



## Full length article

# Bayesian model updating via streamlined Bayesian active learning cubature

Pei-Pei Li <sup>a</sup>, Chao Dang <sup>b</sup>\*, Cristóbal H. Acevedo <sup>b</sup>, Marcos A. Valdebenito <sup>b</sup>, Matthias G.R. Faes <sup>b,c</sup>

<sup>a</sup> Key Laboratory of Urban Security and Disaster Engineering of Ministry of Education, Beijing University of Technology, Beijing 100124, China

<sup>b</sup> Chair for Reliability Engineering, TU Dortmund University, Leonhard-Euler-Straße 5, Dortmund 44227, Germany

<sup>c</sup> International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, Shanghai 200092, China

## ARTICLE INFO

Communicated by I. Kougiumtzoglou

## Keywords:

Bayesian model updating  
Bayesian active learning  
Bayesian cubature  
Gaussian process regression  
Stopping criterion  
Learning function

## ABSTRACT

Bayesian model updating has been widely used for model calibration in engineering and scientific applications, as it not only updates model parameters but also quantifies their uncertainty. However, for computationally expensive models, efficiently estimating the model evidence (marginal likelihood) and the posterior distribution remains challenging. To address this issue, we propose a novel Bayesian active learning method for Bayesian model updating, which is termed as 'Streamlined Bayesian Active Learning Cubature' (SBALC). The core idea is to approximate the log-likelihood function using Gaussian process (GP) regression in a streamlined Bayesian active learning way. Rather than generating many samples from the posterior GP, we only use its mean and standard deviation function to form the model evidence estimator, stopping criterion, and learning function. Specifically, the estimation of model evidence is first treated as a Bayesian cubature problem, with a GP prior assigned over the log-likelihood function. Second, a plug-in estimator for model evidence is proposed based on the posterior mean function of the GP. Third, an upper bound on the expected absolute error between the posterior model evidence and its plug-in estimator is derived. Building on this result, a novel stopping criterion and learning function are proposed using only the posterior mean and standard deviation functions of the GP. Finally, we can obtain the model evidence based on the posterior mean function of the log-likelihood function in conjunction with Monte Carlo simulation, as well as the samples for the posterior distribution of model parameters as a by-product. Four numerical examples are presented to demonstrate the accuracy and efficiency of the proposed method compared to several existing approaches. The results show that the method can significantly reduce the number of model evaluations and the computational time without compromising accuracy.

## 1. Introduction

Computational models are widely used in applied sciences and engineering to simulate complex systems and support real-world decision-making. Although they provide valuable insights, their accuracy is highly dependent on the assumptions and data on which they are based. During system operation and as new measurements are collected, discrepancies often arise between model predictions and observed responses, which result from modeling errors (such as simplifications or neglected physics), parameter uncertainty,

\* Corresponding author.

E-mail address: [chao.dang@tu-dortmund.de](mailto:chao.dang@tu-dortmund.de) (C. Dang).

<https://doi.org/10.1016/j.ymssp.2026.114995>

Received 8 September 2025; Received in revised form 17 July 2026; Accepted 17 September 2026

Available online 4 October 2026

0888-3270/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Open access funding enabled and organized by Projekt DEAL.

measurement errors, and many others. To address these discrepancies, it is often necessary to update or calibrate uncertain model parameters to enhance predictive accuracy. Updating these parameters enhances the ability of models to capture real system behavior and supports more reliable decision-making in design, control, and maintenance processes. Traditional model updating methods, such as deterministic optimization [1] and gradient-based techniques [2,3], are inherently limited by their deterministic nature. They typically yield a single best-fit set of parameters while ignoring uncertainty. In recent years, Bayesian model updating has gained great attention because of its ability to update model parameters while quantifying the associated uncertainty. In this framework, the prior distribution of model parameters is updated to the posterior distribution based on measurements obtained from the real system [4,5]. Although analytical solutions exist in special cases, such as when conjugate priors are employed, computing the posterior is generally challenging in practical engineering. To address this challenge, three main categories of methods are commonly used: approximate analytical methods, simulation-based methods, and surrogate-based methods.

Approximate analytical methods address the Bayesian model updating problem by deriving explicit or semi-explicit expressions for the posterior distribution. One representative method is the Laplace approximation [6], which applies a second-order Taylor expansion to the log-posterior distribution around the maximum a posteriori (MAP) estimate, yielding a local Gaussian approximation. Another important class comprises Kalman filter-based methods. The classical Kalman filter [7,8] obtains the analytical results regarding the Gaussian posterior mean and variance of the model parameters under the assumptions of a linear state transition model, linear observation model, and Gaussian noise. For nonlinear systems, the Extended Kalman filter [9,10] approximates the nonlinear dynamics by applying a first-order Taylor expansion around the current state estimate, while the Unscented Kalman Filter [11] avoids explicit linearization by propagating a deterministically chosen set of sigma points through the nonlinear model and estimating the posterior mean and covariance from the transformed points. In addition to these, variational Bayes [12–14] offers a more general framework by optimizing the evidence lower bound to minimize the Kullback–Leibler divergence between a parameterized variational distribution and the true posterior distribution, thereby yielding a computationally tractable approximation. While these analytical methods are typically fast and interpretable to some extent, they often rely on restrictive modeling assumptions and local approximations and may perform poorly for strongly nonlinear, non-Gaussian, or multimodal problems.

Instead of pursuing explicit or semi-explicit formulas, simulation-based methods estimate the posterior distribution through numerical simulation, making them highly flexible for complex model updating problems. A major subclass comprises Markov Chain Monte Carlo (MCMC) algorithms [15–17], which generate dependent samples whose stationary distribution equals the target posterior distribution, using algorithms such as the Metropolis Hastings sampler [15,16], Gibbs sampler [18], and Hamiltonian Monte Carlo [19,20]. However, when the posterior and prior distributions have little overlap, conventional MCMC suffers from low acceptance rates and poor sampling efficiency. To address this issue, Transitional MCMC (TMCMC) [21] introduces a sequence of intermediate distributions that gradually transform the prior distribution into the posterior distribution, enabling more efficient exploration of high-probability regions without directly sampling from the sharply peaked posterior. Another important subclass is particle filters [22,23], also known as the Sequential Monte Carlo (SMC). It approximates the posterior distributions using a set of weighted particles, which are sequentially propagated and updated over time. By employing sequential importance sampling and resampling, SMC can effectively handle nonlinear dynamic systems and non-Gaussian noise, making it well-suited for complex model updating problems. Beyond these two subclasses, the Bayesian Updating with Structural Reliability methods (BUS) [24,25] offer a complementary perspective by reformulating the model updating problem as an equivalent structural reliability analysis. This framework enables the use of simulation-based reliability methods. While simulation-based methods are flexible and capable of handling complex posterior distributions, they require a large number of forward model evaluations and are computationally intensive.

