\%)$	Sample Size	Analytical Solution
MCS	3522.38	0.750	—	$10^4 \times 770$ [b]	No
RBFNN	2565.57	0.969	29.24	78	No
SVR	1528.44	1.192	58.94	78	No
PCE-LAR	1217.12	1.278	70.39	78	No
Ordinary Kriging	2142.83	1.049	39.92	78	No
The proposed method-Ordinary Kriging	3660.29	0.735	2.01	68.5	Yes
The proposed method-PCK	3593.31	0.743	0.99	75.2	Yes

5. Conclusion

As an important branch of structural optimization, robust design optimization aims at optimizing the structural design while maintaining objective or feasibility robustness. Surrogate model-based analytical robust design optimization opens up a promising avenue to not only reduce the calls of complex computational models by establishing substitutes but also eliminate the repeated sampling estimation for statistical moments during the optimization routine. Motivated by extending the applicable scope of existing analytical method to handle both deterministic and random variables, this paper proposed a Polynomial Chaos Kriging-based methodology for efficient analytical robust design optimization. This work derived the analytical formulas of the statistical moments based on the underlying assumption of PCK surrogate model established on the augmented space. A symmetric investigation was carried out for uniform and Gaussian distributions. The paper also presented an active-learning framework consisting of three stages of initialization, optimization and refinement. Different types of numerical optimization methods, such as gradient-based methods or evolutionary methods, can be seamlessly implemented in the framework in tandem with the adaptively established surrogate model and the proposed analytical robustness formulas.

Several classical numerical examples demonstrated that the proposed analytical formula can be much more efficient compared with simulation methods. Three numerical examples and a practical application assessed the performance of the proposed method. It was noticed that the proposed method can well handle different types of problems (multimodal problems, or discrete problems). Furthermore, the PCK surrogate model generally performed slightly better than the Kriging surrogate model, possibly because of the improvement of the ability to capture the global trend.

The paper is concluded by discussing some limitations and possible extensions of the proposed RDO method. This work adopts the PC-Kriging, which introduces the PCE into the interpolation-type Kriging, as the surrogate model. Compared with the traditional sparse PCE, PCK improves the approximation ability to capture local variations of responses; however, PCK also suffers from the computational burden for inversion of the covariance matrix. Therefore, the proposed method is not suitable for very high-dimensional problems. Future works will investigate incorporation of dimension reduction techniques into the proposed framework. As reduction of the design space also alleviates the computational burden of model training, an adaptive decomposition framework could be investigated to enhance the method's performance.

[a] $\mathbb{E}(\ddot{u}_{h,max}) + 3\mathbb{S}(\ddot{u}_{h,max})$ is the average value over 10 independent runs.
[b] the sample size for MCS indicates (the number of realizations to estimate the statistics) × (the number of evaluations of the robustness index during the optimization loop).

CRediT authorship contribution statement

Chaolin Song: Conceptualization, Methodology, Formal Analysis, Writing - original draft. **Abdollah Shafieezadeh**: Conceptualization, Supervision, Methodology, Writing - review & editing. **Rucheng Xiao**: Conceptualization, Supervision. **Bin Sun**: Conceptualization, Writing - review & editing.

Acknowledgements

The authors appreciate the financial support from the National Natural Science Foundation of China (52308196), and the U.S. National Science Foundation (NSF) through awards CMMI-2000156. The first author would like to thank the support from the Postdoctoral Fellowship Program of CPSF under Grant Number GZB20230528 and 2023M742664. The opinions and statements do not necessarily represent those of the sponsors. In addition, the authors would like to express their gratitude to the Research Division on Structural Health Monitoring and Vibration Control, Department of Bridge Engineering, Tongji University and Mr. Zuqian Jiang, for the information about the bridge and the help in the modeling.

Appendix A

In this Appendix, following the derivations in Section 3.1 and 3.2, the statistical moments estimated by PCK with $\boldsymbol{\xi}$ following independent Gaussian distribution is provided. It is noticed that the distribution of random variables mainly influences Eq. (13) and Eq. (15) when computing the first moment, and influences Eq. (19), Eq. (22) and Eq. (25) when computing the second moment.

For an arbitrary univariate orthogonal polynomial, the function, i.e., $\psi_{\alpha}(\xi_k)$, can be typically expressed as the consolidation of several power functions. For instance, a second-order probabilistic Hermite polynomial is expressed as $\psi_{\alpha=2}(\xi_k) = \sqrt{2}[b_{\hbar=2}(\xi_k) - b_{\hbar=0}(\xi_k)]/2$. When the random variable ξ_k follow Gaussian distribution $\mathcal{N}[\xi_k|\mu_k,(\sigma_k)^2]$ where $k = 1,\cdots,dim_{\xi}$ denotes the dimension number, and let $b_{\hbar}(x) = x^{\hbar}$ denote an $\hbar$-order function. It is noticed that the following equation can be obtained:

$$\mathbb{E}[b_{\hbar}(\xi_k)] = \int b_{\hbar}(\xi_k)\cdot f(\xi_k|\boldsymbol{d}_{\mu})\,\mathrm{d}\xi_k = \int b_{\hbar}(\xi_k)\cdot \mathcal{N}[\xi_k|\mu_k,(\sigma_k)^2]\mathrm{d}\xi_k \tag{A.1}$$

The following recursion table can be then obtained.

Table A.1. The analytical formula of integral with Gaussian distribution

Integration Function	Analytical Expression
$\int f(\xi_k\|\boldsymbol{d}_{\mu})\cdot b_{\hbar=0}(\xi_k)\mathrm{d}\xi_k$	1
$\int f(\xi_k\|\boldsymbol{d}_{\mu})\cdot b_{\hbar=1}(\xi_k)\mathrm{d}\xi_k$	μ_k
⋮	⋮
$\int f(\xi_k\|\boldsymbol{d}_{\mu})\cdot b_{\hbar=n}(\xi_k)\mathrm{d}\xi_k$	$(\sigma_k)^2(n-1)\int f(\xi_k\|\boldsymbol{d}_{\mu})\cdot b_{\hbar=n-2}(\xi_k)\mathrm{d}\xi_k + \mu_k\int f(\xi_k\|\boldsymbol{d}_{\mu})\cdot b_{\hbar=n-1}(\xi_k)\mathrm{d}\xi_k$

Thus, Eq. (13), Eq. (19) can be efficiently analytically determined based on the optimized coefficients and corresponding order of the polynomial basis functions. (Note that the raw moments of some distributions have been discussed in the literature [42]; the above formulas are not all new.)