Surrogate-based methods have attracted growing interest in recent years due to their ability to balance computational efficiency and estimation accuracy. These methods replace expensive models with computationally inexpensive approximations, thus improving the efficiency of model updating while maintaining acceptable accuracy. Among various surrogate modeling techniques such as polynomial chaos expansion (PCE) [26], artificial neural networks (ANN) [27], and Gaussian process regression (GPR) methods [28], the GPR (or Kriging) framework [29] has been extensively studied due to its modeling flexibility and its ability to provide both predictions and uncertainty quantification. The first subtype employs an active learning GPR to approximate expensive model evaluations during sampling, thereby improving the efficiency of Bayesian model updating. Owing to the success of active learning strategies in the reliability analysis field [30–32], the combination of the GPR model with the sampling method within the BUS framework has been widely investigated to improve the computational efficiency [33–36]. Beyond the BUS framework, several studies have directly coupled GPR models with simulation-based sampling techniques. For example, Angelikopoulos et al. [37] combined adaptive Kriging with TMCMC to reduce the computational cost of approximating the posterior probability distribution. Yoshida et al. [38] proposed an adaptive Gaussian process (GP) particle filters framework with a tailored learning function for efficient posterior PDF estimation; and Taflanidis et al. [39] embedded an adaptive GPR model within an SMC scheme, using intermediate auxiliary densities to efficiently sample from the target posterior distribution. The second subtype is GPR-based Bayesian cubature, initially proposed for multidimensional numerical integration [40]. It was later extended to model evidence estimation [41,42] by probabilistically integrating the likelihood function, and further to directly approximate the posterior distribution [43]. Model evidence measures the plausibility of the observed data under a specified prior distribution and likelihood function, thus providing a rigorous criterion for model assessment. As it is also regarded as an important part in model updating, recent research efforts [44–46] have been focused on its evaluation, with the posterior distribution provided as a by-product. These studies further demonstrated that coupling Bayesian cubature with active learning for model evidence estimation can substantially reduce the number of model evaluations and yield more robust uncertainty quantification of the posterior distributions. However, they typically rely on either generating many sample paths from the GP posterior or on applying a linearization trick to derive

closed-form mean and covariance, which is often computationally expensive and cumbersome, while the linearization may introduce approximation errors and lose accuracy for highly nonlinear dynamics. Therefore, it is worthwhile to investigate efficient sample-free Bayesian cubature methods for Bayesian model updating.

In this study, we propose a novel Bayesian active learning method for Bayesian model updating, referred to as ‘Streamlined Bayesian Active Learning Cubature’ (SBALC). The proposed method does not require generating samples from the posterior GP for producing the model evidence and posterior distribution. The main contributions of this work can be summarized as follows. First, we propose a plug-in estimator for the model evidence based on the posterior mean function of the log-likelihood function, which is computationally inexpensive and does not require samples from the posterior GP. Second, an upper bound on the expected absolute error is derived between the posterior model evidence and its plug-in estimator. This enables the quantification of epistemic uncertainty of the plug-in estimator using only the upper and lower credible bounds of the posterior GP. Third, a novel stopping criterion and a new learning function are defined merely using the posterior mean and standard deviation functions of the GP, providing efficient and reliable components for the active learning process. Finally, the model evidence can be obtained from the posterior mean function of the GP in conjunction with Monte Carlo simulation (MCS), as well as the posterior distribution of model parameters as a by-product. It is worth noting that, at the time of submission of this paper, a related approach has been reported by Wei [47]. Both studies employ a Gaussian process surrogate of the log-likelihood function and make use of the upper and lower bounds derived from the posterior mean and standard deviation of the GP. However, the present work differs in its theoretical development and the resulting active learning strategy. Specifically, this study derives an upper bound on the expected absolute error between the posterior model evidence and the proposed plug-in estimator. This result establishes a theoretical connection between the model evidence error and the GP-based bounds, which serves as the foundation for the proposed stopping criterion and learning function. Consequently, the active learning procedure is directly linked to the accuracy of model evidence estimation.

The remainder of this study is organized as follows. Section 2 presents the formulation of the problem. Section 3 details the proposed SBALC method. Section 4 provides four numerical examples to demonstrate the performance of the proposed method. Finally, Section 5 summarizes the main conclusions of this study.

## 2. Problem formulation

Bayesian model updating facilitates the calibration of model parameters and the improvement of predictive accuracy of computational models by integrating prior knowledge with newly observed data [4]. The primary objective of Bayesian model updating is to estimate the posterior probability density function (PDF) of the parameters, conditional on the observed data. This objective is achieved through Bayes’ theorem:

$$f(\mathbf{x} | \mathbf{Y}_{\text{obs}}) = \frac{L(\mathbf{Y}_{\text{obs}} | \mathbf{x})f(\mathbf{x})}{c}, \quad (1)$$

where  $f(\mathbf{x} | \mathbf{Y}_{\text{obs}})$  denotes the posterior PDF;  $\mathbf{x}$  denotes a realization of the  $d$ -dimensional model parameter vector  $\mathbf{X} = [X_1, X_2, \dots, X_d] \in D_{\mathbf{X}} \subseteq \mathbb{R}^d$  to be updated;  $\mathbf{Y}_{\text{obs}}$  represents the set of observations obtained from experiments or measurements;  $L(\mathbf{Y}_{\text{obs}} | \mathbf{x})$  is the likelihood function;  $f(\mathbf{x})$  denotes the prior PDF of the model parameters; and  $c$  is the model evidence, also known as the marginal likelihood.