For Eq. (15), when random variables are assumed to follow independent Gaussian distributions, referring to Appendix C, the following equation can be recast:

$$\mathbb{E}\left(\exp\left[-\frac{1}{2}\left(\frac{\xi_k-\xi_k^{(i)}}{\theta_{\xi,k}^*}\right)^2\right]\right) = \int \sqrt{2\pi}\theta_{\xi,k}^*\cdot f(\xi_k|\boldsymbol{d}_{\mu})\cdot\mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta_{\xi,k}^*\right)^2\right]\mathrm{d}\xi_k \tag{A.2}$$

$$= \int \sqrt{2\pi}\theta^*_{\xi,k} \cdot \mathcal{N}[\xi_k|\mu_k,(\sigma_k)^2] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k$$
$$= \sqrt{2\pi}\theta^*_{\xi,k} \cdot s_{n1}$$

where s_{n1} can be determined by:

$$s_{n1} = \frac{1}{\sqrt{2\pi\left[(\sigma_k)^2 + \left(\theta^*_{\xi,k}\right)^2\right]}} \exp\left[-\frac{(\xi_k^{(i)} - \mu_k)^2}{2\left[(\sigma_k)^2 + \left(\theta^*_{\xi,k}\right)^2\right]}\right] \tag{A.3}$$

Therefore, Eq. (15) can be efficiently analytically determined.

Furthermore, Eq. (22) can be recast as:

$$\begin{aligned}\mathbb{E}\left[\psi_\alpha(\xi_k) r\left(\theta^*_{\xi,k},\xi_k,\xi_k^{(i)}\right)\right] &= \sqrt{2\pi}\theta^*_{\xi,k} \int f(\xi_k|\boldsymbol{d}_\mu) \cdot \psi_\alpha(\xi_k) \cdot \mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \\ &= \sqrt{2\pi}\theta^*_{\xi,k} \int \psi_\alpha(\xi_k) \cdot \mathcal{N}[\xi_k|\mu_k,(\sigma_k)^2] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \\ &= \sqrt{2\pi}\theta^*_{\xi,k} s_{n1} \int \psi_\alpha(\xi_k) \cdot \mathcal{N}\left[\xi_k|\xi_k^{(n1)},\left(\sigma_k^{(n1)}\right)^2\right] \mathrm{d}\xi_k\end{aligned} \tag{A.4}$$

where $\xi_k^{(n1)}$ and $\sigma_k^{(n1)}$ can be determined by:

$$\xi_k^{(n1)} = \frac{\xi_k^{(i)}(\sigma_k)^2 + \mu_k\left(\theta^*_{\xi,k}\right)^2}{(\sigma_k)^2 + \left(\theta^*_{\xi,k}\right)^2} \tag{A.5}$$

$$\sigma_k^{(n1)} = \sqrt{\frac{(\sigma_k)^2\left(\theta^*_{\xi,k}\right)^2}{(\sigma_k)^2 + \left(\theta^*_{\xi,k}\right)^2}} \tag{A.6}$$

Subsequently, the Table A.1 can be applied again based on the optimized coefficients and corresponding order of the polynomial basis functions to analytically calculate Eq. (22).

For the last part required when calculating the second moment of statistics, Eq. (25) can be formulated as:

$$\begin{aligned}\mathbb{E}\left[r\left(\theta^*_{\xi,k},\xi_k,\xi_k^{(i)}\right) r\left(\theta^*_{\xi,k},\xi_k,\xi_k^{(t)}\right)\right] &= \int 2\pi\left(\theta^*_{\xi,k}\right)^2 \cdot f(\xi_k|\boldsymbol{d}_\mu) \cdot \mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta^*_{\xi,k}\right)^2\right] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(t)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \\ &= \int 2\pi\left(\theta^*_{\xi,k}\right)^2 \cdot \mathcal{N}[\xi_k|\mu_k,(\sigma_k)^2] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(i)},\left(\theta^*_{\xi,k}\right)^2\right] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(t)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \\ &= \int 2\pi\left(\theta^*_{\xi,k}\right)^2 \cdot s_{n1} \cdot \mathcal{N}\left[\xi_k|\xi_k^{(n1)},\left(\sigma_k^{(n1)}\right)^2\right] \cdot \mathcal{N}\left[\xi_k|\xi_k^{(t)},\left(\theta^*_{\xi,k}\right)^2\right] \mathrm{d}\xi_k \\ &= \int 2\pi\left(\theta^*_{\xi,k}\right)^2 \cdot s_{n1} \cdot s_{n2} \cdot \mathcal{N}\left[\xi_k|\xi_k^{(n2)},\left(\sigma_k^{(n2)}\right)^2\right] \mathrm{d}\xi_k = 2\pi\left(\theta^*_{\xi,k}\right)^2 \cdot s_{n1} \cdot s_{n2}\end{aligned} \tag{A.7}$$

where s_{n2}, ξ_{n2}^k and σ_{n2}^k can be similarly determined referring to Appendix C.

$$s_{n2} = \frac{1}{\sqrt{2\pi\left[\left(\sigma_k^{(n1)}\right)^2 + \left(\theta^*_{\xi,k}\right)^2\right]}} \exp\left[-\frac{\left(\xi_k^{(t)} - \xi_k^{(n1)}\right)^2}{2\left[\left(\sigma_k^{(n1)}\right)^2 + \left(\theta^*_{\xi,k}\right)^2\right]}\right] \tag{A.8}$$

$$\xi_k^{(n2)} = \frac{\xi_k^{(t)}\left(\sigma_k^{(n1)}\right)^2 + \xi_k^{(n1)}\left(\theta^*_{\xi,k}\right)^2}{\left(\sigma_k^{(n1)}\right)^2 + \left(\theta^*_{\xi,k}\right)^2} \tag{A.9}$$

$$\sigma_k^{(n2)} = \sqrt{\frac{\left(\sigma_k^{(n1)}\right)^2\left(\theta^*_{\xi,k}\right)^2}{\left(\sigma_k^{(n1)}\right)^2 + \left(\theta^*_{\xi,k}\right)^2}} \tag{A.10}$$

Based on the above formula and Table A.1, the integrals required in the formulas of the statistical moments have also been analytically expressed for the case of independent Gaussian distributions.