The likelihood function  $L(\mathbf{Y}_{\text{obs}} | \mathbf{x})$  quantifies the probability of observing the data  $\mathbf{Y}_{\text{obs}}$  given the model parameters  $\mathbf{x}$  of the computational model  $\mathcal{R}(\cdot)$ . For simplicity,  $\mathcal{R}(\cdot)$  is assumed to be a deterministic function, so that all uncertainty arises only from the input model parameters. The likelihood function typically represents the discrepancy between the response of the computational model  $Y = \mathcal{R}(\mathbf{X}) : \mathbb{R}^d \mapsto \mathbb{R}$  and the observed data  $\mathbf{Y}_{\text{obs}}$ . If an additive representation of the measurement error is appropriate, the error term  $\varepsilon$  can be expressed as [24,48]:  $\varepsilon = \mathbf{Y}_{\text{obs}} - \mathcal{R}(\mathbf{X})$ . Alternatively, if the magnitude of the measurement error is expected to scale with the model observations, the error can be expressed in relative terms [17]:  $\varepsilon = \frac{Y_{\text{obs}}}{\mathcal{R}(\mathbf{X})} - 1$ . These two error models, absolute and relative, are commonly used to construct the likelihood function. The choice of an appropriate error model depends on the nature of the observed data. To unify these formulations, an error mapping function  $\mathcal{J}(\mathbf{X})$ , defined as  $\mathcal{J}(\mathbf{X}) = \mathcal{J}(\mathbf{Y}_{\text{obs}}, \mathcal{R}(\mathbf{X}))$ , is introduced. This function provides a consistent approach for representing errors, regardless of whether absolute or relative errors are considered. When the PDF of the measurement error is known, the likelihood function can be parameterized by the measurement data and computational model as  $L(\mathbf{Y}_{\text{obs}} | \mathbf{x}) = \varphi_{\varepsilon}(\mathcal{J}(\mathbf{x}))$ , where  $\varphi_{\varepsilon}$  represents the PDF of the measurement error  $\varepsilon$ .

Various models have been investigated to characterize the PDF of measurement error. Among these, the Gaussian model is most widely adopted, resulting in the following form of likelihood function:

$$L(\mathbf{Y}_{\text{obs}} | \mathbf{x}) \propto \exp\left(-\frac{\mathcal{J}(\mathbf{x})^2}{2\sigma_{\varepsilon}^2}\right), \quad (2)$$

where  $\sigma_{\varepsilon}$  denotes the standard deviation of the measurement error, which is assumed to be known in this study.

The model evidence  $c$  acts as the normalization constant in Bayes’ theorem, ensuring that the posterior distribution integrates into one. It is defined as the integral of the likelihood function with respect to the prior PDF over the parameter space  $D_{\mathbf{X}}$ :

$$c = \int_{D_{\mathbf{X}}} L(\mathbf{Y}_{\text{obs}} | \mathbf{x}) f(\mathbf{x}) d\mathbf{x}. \quad (3)$$

Once the prior  $f(\mathbf{x})$  and likelihood  $L(\mathbf{Y}_{\text{obs}} | \mathbf{x})$  are specified, obtaining the model evidence  $c$  and posterior distribution  $f(\mathbf{x} | \mathbf{Y}_{\text{obs}})$  are two key tasks in Bayesian model updating. Analytical solutions are typically not available except in special cases where the prior and likelihood are conjugate. In most engineering applications, however, such conjugacy does not hold, and the underlying models are often nonlinear and computationally expensive to evaluate, necessitating the use of efficient numerical or approximate methods.

### 3. Proposed SBALC method

This section introduces the proposed SBALC method for Bayesian model updating, which primarily focuses on model evidence estimation, while also providing samples of the posterior distribution as a byproduct. Section 3.1 introduces GPR for modeling the log-likelihood function. Section 3.2 proposes the plug-in model evidence estimator. Section 3.3 derives the upper bound of the expected absolute error between the posterior model evidence and its plug-in model evidence estimator, and proposes a novel stopping criterion. Section 3.4 presents the new learning function. Finally, Section 3.5 outlines the complete implementation procedure of the proposed method.

#### 3.1. Gaussian process regression

To evaluate the integral in Eq. (3), we adopt the Bayesian cubature principle, which interprets the model evidence estimation as a Bayesian inference problem. In this setting, the integrand is modeled as a random function through a GP. However, the likelihood function in Eq. (3) is strictly non-negative and often exhibits strong nonlinearity, which makes direct modeling difficult. To address this issue, we instead consider its logarithm, referred to as the log-likelihood function  $\mathcal{L}(\mathbf{x}) = \log(L(\mathbf{Y}_{\text{obs}}|\mathbf{x})) : \mathbb{R}^d \mapsto \mathbb{R}$ . Specifically, a GP prior is placed over  $\mathcal{L}(\mathbf{x})$ :

$$\mathcal{L}_0(\omega, \mathbf{x}) \sim \mathcal{GP}\left(m_{\mathcal{L}_0}(\mathbf{x}), \kappa_{\mathcal{L}_0}(\mathbf{x}, \mathbf{x}')\right), \quad (4)$$

where  $\mathcal{L}_0(\omega, \mathbf{x}) : \Omega \times D_{\mathbf{x}} \rightarrow \mathbb{R}$  denotes the GP prior over  $\mathcal{L}(\mathbf{x})$  and  $\omega \in \Omega$  denotes an elementary event in the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with  $\Omega$  is the sample space,  $\mathcal{F}$  is the  $\sigma$ -algebra of events, and  $\mathbb{P}$  is the associated probability measure;  $m_{\mathcal{L}_0}(\mathbf{x})$  and  $\kappa_{\mathcal{L}_0}(\mathbf{x}, \mathbf{x}')$  denote the prior mean and covariance functions, respectively. Since the log-likelihood is a deterministic function for fixed observed data, no observation noise is considered in the GP prior.