Appendix B

In this Appendix, the statistical moments estimated by PCK with $\boldsymbol{\xi}$ following independent uniform distribution is provided. Throughout this section, it is generally assumed that a random variable ξ_k follows a uniform distribution between $\left[\xi_k^{(u)}, \xi_k^{(l)}\right]$ ($\left[\xi_k^{(u)}, \xi_k^{(l)}\right]$ can depend on $\boldsymbol{d}_\mu$). As a reminder, note that the distribution of random variables mainly influences Eq. (13) and Eq. (15) when computing the first moment, and influences Eq. (19), Eq. (22) and Eq. (25) when computing the second moment.

Following the assumption in the above derivations, Let $b_\hbar(x) = x^\hbar$ denote an $\hbar$ order function. For these conditions, the following equation can be obtained:

$$\int f(\xi_k|\boldsymbol{d}_\mu) \cdot b_\hbar(\xi_k) \mathrm{d}\xi_k = \frac{1}{\xi_k^{(u)} - \xi_k^{(l)}} \frac{1}{\hbar + 1} \left[\left(\xi_k^{(u)}\right)^{h+1} - \left(\xi_k^{(l)}\right)^{h+1} \right] \tag{B.1}$$

The integral required in Eq. (13) and (19) can be therefore analytically calculated combining the integrals for all $b_\hbar(x)$ components that are determined based on the type and degree of the polynomial basis.

Additionally, with the uniform distribution, Eq. (15) can be recast as:

$$\begin{aligned}
\mathbb{E}\left(\exp\left[-\frac{1}{2} \left(\frac{\xi_k - \xi_k^{(i)}}{\theta_{\xi,k}^*} \right)^2 \right] \right) &= \frac{\sqrt{2\pi}\theta_{\xi,k}^*}{\xi_k^{(u)} - \xi_k^{(l)}} \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} \mathcal{N}\left[\xi_k | \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2 \right] \mathrm{d}\xi_k \\
&= \frac{\sqrt{2\pi}\theta_{\xi,k}^*}{\xi_k^{(u)} - \xi_k^{(l)}} \left[\Phi\left(\frac{\xi_k^{(u)} - \xi_k^{(i)}}{\theta_{\xi,k}^*} \right) - \Phi\left(\frac{\xi_k^{(l)} - \xi_k^{(i)}}{\theta_{\xi,k}^*} \right) \right]
\end{aligned} \tag{B.2}$$

where $\Phi(\cdot)$ denotes the Gaussian cumulative density function.

For the analytical formula of Eq. (22), the following relationship can be observed:

$$\int f(\xi_k|\boldsymbol{d}_\mu) \cdot b_\hbar(\xi_k) \cdot \mathcal{N}\left[\xi_k | \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2 \right] \mathrm{d}\xi_k = \frac{1}{\sqrt{2\pi}\theta_{\xi,k}^*} \frac{1}{\xi_k^{(u)} - \xi_k^{(l)}} \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^\hbar \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\xi_k \tag{B.3}$$

$$\begin{aligned}
&\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^\hbar \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\xi_k \\
&= \left(\theta_{\xi,k}^*\right)^2 \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^{\hbar-1} \cdot \frac{\left(\xi_k - \xi_k^{(i)}\right)}{\left(\theta_{\xi,k}^*\right)^2} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\xi_k + \xi_k^{(i)} \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^{\hbar-1} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\xi_k \\
&= \left(\theta_{\xi,k}^*\right)^2 \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^{\hbar-1} \mathrm{d}\left(-\exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \right) + \xi_k^{(i)} \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^{\hbar-1} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\xi_k
\end{aligned} \tag{B.4}$$

The method of integration by parts is applied herein:

$$\begin{aligned}
&\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} \left(\xi_k\right)^{\hbar-1} \mathrm{d}\left(-\exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \right) \\
&= \int_{\xi_k^{(l)}}^{\xi_k^{(u)}} \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \mathrm{d}\left((\xi_k)^{\hbar-1}\right) - \left[(\xi_k)^{\hbar-1} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2} \right] \right]_{\xi_k^{(l)}}^{\xi_k^{(u)}}
\end{aligned} \tag{B.5}$$

$$= (\hbar - 1)\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} (\xi_k)^{\hbar-2} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2}\right] \mathrm{d}\xi_k - \left[(\xi_k)^{\hbar-1} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2}\right]\right]_{\xi_k^{(l)}}^{\xi_k^{(u)}}$$

Based on these derivations, Table B.1 presents the analytical solutions of the integral terms that appear in Eq. (22).

Table B.1. The analytical formula of integral with uniform distribution

Integration Function	Analytical Expression
$\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} f(\xi_k \mid \boldsymbol{d}_\mu) \cdot b_{\hbar=0}(\xi_k) \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k$	$\frac{1}{\xi_k^{(u)} - \xi_k^{(l)}}\left[\Phi\left(\frac{\xi_k^{(u)} - \xi_k^{(i)}}{\theta_{\xi,k}^*}\right) - \Phi\left(\frac{\xi_k^{(l)} - \xi_k^{(i)}}{\theta_{\xi,k}^*}\right)\right]$
$\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} f(\xi_k \mid \boldsymbol{d}_\mu) \cdot b_{\hbar=1}(\xi_k) \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k$	$\frac{\xi_k^{(i)}}{\xi_k^{(u)} - \xi_k^{(l)}}\left[\Phi\left(\frac{\xi_k^{(u)} - \xi_k^{(i)}}{\theta_{\xi,k}^*}\right) - \Phi\left(\frac{\xi_k^{(l)} - \xi_k^{(i)}}{\theta_{\xi,k}^*}\right)\right] - \frac{1}{\xi_k^{(u)} - \xi_k^{(l)}}\frac{\theta_{\xi,k}^*}{\sqrt{2\pi}}\left[\exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2}\right]\right]_{\xi_k^{(l)}}^{\xi_k^{(u)}}$
⋮	⋮
$\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} f(\xi_k \mid \boldsymbol{d}_\mu) \cdot b_{\hbar=n}(\xi_k) \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k$	$\left(\theta_{\xi,k}^*\right)^2(\hbar - 1)\int f(\xi_k \mid \boldsymbol{d}_\mu) \cdot b_{\hbar=n-2}(\xi_k) \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k + \xi_k^{(i)}\int f(\xi_k \mid \boldsymbol{d}_\mu) \cdot b_{\hbar=n-1}(\xi_k) \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k - \frac{1}{\xi_k^{(u)} - \xi_k^{(l)}}\frac{\theta_{\xi,k}^*}{\sqrt{2\pi}}\left[(\xi_k)^{n-1} \cdot \exp\left[-\frac{\left(\xi_k - \xi_k^{(i)}\right)^2}{2\left(\theta_{\xi,k}^*\right)^2}\right]\right]_{\xi_k^{(l)}}^{\xi_k^{(u)}}$