Without loss of generality, the prior mean is taken as a constant, and the covariance function is modeled by a Gaussian kernel given by:

$$m_{\mathcal{L}_0}(\mathbf{x}) = \beta, \quad (5)$$

$$\kappa_{\mathcal{L}_0}(\mathbf{x}, \mathbf{x}') = \sigma_0^2 \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \Sigma^{-1}(\mathbf{x} - \mathbf{x}')\right), \quad (6)$$

where  $\beta \in \mathbb{R}$ ;  $\sigma_0 > 0$  is the standard deviation of the process;  $\Sigma = \text{diag}(l_1^2, l_2^2, \dots, l_d^2)$  with  $l_i > 0$  being the length scale in the  $i$ th dimension. The  $d+2$  parameters collected in  $\mathcal{V} = [\beta, \sigma_0, l_1, l_2, \dots, l_d]$  are referred to as hyperparameters. The selection of a constant mean is a standard choice in Gaussian process modeling when no reliable prior knowledge about the global trend is available. For the covariance function in Eq. (6), the squared exponential kernel is adopted due to its smoothness properties and widespread use in Bayesian inference and surrogate modeling. Finally, the assumption of a diagonal length-scale matrix corresponds to the standard automatic relevance determination (ARD) formulation, which is widely used in Gaussian process based Bayesian model updating.

Suppose that we have a supervised training dataset  $\mathcal{D} = \{\mathcal{X}, \mathcal{Y}\}$ , where  $\mathcal{X} = \{\mathbf{x}^{(i)}\}_{i=1}^n$  is an  $n \times d$  matrix with its  $i$ th row being  $\mathbf{x}^{(i)}$ .  $\mathcal{Y} = [\mathcal{Y}^{(1)}, \mathcal{Y}^{(2)}, \dots, \mathcal{Y}^{(n)}]^T$  contains the corresponding outputs given by the computational model, i.e.,  $\mathcal{Y}^{(i)} = \mathcal{L}(\mathbf{x}^{(i)})$ . It is noted that  $n$  denotes the number of training samples used for GP surrogate construction, rather than the number of observed data points. Each  $\mathcal{Y}^{(i)}$  corresponds to the log-likelihood evaluated at the training input  $\mathbf{x}^{(i)}$ . The hyperparameters of the prior mean and covariance functions are estimated via maximum likelihood estimation. The posterior of the log-likelihood function conditional on  $\mathcal{D}$  remains a GP:

$$\mathcal{L}_n(\omega, \mathbf{x}) \sim \mathcal{GP}(m_{\mathcal{L}_n}(\mathbf{x}), \kappa_{\mathcal{L}_n}(\mathbf{x}, \mathbf{x}')), \quad (7)$$

with mean and covariance:

$$m_{\mathcal{L}_n}(\mathbf{x}) = m_{\mathcal{L}_0}(\mathbf{x}) + \kappa_{\mathcal{L}_0}(\mathbf{x}, \mathcal{X})^T \mathbf{K}_{\mathcal{L}_0}^{-1} (\mathcal{Y} - m_{\mathcal{L}_0}(\mathcal{X})), \quad (8)$$

$$\kappa_{\mathcal{L}_n}(\mathbf{x}, \mathbf{x}') = \kappa_{\mathcal{L}_0}(\mathbf{x}, \mathbf{x}') - \kappa_{\mathcal{L}_0}(\mathbf{x}, \mathcal{X})^T \mathbf{K}_{\mathcal{L}_0}^{-1} \kappa_{\mathcal{L}_0}(\mathcal{X}, \mathbf{x}'), \quad (9)$$

where  $m_{\mathcal{L}_0}(\mathcal{X})$  is an  $n \times 1$  mean vector with  $i$ th element being  $m_{\mathcal{L}_0}(\mathbf{x}^{(i)})$ ;  $\mathbf{K}_{\mathcal{L}_0}$  is an  $n$ -by- $n$  covariance matrix whose  $(i_1, i_2)$ -th element is  $\kappa_{\mathcal{L}_0}(\mathbf{x}^{(i_1)}, \mathbf{x}^{(i_2)})$  evaluated at the  $i_1$ -th and  $i_2$ -th training samples;  $\kappa_{\mathcal{L}_0}(\mathbf{x}, \mathcal{X})$  and  $\kappa_{\mathcal{L}_0}(\mathcal{X}, \mathbf{x}')$  are two  $n \times 1$  covariance vectors with  $i$ th element being  $\kappa_{\mathcal{L}_0}(\mathbf{x}, \mathbf{x}^{(i)})$  and  $\kappa_{\mathcal{L}_0}(\mathbf{x}^{(i)}, \mathbf{x}')$ , respectively.

The posterior GP provides not only a mean prediction, but also a measure of uncertainty in the prediction at any unseen point, given by the posterior variance  $\sigma_{\mathcal{L}_n}^2(\mathbf{x}) = \kappa_{\mathcal{L}_n}(\mathbf{x}, \mathbf{x})$ . As a final remark, note that the vector  $\mathbf{x}$  which is the argument of the GPR may involve quantities with different physical meaning. To avoid potential issues regarding their order of magnitude, the GPR is trained considered standardization of the input data.

### 3.2. Plug-in estimator for the model evidence

Since the  $\mathcal{L}_n(\omega, \mathbf{x})$  follows a GP, exponentiating it yields the posterior likelihood process  $e^{\mathcal{L}_n(\omega, \mathbf{x})}$ , which is a lognormal process. The posterior model evidence is defined as:

$$c_n(\omega) = \int_{D_{\mathbf{x}}} e^{\mathcal{L}_n(\omega, \mathbf{x})} f(\mathbf{x}) d\mathbf{x}. \quad (10)$$

Since the posterior model evidence is a function of  $\omega$ , it is also a random variable. The randomness of posterior model evidence originates from the GP model itself, reflecting epistemic uncertainty due to the limited number of log-likelihood evaluations. As more evaluations of the log-likelihood function are collected, this uncertainty may diminish and the posterior model evidence  $c_n(\omega)$  converges to the true model evidence  $c^*$ . However, since the exact distribution of  $c_n(\omega)$  is analytically intractable, we introduce a plug-in model evidence estimator by replacing  $\mathcal{L}_n(\omega, \mathbf{x})$  with its posterior mean function:

$$\tilde{c}_n = \int_{D_{\mathbf{x}}} e^{m_{\mathcal{L}_n}(\mathbf{x})} f(\mathbf{x}) d\mathbf{x}. \quad (11)$$