Furthermore, for Eq. (25), When the random variable ξ_k follows a uniform distribution between $\left[\xi_k^{(u)}, \xi_k^{(l)}\right]$, and the above equation can be recast as:

$$\mathbb{E}\left[r\left(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)}\right) r\left(\theta_{\xi,k}^*, \xi_k, \xi_k^{(t)}\right)\right] = \frac{2\pi\left(\theta_{\xi,k}^*\right)^2}{\xi_k^{(u)} - \xi_k^{(l)}}\int_{\xi_k^{(l)}}^{\xi_k^{(u)}} \mathcal{N}\left[\xi_k \mid \xi_k^{(i)}, \left(\theta_{\xi,k}^*\right)^2\right] \cdot \mathcal{N}\left[\xi_k \mid \xi_k^{(t)}, \left(\theta_{\xi,k}^*\right)^2\right] \mathrm{d}\xi_k \tag{B.6}$$

Referring to Appendix C again, the product of two normal distribution also follows a normal distribution. The following equations can be obtained:

$$\mu_{u,k} = \frac{\xi_k^{(i)} + \xi_k^{(t)}}{2} \tag{B.7}$$

$$\sigma_{u,k} = \sqrt{\frac{\left(\theta_{\xi,k}^*\right)^2}{2}} \tag{B.8}$$

$$s_{u1} = \frac{1}{\sqrt{4\pi\left(\theta_{\xi,k}^*\right)^2}} \exp\left[-\frac{\left(\xi_k^{(t)} - \xi_k^{(i)}\right)^2}{4\left(\theta_{\xi,k}^*\right)^2}\right] \tag{B.9}$$

$$\mathbb{E}\left[r\left(\theta_{\xi,k}^*, \xi_k, \xi_k^{(i)}\right) r\left(\theta_{\xi,k}^*, \xi_k, \xi_k^{(t)}\right)\right] = \frac{2\pi\left(\theta_{\xi,k}^*\right)^2}{\xi_k^{(u)} - \xi_k^{(l)}} s_{u1}\left[\Phi\left(\frac{\xi_k^{(u)} - \mu_{u,k}}{\sigma_{u,k}}\right) - \Phi\left(\frac{\xi_k^{(l)} - \mu_{u,k}}{\sigma_{u,k}}\right)\right] \tag{B.10}$$

Based on the above formulas and Table B.1, the integrals required in the formulas of the statistical moments have also been analytically expressed for the case of independent uniform distributions.

Appendix C

Let $\mathcal{N}\left(\xi\middle|\xi_a, {\theta_a}^2\right)$ denote a Gaussian probability density function with mean of ξ_a and standard deviation of θ_a; $\mathcal{N}\left(\xi\middle|\xi_b, {\theta_b}^2\right)$ denotes a Gaussian probability density function with mean of ξ_b and standard deviation of θ_b. Let $\xi_n = \frac{\xi_b {\theta_a}^2 + \xi_a {\theta_b}^2}{{\theta_a}^2 + {\theta_b}^2}$ and $\theta_n = \sqrt{\frac{{\theta_a}^2 {\theta_b}^2}{{\theta_a}^2 + {\theta_b}^2}}$. It has been proven that the product of $\mathcal{N}\left(\xi\middle|\xi_a, {\theta_a}^2\right)$ and $\mathcal{N}\left(\xi\middle|\xi_b, {\theta_b}^2\right)$ follows a scaled Gaussian distribution:

$$\mathcal{N}\left(\xi\middle|\xi_a, {\theta_a}^2\right) \cdot \mathcal{N}\left(\xi\middle|\xi_b, {\theta_b}^2\right) = s_g \cdot \mathcal{N}\left(\xi\middle|\xi_n, {\theta_n}^2\right) \tag{C.1}$$

where s_g denotes the scaled factor that can be formulated as:

$$s_g = \frac{1}{\sqrt{2\pi({\theta_a}^2 + {\theta_b}^2)}} \exp\left[-\frac{(\xi_a - \xi_b)^2}{2({\theta_a}^2 + {\theta_b}^2)}\right] \tag{C.2}$$

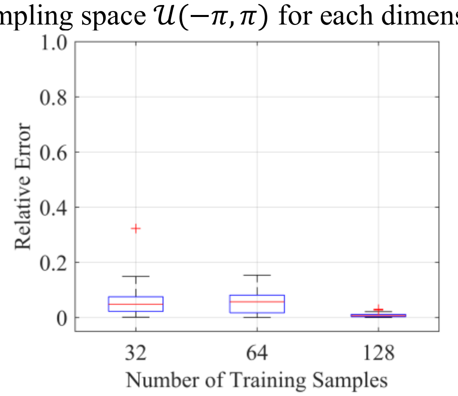
Appendix D

To validate the accuracy of the analytical formula for obtaining the statistics of the performance function and to compare the PCK and Kriging surrogate model, five examples are investigated here. The analytical expressions derived in Section 3.1 and 3.2 are applied. For a fair comparison, 20 independent runs are executed for obtaining the mean relative error of the estimations. The UQLab toolbox (version 2.0.0) is adopted for creating the surrogate model where HGA is selected as the optimization method for both ordinary Kriging and PCK.

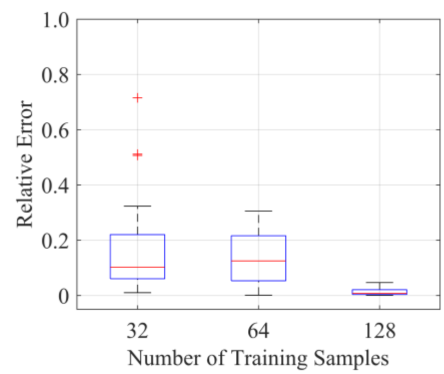
The first example is the so-called Ishigami function with three uniformly distributed random variables as the inputs. The function has the following form:

$$g_1 = \sin(\xi_1) + 7\sin^2(\xi_2) + 0.1x_3^4 \sin(\xi_1) \tag{D.1}$$

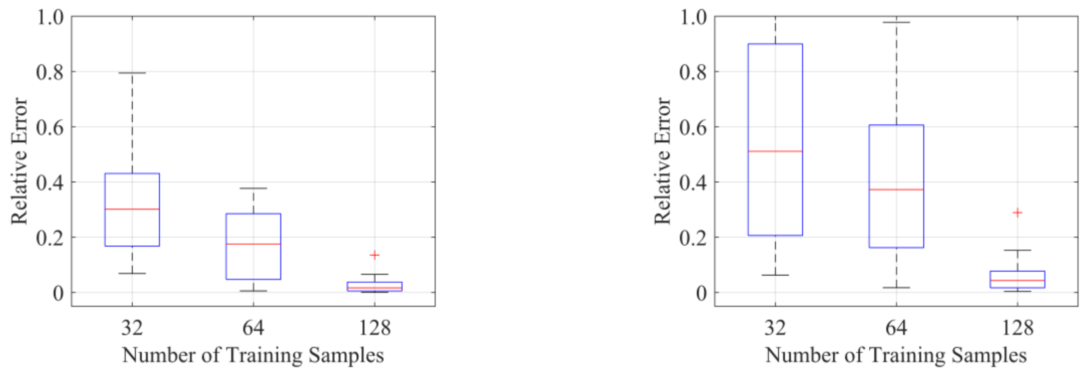
where $\xi_i \sim \mathcal{U}(1,2)$, $i = 1,2,3$. Fig. D.1 shows the analytical prediction of the statistics of the output for different number of training samples from 32 to 128. The training samples are uniformly generated from the sampling space $\mathcal{U}(-\pi, \pi)$ for each dimension.



(a) Relative error for estimating the mean by PCK



(b) Relative error for estimating the mean by ordinary Kriging



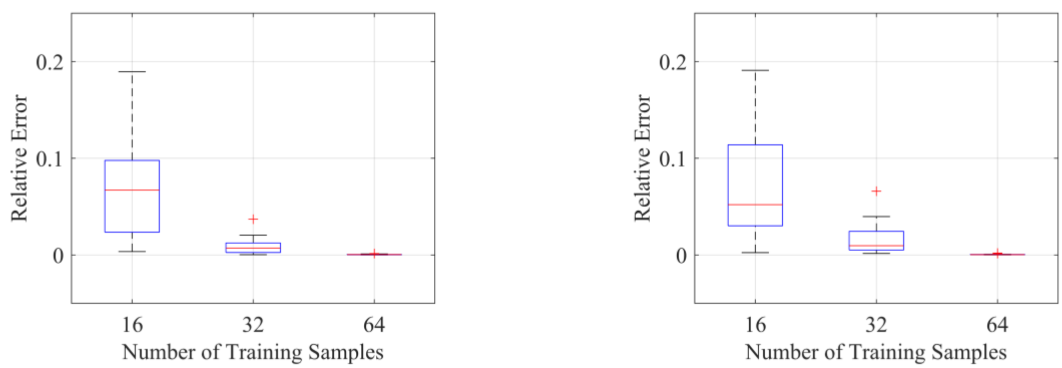
(c) Relative error for estimating SD by PCK (d) Relative error for estimating SD by ordinary Kriging

Fig. D.1 Function One: Relative error for estimating the first two moments by PCK and ordinary Kriging.

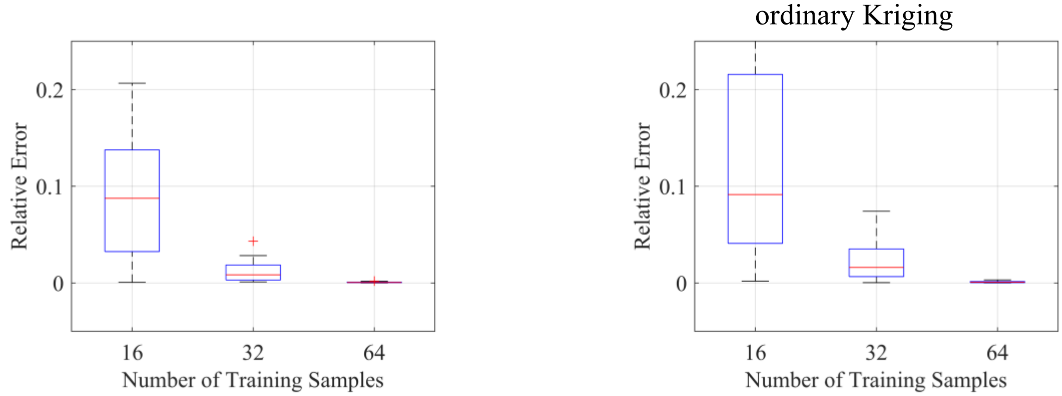
Function two is the so-called Rosenbrock function with two uniformly distributed random variables as the input. The function reads:

$$g_2 = 100(\xi_2 - \xi_1^2)^2 + (1 - \xi_1)^2 \tag{D.2}$$

where $\xi_1 \sim \mathcal{U}(1,2)$ and $\xi_2 \sim \mathcal{U}(0,1)$. Fig. D.2 shows the analytical prediction of the statistics of the output for different numbers of training samples ranging from 16 to 64. The training samples are uniformly generated from the sample space $\mathcal{U}(-2,2)$ for each dimension.



(a) Relative error for estimating the mean by PCK (b) Relative error for estimating the mean by ordinary Kriging



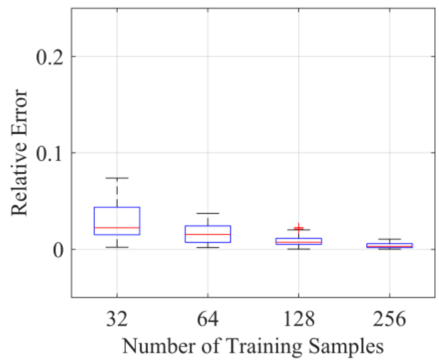
(c) Relative error for estimating SD by PCK (d) Relative error for estimating SD by ordinary Kriging

Fig. D.2 Function Two: Relative error for estimating the first two moments by PCK and ordinary Kriging.