This plug-in estimator will also converge to the true model evidence  $c^*$  as more evaluations of the log-likelihood function become available (when  $\sigma_{\mathcal{L}_n}(\mathbf{x}) \rightarrow 0$  and  $m_{\mathcal{L}_n}(\mathbf{x}) \rightarrow \mathcal{L}(\mathbf{x})$  for all  $\mathbf{x} \in D_{\mathbf{x}}$ ). However, the integral in Eq. (11) is generally intractable in closed form. Therefore, Monte Carlo simulation (MCS) is adopted to evaluate the plug-in model evidence  $\tilde{c}_n$ . The MCS estimator of  $\tilde{c}_n$  is given by:

$$\hat{\tilde{c}}_n = \frac{1}{N_M} \sum_{j=1}^{N_M} e^{m_{\mathcal{L}_n}(\mathbf{x}^{(j)})}, \quad (12)$$

where  $\mathbf{x}^{(j)}$  is the  $j$ th random sample in the sample set  $\mathbb{S} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N_M)}\}$  generated from the prior PDF;  $N_M$  denotes the number of random samples.

The variance of the estimator  $\hat{\tilde{c}}_n$  is expressed as:

$$\text{VAR}[\hat{\tilde{c}}_n] = \frac{1}{N_M(N_M - 1)} \sum_{j=1}^{N_M} \left( e^{m_{\mathcal{L}_n}(\mathbf{x}^{(j)})} - \hat{\tilde{c}}_n \right)^2, \quad (13)$$

where VAR denotes the variance operator.

Furthermore, the coefficient of variation (CoV) of the estimator  $\hat{\tilde{c}}_n$  is defined as  $\text{CoV}[\hat{\tilde{c}}_n] = \sqrt{\text{VAR}[\hat{\tilde{c}}_n] / \hat{\tilde{c}}_n}$ . This indicates that a large sample size leads to a more reliable estimate.

### 3.3. Uncertainty measure and stopping criterion

After obtaining the model evidence estimate, a stopping criterion is needed to evaluate whether the estimate reaches a desired level of accuracy. Based on the predictions of the posterior GP for the log-likelihood function, the upper credible bound and the lower credible bound of the posterior GP are defined as  $u_n(\mathbf{x}) = m_{\mathcal{L}_n}(\mathbf{x}) + b\sigma_{\mathcal{L}_n}(\mathbf{x})$  and  $l_n(\mathbf{x}) = m_{\mathcal{L}_n}(\mathbf{x}) - b\sigma_{\mathcal{L}_n}(\mathbf{x})$ . Here, the parameter  $b$  ( $b > 0$ ) is a real number that controls the width of the credible interval; in particular, setting  $b = \Phi^{-1}(1 - \alpha/2)$  yields a  $(1 - \alpha) \times 100\%$  credible interval, where  $\Phi^{-1}(\cdot)$  denotes the inverse cumulative distribution function (CDF) of the standard normal variable. If the posterior GP  $\mathcal{L}_n(\omega, \mathbf{x})$  in Eq. (10) is replaced by its upper and lower credible bounds, two new quantities can be defined:

$$\bar{c}_n = \int_{D_{\mathbf{x}}} \left( e^{(m_{\mathcal{L}_n}(\mathbf{x}) + b\sigma_{\mathcal{L}_n}(\mathbf{x}))} \right) f(\mathbf{x}) d\mathbf{x}, \quad (14)$$

$$\underline{c}_n = \int_{D_{\mathbf{x}}} \left( e^{(m_{\mathcal{L}_n}(\mathbf{x}) - b\sigma_{\mathcal{L}_n}(\mathbf{x}))} \right) f(\mathbf{x}) d\mathbf{x}, \quad (15)$$

where  $\bar{c}_n$  and  $\underline{c}_n$  are the upper and lower bounds of the posterior model evidence  $c_n(\omega)$ . These bounds cannot be regarded as  $(1 - \alpha) \times 100\%$  credible bound for  $c_n(\omega)$ , because the integration operator does not, in general, preserve the coverage probability implied by the point-wise credible bounds of  $\mathcal{L}_n(\omega, \mathbf{x})$ .

**Proposition 1.** For  $b > 0$ , there exists  $\underline{c}_n \leq \tilde{c}_n \leq \bar{c}_n$ .

**Proof.** Based on Eqs. (11) and (15), we have:

$$\underline{c}_n - \tilde{c}_n = \int_{D_{\mathbf{x}}} \left[ e^{-b\sigma_{\mathcal{L}_n}(\mathbf{x})} - 1 \right] e^{m_{\mathcal{L}_n}(\mathbf{x})} f(\mathbf{x}) d\mathbf{x}. \quad (16)$$

Since  $b > 0$  and  $\sigma_{\mathcal{L}_n}(\mathbf{x}) \geq 0$ , we have  $e^{-b\sigma_{\mathcal{L}_n}(\mathbf{x})} - 1 \leq 0$ , with equality if and only if  $\sigma_{\mathcal{L}_n}(\mathbf{x}) = 0$ . Multiplying this inequality by the non-negative function  $e^{m_{\mathcal{L}_n}(\mathbf{x})} f(\mathbf{x})$  preserves the sign, and integration over the input space yields  $\underline{c}_n \leq \tilde{c}_n$ , with equality only when  $\sigma_{\mathcal{L}_n}(\mathbf{x}) = 0$ . Similarly to the proof of the first inequality  $\underline{c}_n \leq \tilde{c}_n$ , the second inequality  $\tilde{c}_n \leq \bar{c}_n$  can also be proved. Combining  $\underline{c}_n \leq \tilde{c}_n$  and  $\tilde{c}_n \leq \bar{c}_n$  completes the proof, holding if and only if  $\sigma_{\mathcal{L}_n}(\mathbf{x}) = 0$ .

**Proposition 1** indicates that for any  $b > 0$ , the plug-in model evidence estimator  $\tilde{c}_n$  always lies between its lower and upper bounds. Moreover, as the standard deviation  $\sigma_{\mathcal{L}_n}(\mathbf{x})$  of the posterior GP tends to zero, both bounds converge to the plug-in model evidence estimator:  $\underline{c}_n \rightarrow \tilde{c}_n$  and  $\bar{c}_n \rightarrow \tilde{c}_n$ .