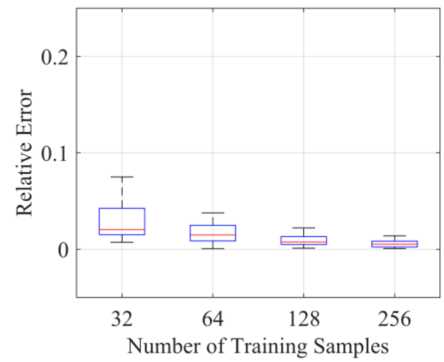
Function three is the modified Sobol function with four uniformly distributed random variables as the input. The function reads:

$$g_3 = \prod_{i=1}^{4} \frac{|4\xi_i - 2| + c_i}{1 + c_i} \tag{D.3}$$

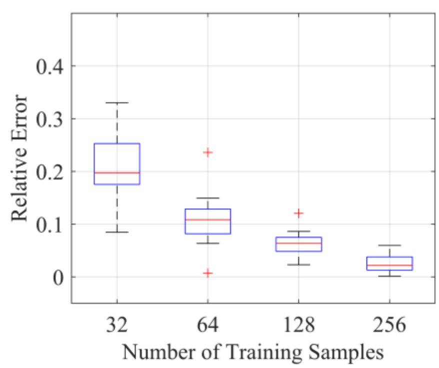
where $\xi_i \sim \mathcal{U}(0,1)$, $i = 1,2,3,4$, $c = (1,2,5,10)$. Fig. D.3 shows the analytical prediction of the statistics of the output, by changing the number of training samples from 32 to 256. The training samples are uniformly generated from the sampling space $\mathcal{U}(0,1)$ for each dimension.



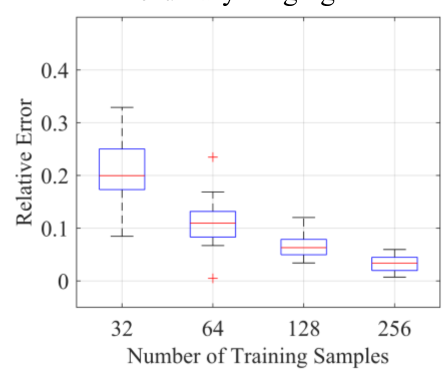
(a) Relative error for estimating the mean by PCK



(b) Relative error for estimating the mean by ordinary Kriging



(c) Relative error for estimating SD by PCK



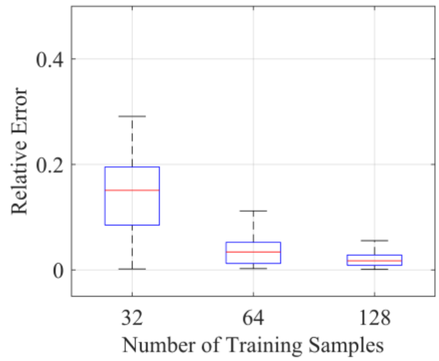
(d) Relative error for estimating SD by ordinary Kriging

Fig. D.3 Function Three: Relative error for estimating the first two moments by PCK and ordinary Kriging.

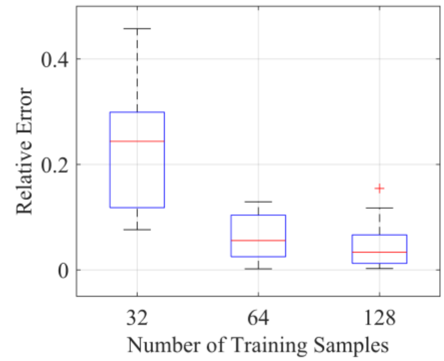
Function four is a four-dimensional function with four normally distributed random variables as the input. The function reads:

$$g_4 = 2/3 \exp(\xi_1^2 + \xi_2^2) + \xi_4 \cos(\xi_3) + \xi_3 \tag{D.4}$$

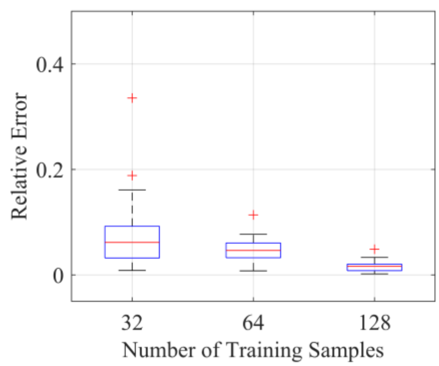
where $\xi_i \sim \mathcal{N}(0.1, 0.1^2)$, $i = 1,2,3,4$. Fig. D.4 shows the analytical prediction of the statistics of the output, by changing the number of training samples from 32 to 128. The training samples are uniformly generated from the sampling space $\mathcal{U}(-1,1)$ for each dimension.



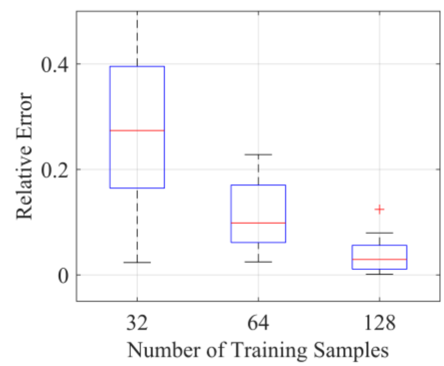
(a) Relative error for estimating the mean by PCK



(b) Relative error for estimating the mean by ordinary Kriging



(c) Relative error for estimating SD by PCK



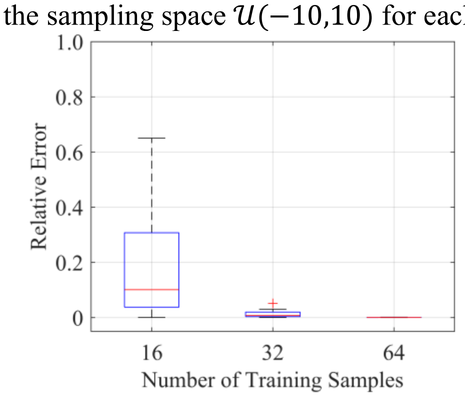
(d) Relative error for estimating SD by ordinary Kriging

Fig. D.4 Function Four: Relative error for estimating the first two moments by PCK and ordinary Kriging.

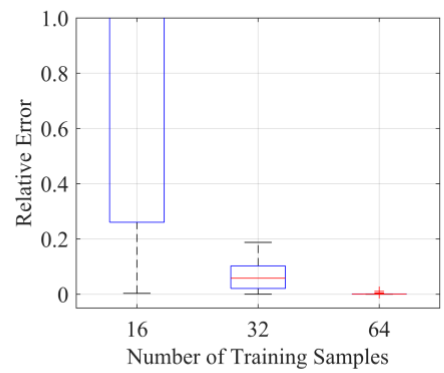
Function five is a two-dimensional function with two normally distributed random variables as the input. The function reads:

$$g_5 = \xi_1 + 5\sin(\xi_1) + 0.1\xi_1\xi_2^2 \tag{D.5}$$

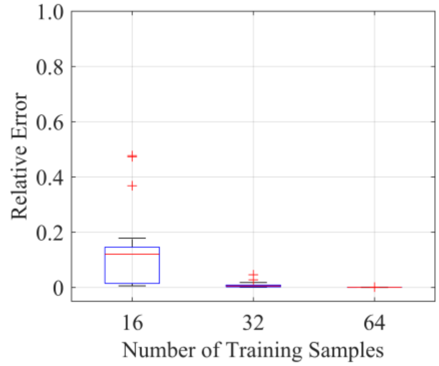
where $\xi_i \sim \mathcal{N}(1,1)$, $i = 1,2$. Fig. D.5 shows the analytical prediction of the statistics of the output, by changing the number of training samples from 16 to 64. The training samples are uniformly generated from the sampling space $\mathcal{U}(-10,10)$ for each dimension.



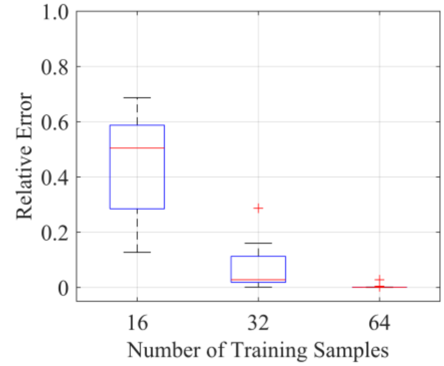
(a) Relative error for estimating the mean by PCK



(b) Relative error for estimating the mean by ordinary Kriging



(c) Relative error for estimating SD by PCK

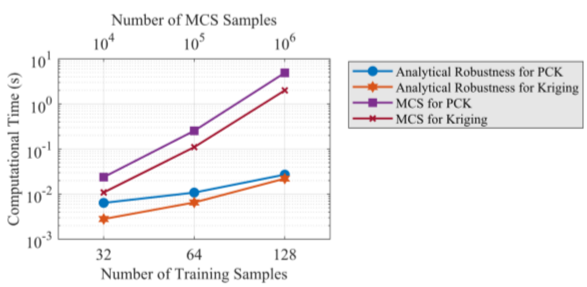


(d) Relative error for estimating SD by ordinary Kriging

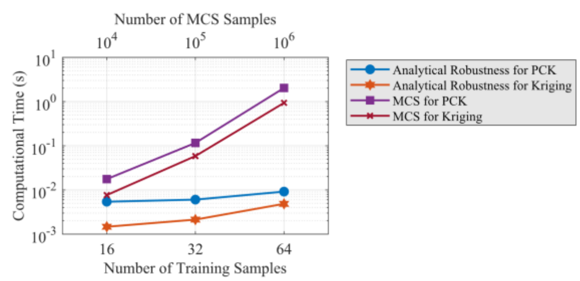
Fig. D.5 Function Five: Relative error for estimating the first two moments by PCK and ordinary Kriging.

Appendix E

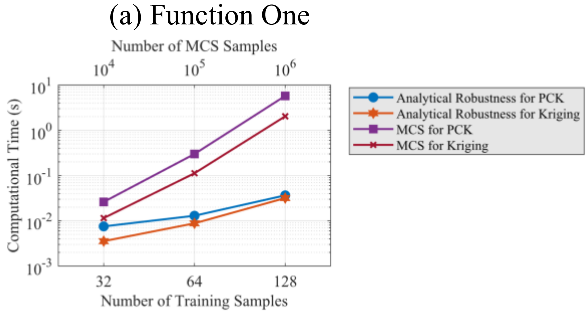
To highlight the motivation for formulating the analytical expression for robustness index, this appendix further compares the computational burden for calculating the analytical equations and implementing MCS. The same functions presented in Appendix A are selected for investigation. The results are averaged on 20 runs. The UQLab toolbox (version 2.0.0) is adopted for the establishment of the surrogate model; HGA is selected as the optimization method for both ordinary Kriging and PCK. As an established surrogate model can be used multiple times during the optimization loop, this appendix therefore does not include the computational burden for training the surrogate model, and only compares the consumed computational time for evaluating the first and second statistical moments.



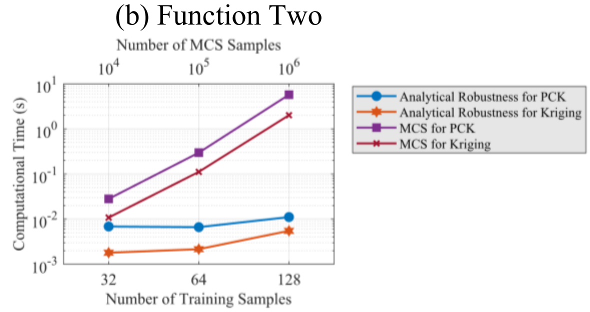
(a) Function One



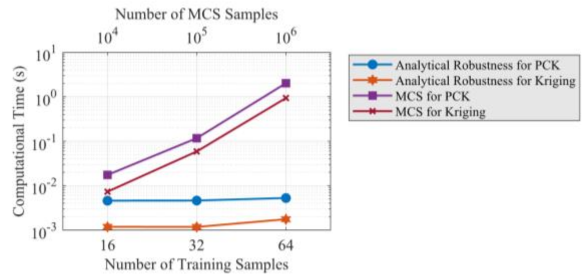
(b) Function Two



(c) Function Three



(d) Function Four



(e) Function Five

Fig. E.1 Comparison between the computational burdens for estimating the first two moments by analytical equations and MCS.

References

[1] Der Kiureghian A, Ditlevsen O. Aleatory or epistemic? Does it matter? Structural safety. 2009;31:105-12.

[2] Paiva RM, Crawford C, Suleman A. Robust and reliability-based design optimization framework for wing design. AIAA journal. 2014;52:711-24.

[3] Beck AT, de Santana Gomes WJ. A comparison of deterministic, reliability-based and risk-based structural optimization under uncertainty. Probabilistic Engineering Mechanics. 2012;28:18-29.

[4] Kanno Y. On three concepts in robust design optimization: absolute robustness, relative robustness, and less variance. Structural and Multidisciplinary Optimization. 2020;62:979-1000.

[5] Taguchi G, Phadke MS. Quality engineering through design optimization. Quality control, robust design, and the Taguchi method: Springer; 1989. p. 77-96.