**Proposition 2.** For  $b > 0$ , there exists a finite constant  $M_b$  such that  $\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] \leq M_b \cdot (\bar{c}_n - \underline{c}_n)$ .

**Proof.** From Eqs. (10) and (11), the expected absolute error between  $c_n(\omega)$  and  $\tilde{c}_n$  is given by:

$$\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] = \mathbb{E}_\omega \left[ \left| \int_{D_X} (e^{\mathcal{L}_n(\omega, \mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x})}) f(\mathbf{x}) d\mathbf{x} \right| \right], \quad (17)$$

where  $\mathbb{E}_\omega[\cdot]$  denotes the expectation over  $\omega$ .

By the triangle inequality and Tonelli's theorem [49], it follows that:

$$\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] \leq \int_{D_X} \mathbb{E}_\omega \left[ \left| e^{\mathcal{L}_n(\omega, \mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x})} \right| \right] f(\mathbf{x}) d\mathbf{x}. \quad (18)$$

To evaluate the expectation in the integrand of Eq. (18), we fix  $\mathbf{x}^* \in D_X$ . By the properties of GPR,  $\mathcal{L}_n(\omega, \mathbf{x}^*)$  follows a Gaussian distribution:

$$\mathcal{L}_n(\omega, \mathbf{x}^*) \stackrel{d}{=} m_{\mathcal{L}_n}(\mathbf{x}^*) + \sigma_{\mathcal{L}_n}(\mathbf{x}^*)Z, \quad Z \sim \mathcal{N}(0, 1). \quad (19)$$

Hence, the expectation in the integrand of Eq. (18) can be reformulated in terms of  $Z$ :

$$\mathbb{E}_\omega \left[ \left| e^{\mathcal{L}_n(\omega, \mathbf{x}^*)} - e^{m_{\mathcal{L}_n}(\mathbf{x}^*)} \right| \right] = e^{m_{\mathcal{L}_n}(\mathbf{x}^*)} \mathbb{E}_Z \left[ \left| e^{\sigma_{\mathcal{L}_n}(\mathbf{x}^*)Z} - 1 \right| \right]. \quad (20)$$

Let  $a := \sigma_{\mathcal{L}_n}(\mathbf{x}^*) \geq 0$ . The detailed derivation of  $\mathbb{E}_Z[|e^{aZ} - 1|]$  is provided in Appendix A.1. The resulting closed-form expression is:

$$\mathbb{E}_Z[|e^{aZ} - 1|] = e^{a^2/2} (\Phi(a) - \Phi(-a)) = e^{a^2/2} (2\Phi(a) - 1). \quad (21)$$

Substituting Eq. (21) into Eq. (20), and further into Eq. (18), leads to:

$$\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] \leq \int_{D_X} e^{m_{\mathcal{L}_n}(\mathbf{x}) + \frac{1}{2}\sigma_{\mathcal{L}_n}^2(\mathbf{x})} [2\Phi(\sigma_{\mathcal{L}_n}(\mathbf{x})) - 1] f(\mathbf{x}) d\mathbf{x}. \quad (22)$$

On the other hand, the difference between  $U_e(\mathbf{x}) = \exp(u_n(\mathbf{x}))$  and  $L_e(\mathbf{x}) = \exp(l_n(\mathbf{x}))$  is given as:

$$U_e(\mathbf{x}) - L_e(\mathbf{x}) = e^{m_{\mathcal{L}_n}(\mathbf{x})} \left( e^{b\sigma_{\mathcal{L}_n}(\mathbf{x})} - e^{-b\sigma_{\mathcal{L}_n}(\mathbf{x})} \right) = 2e^{m_{\mathcal{L}_n}(\mathbf{x})} \sinh(b\sigma_{\mathcal{L}_n}(\mathbf{x})). \quad (23)$$

By rearranging Eq. (23), we can isolate  $e^{m_{\mathcal{L}_n}(\mathbf{x})}$  as

$$e^{m_{\mathcal{L}_n}(\mathbf{x})} = \frac{U_e(\mathbf{x}) - L_e(\mathbf{x})}{2 \sinh(b\sigma_{\mathcal{L}_n}(\mathbf{x}))}. \quad (24)$$

Inserting Eq. (24) into Eq. (22) yields:

$$\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] \leq \int_{D_X} (U_e(\mathbf{x}) - L_e(\mathbf{x})) h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) f(\mathbf{x}) d\mathbf{x}, \quad (25)$$

where

$$h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) = \frac{e^{\frac{1}{2}\sigma_{\mathcal{L}_n}^2(\mathbf{x})} [2\Phi(\sigma_{\mathcal{L}_n}(\mathbf{x})) - 1]}{2 \sinh(b\sigma_{\mathcal{L}_n}(\mathbf{x}))}. \quad (26)$$

The boundedness of  $h_b(\cdot)$  on  $[0, \sigma_0]$  is established in Appendix A.2. Therefore, there exists a finite constant  $M_b := \sup_{\sigma_{\mathcal{L}_n}(\mathbf{x}) \in [0, \sigma_0]} h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) < \infty$ . Consequently,

$$\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|] \leq M_b \int_{D_X} \left( e^{m_{\mathcal{L}_n}(\mathbf{x}) + b\sigma_{\mathcal{L}_n}(\mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x}) - b\sigma_{\mathcal{L}_n}(\mathbf{x})} \right) f(\mathbf{x}) d\mathbf{x} = M_b \cdot (\bar{c}_n - \underline{c}_n). \quad (27)$$

This completes the proof.

**Proposition 2** shows that the expected absolute error between  $c_n(\omega)$  and  $\tilde{c}_n$  is bounded by a constant multiple of the difference between the upper and lower bounds of  $c_n(\omega)$ , namely  $\bar{c}_n - \underline{c}_n$ . This result implies that the epistemic uncertainty in  $c_n(\omega)$  can be quantified solely through  $\bar{c}_n - \underline{c}_n$ . Moreover, if  $\bar{c}_n - \underline{c}_n$  converges to zero, then the expected absolute error  $\mathbb{E}_\omega[|c_n(\omega) - \tilde{c}_n|]$  also converges to zero, thereby providing a theoretical guarantee that the plug-in model evidence estimator converges to the true model evidence  $c^*$ .

Taken together, Propositions 1 and 2 provide a rigorous theoretical foundation for designing the stopping criterion and learning function.

Based on the above propositions, a stopping criterion is proposed as follows:

$$\frac{\bar{c}_n - \underline{c}_n}{\bar{c}_n} < \epsilon_{\text{RE}}, \quad (28)$$

where  $\epsilon_{\text{RE}}$  denotes the user-defined threshold of the relative error.

The stopping criterion in Eq. (28) requires that the difference between the upper and lower bounds of the posterior model evidence  $\bar{c}_n - \underline{c}_n$  relative to the plug-in estimator  $\bar{c}_n$  needs to be smaller than a specified threshold  $\epsilon_{\text{RE}}$ . In other words, the iterative procedure ends once the relative epistemic uncertainty in  $c_n(\omega)$  is sufficiently small. To implement the proposed stopping criterion, all components in Eq. (28) need to be estimated. The denominator  $\bar{c}_n$  can be computed using MCS as shown in Eq. (12). Likewise, the  $\bar{c}_n$  and  $\underline{c}_n$  in the numerator can also be estimated via MCS, and are given by:

$$\hat{\bar{c}}_n = \frac{1}{N_M} \sum_{j=1}^{N_M} e^{(m_{\mathcal{L}_n}(\mathbf{x}^{(j)}) + b\sigma_{\mathcal{L}_n}(\mathbf{x}^{(j)}))}, \quad (29)$$

$$\hat{\underline{c}}_n = \frac{1}{N_M} \sum_{j=1}^{N_M} e^{(m_{\mathcal{L}_n}(\mathbf{x}^{(j)}) - b\sigma_{\mathcal{L}_n}(\mathbf{x}^{(j)}))}. \quad (30)$$

**Remark 1.** In addition to the stopping criterion in Eq. (28), other two possible options are:  $\frac{\bar{c}_n - \underline{c}_n}{\bar{c}_n} < \epsilon_{\text{RE}}$  and  $\frac{\bar{c}_n - \underline{c}_n}{\underline{c}_n} < \epsilon_{\text{RE}}$ .

### 3.4. Learning function and the best next point selection

If the stopping criterion has not yet been met, a well-designed learning function is needed to select the best next point for improving the quality of the GPR surrogate model of the log-likelihood function, aiming to further improve the accuracy of the model evidence estimate. For this purpose, we introduce a new learning function defined as:

$$\mathcal{LF}(\mathbf{x}) = \underbrace{\sigma_{\mathcal{L}_n}^2(\mathbf{x})}_{\textcircled{1}} \cdot \underbrace{\left( e^{m_{\mathcal{L}_n}(\mathbf{x}) + b\sigma_{\mathcal{L}_n}(\mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x}) - b\sigma_{\mathcal{L}_n}(\mathbf{x})} \right)}_{\textcircled{2}} \cdot f(\mathbf{x}), \quad (31)$$

where the first term  $\textcircled{1}$  serves as a weighting factor that emphasizes regions of high uncertainty, while the second term  $\textcircled{2}$  contributes to the difference between the upper and lower bounds of the posterior model evidence. Actually, it is the integrand of  $\bar{c}_n - \underline{c}_n$ . A larger value of  $\mathcal{LF}(\mathbf{x})$  indicates that the candidate point is more promising.

It is worth noting that the proposed learning function is directly connected to the uncertainty measure introduced in Proposition 2. The second term corresponds exactly to the integrand of the upper-lower evidence gap  $\bar{c}_n - \underline{c}_n$ , which quantifies the epistemic uncertainty of the model evidence estimate. Consequently, sampling points in regions where this quantity is large is expected to produce the greatest reduction in the global uncertainty measure  $\bar{c}_n - \underline{c}_n$ , thereby improving the accuracy of the model evidence estimation. Although the second term already depends on the GP standard deviation through the credible bounds, its magnitude is also strongly influenced by the likelihood level  $e^{m_{\mathcal{L}_n}(\mathbf{x})}$ . As a result, regions with high likelihood values may dominate the learning process even when the surrogate uncertainty is relatively small. The additional variance weighting term  $\sigma_{\mathcal{L}_n}^2(\mathbf{x})$  is therefore introduced to further emphasize regions where the GP surrogate remains uncertain and to achieve a better balance between exploitation and exploration. By combining both terms, the proposed learning function simultaneously accounts for the local contribution to the model evidence uncertainty and the uncertainty of the surrogate model itself, resulting in more effective sample selection. This differs from many existing active learning criteria, which are typically constructed based on surrogate uncertainty, likelihood magnitude, or numerical integration error indicators separately.

The next evaluation point is selected by maximizing the learning function:

$$\mathbf{x}^{(n+1)} = \arg \max_{\mathbf{x} \in D_{\delta_1}} \mathcal{LF}(\mathbf{x}), \quad (32)$$

where the feasible domain  $D_{\delta_1}$  is defined as  $D_{\delta_1} = \prod_{j=1}^d [F_j^{-1}(\delta_1), F_j^{-1}(1 - \delta_1)]$ , where  $F_j^{-1}$  is the inverse prior CDF of  $X_j$  and  $\delta_1$  is a small truncation probability.

**Remark 2.** If  $\frac{\bar{c}_n - \underline{c}_n}{\bar{c}_n} < \epsilon_{\text{RE}}$  is adopted as the stopping criterion, the corresponding learning function can be:  $\mathcal{LF}(\mathbf{x}) = \sigma_{\mathcal{L}_n}^2(\mathbf{x}) \cdot \left( e^{m_{\mathcal{L}_n}(\mathbf{x}) + b\sigma_{\mathcal{L}_n}(\mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x})} \right) \cdot f(\mathbf{x})$ .

**Remark 3.** If  $\frac{\bar{c}_n - \underline{c}_n}{\underline{c}_n} < \epsilon_{\text{RE}}$  is adopted as the stopping criterion, the corresponding learning function can be:  $\mathcal{LF}(\mathbf{x}) = \sigma_{\mathcal{L}_n}^2(\mathbf{x}) \cdot \left( e^{m_{\mathcal{L}_n}(\mathbf{x})} - e^{m_{\mathcal{L}_n}(\mathbf{x}) - b\sigma_{\mathcal{L}_n}(\mathbf{x})} \right) \cdot f(\mathbf{x})$ .