[6] Beyer H-G, Sendhoff B. Robust optimization–a comprehensive survey. Computer methods in applied mechanics and engineering. 2007;196:3190-218.

[7] Zang C, Friswell M, Mottershead J. A review of robust optimal design and its application in dynamics. Comput Struct. 2005;83:315-26.

[8] Park G-J, Lee T-H, Lee KH, Hwang K-H. Robust design: an overview. AIAA journal. 2006;44:181-91.

[9] Huang B, Du X. Analytical robustness assessment for robust design. Structural and multidisciplinary optimization. 2007;34:123-37.

[10] Li M, Azarm S. Multiobjective collaborative robust optimization with interval uncertainty and interdisciplinary uncertainty propagation. 2008.

[11] Roach T, Kapelan Z, Ledbetter R, Ledbetter M. Comparison of robust optimization and info-gap methods for water resource management under deep uncertainty. Journal of Water Resources Planning and Management. 2016;142:04016028.

[12] Song C, Shafieezadeh A, Xiao R. High-dimensional reliability analysis with error-guided active-learning probabilistic support vector machine: Application to wind-reliability analysis of transmission towers. Journal of Structural Engineering. 2022;148:04022036.

[13] Wang J, Sun Z, Cao R. An efficient and robust Kriging-based method for system reliability analysis. Reliability Engineering & System Safety. 2021;216:107953.

[14] Yang M, Zhang D, Jiang C, Han X, Li Q. A hybrid adaptive Kriging-based single loop approach for complex reliability-based design optimization problems. Reliability Engineering & System Safety. 2021;215:107736.

[15] Jin R, Chen W, Sudjianto A. Analytical metamodel-based global sensitivity analysis and uncertainty propagation for robust design. SAE transactions. 2004:121-8.

[16] Chatterjee T, Chakraborty S, Chowdhury R. A critical review of surrogate assisted robust design optimization. Archives of Computational Methods in Engineering. 2019;26:245-74.

[17] Ren X, Rahman S. Robust design optimization by polynomial dimensional decomposition. Structural and Multidisciplinary Optimization. 2013;48:127-48.

[18] Lee D, Rahman S. Robust design optimization under dependent random variables by a generalized polynomial chaos expansion. Structural and Multidisciplinary Optimization. 2021;63:2425-57.

[19] Chatterjee T, Chakraborty S, Chowdhury R. Analytical moment based approximation for robust design optimization. Structural and Multidisciplinary Optimization. 2018;58:2135-62.

[20] Zhou Y, Lu Z, Cheng K. A new surrogate modeling method combining polynomial chaos expansion and Gaussian kernel in a sparse Bayesian learning framework. International Journal for Numerical Methods in Engineering. 2019;120:498-516.
[21] Liu Y, Zhao G, Li G, He W, Zhong C. Analytical robust design optimization based on a hybrid surrogate model by combining polynomial chaos expansion and Gaussian kernel. Structural and Multidisciplinary Optimization. 2022;65:335.
[22] Schobi R, Sudret B, Wiart J. Polynomial-chaos-based Kriging. International Journal for Uncertainty Quantification. 2015;5.
[23] Song C, Xiao R, Jiang Z, Sun B. Active-learning Kriging-assisted robust design optimization of tuned mass dampers: Vibration mitigation of a steel-arch footbridge. Eng Struct. 2024;303:117502.
[24] Ribaud M, Blanchet-Scalliet C, Helbert C, Gillot F. Robust optimization: a kriging-based multi-objective optimization approach. Reliability Engineering & System Safety. 2020;200:106913.
[25] Lin P, Zhang L, Tiong RL. Multi-objective robust optimization for enhanced safety in large-diameter tunnel construction with interactive and explainable AI. Reliability Engineering & System Safety. 2023;234:109172.
[26] Krige DG. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy. 1951;52:119-39.
[27] Matheron G. The intrinsic random functions and their applications. Advances in applied probability. 1973;5:439-68.
[28] Marelli S, Sudret B. UQLab: A framework for uncertainty quantification in Matlab. Vulnerability, uncertainty, and risk: quantification, mitigation, and management2014. p. 2554-63.
[29] Lophaven SN, Nielsen HB, Søndergaard J. DACE: a Matlab kriging toolbox: Citeseer; 2002.
[30] Xiu D, Karniadakis GE. The Wiener--Askey polynomial chaos for stochastic differential equations. SIAM journal on scientific computing. 2002;24:619-44.
[31] Zhang J, Gong W, Yue X, Shi M, Chen L. Efficient reliability analysis using prediction-oriented active sparse polynomial chaos expansion. Reliability Engineering & System Safety. 2022;228:108749.
[32] Schöbi R, Sudret B. Global sensitivity analysis in the context of imprecise probabilities (p-boxes) using sparse polynomial chaos expansions. Reliability Engineering & System Safety. 2019;187:129-41.
[33] Jia G, Wang Y, Cai Z, Jin Y. An improved (μ+λ)-constrained differential evolution for constrained optimization. Information Sciences. 2013;222:302-22.
[34] Stolpe M. Truss optimization with discrete design variables: a critical review. Structural and Multidisciplinary Optimization. 2016;53:349-74.
[35] Sandgren E, Cameron TM. Robust design optimization of structures through consideration of variation. Comput Struct. 2002;80:1605-13.
[36] Li L. Study on Mechanical Properties of Single Cable Plane Curved Beam Suspension Bridge (In Chinese) [D]: Nanjing Forestry University; 2017.
[37] Lievens K, Lombaert G, De Roeck G, Van den Broeck P. Robust design of a TMD for the vibration serviceability of a footbridge. Eng Struct. 2016;123:408-18.
[38] Heinemeyer C, Feldmann M. European design guide for footbridge vibration. Footbridge vibration design: CRC Press; 2009. p. 13-30.
[39] Ricciardelli F, Demartino C. Design of footbridges against pedestrian-induced vibrations. Journal of Bridge Engineering. 2016;21:C4015003.
[40] Beale MH, Hagan MT, Demuth HB. Neural network toolbox. User's Guide, MathWorks. 2010;2:77-81.
[41] Weber F, Feltrin G, Huth O. Guidelines for structural control. Structural Engineering Research Laboratory, Swiss Federal Laboratories for Material Testing and Research, Dubendorf, Switzerland. 2006;173.
[42] Winkelbauer A. Moments and absolute moments of the normal distribution. arXiv preprint arXiv:12094340. 2012.