### 3.5. Implementation procedure of the proposed method

The implementation procedure of the proposed SBALC method is summarized below, along with a flowchart in Fig. 1.

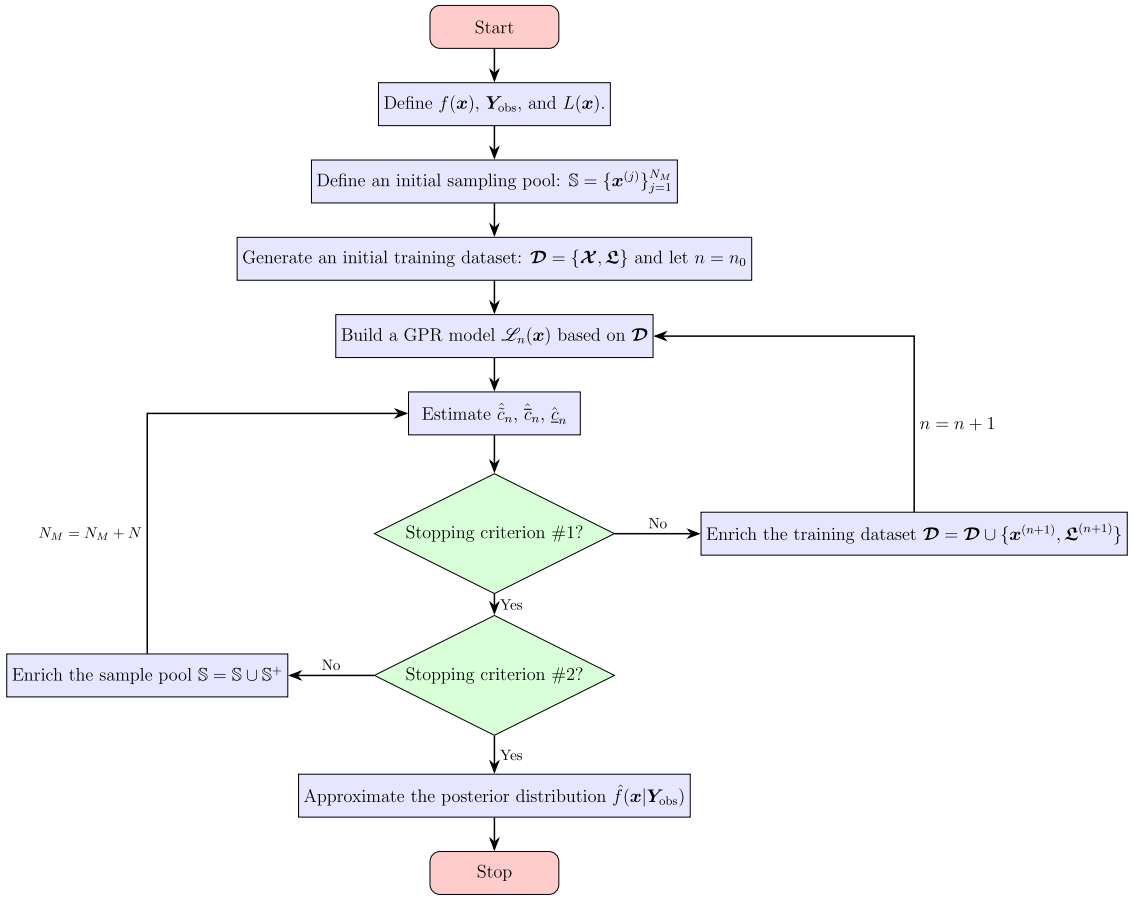


Fig. 1. Flowchart of the proposed method.

**Step 1: Initialization**

Define the prior distribution  $f(\mathbf{x})$  of model parameters, and specify the likelihood function  $L(\mathbf{x})$  based on observed data  $\mathbf{Y}_{\text{obs}}$ , the specified measurement error  $\varepsilon$ , and the response function  $R(\mathbf{X})$ .

**Step 2: Generate prior samples**

Generate a set of  $N_M$  random samples according to the prior PDF  $f(\mathbf{x})$ , denoted as  $\mathbb{S} = \{\mathbf{x}^{(j)}\}$  for  $j = 1, 2, \dots, N_M$ .

**Step 3: Generate an initial training dataset**

First, a small number of  $n_0$  input samples  $\mathcal{X} = \{\mathbf{x}^{(i)}\}_{i=1}^{n_0}$  are generated using Hammersley sequence, which is a low-discrepancy quasi-random sequence that provides a more uniform coverage of the input space than purely random sampling within the feasible truncated domain  $D_{\delta_0}$ .  $D_{\delta_0}$  and  $D_{\delta_1}$  have the same mathematical form but are used in different stages of the algorithm.  $\delta_0$  is employed for generating the initial training samples, whereas  $\delta_1$  is used during the adaptive sampling stage. Then, the log-likelihood function  $\mathcal{L}(\mathbf{x})$  is evaluated at these points to produce the corresponding response values, i.e.,  $\mathcal{Z} = \{\mathcal{L}(\mathbf{x}^{(i)})\}_{i=1}^{n_0}$ . Finally, the initial training set is constructed by  $\mathcal{D} = \{\mathcal{X}, \mathcal{Z}\}$ . Let  $n = n_0$ .

**Step 4: Build a GPR model**

Train a GP model  $\mathcal{L}_n(\omega, \mathbf{x})$  to approximate the log-likelihood function using the training dataset  $\mathcal{D} = \{\mathcal{X}, \mathcal{Z}\}$ . In the present study, the fitrgp function from Matlab R2024a with a constant prior mean and an squared exponential kernel (see Eq. (6)) is employed.

**Step 5: Estimate the model evidence**

Estimate the plug-in model evidence estimator  $\hat{c}_n$  using MCS over  $\mathbb{S}$ , as well as the lower bound  $\hat{c}_n$  and the upper bound  $\hat{c}_n$ .

**Step 6: Check the stopping criterion #1**

If  $\frac{\hat{c}_n - \hat{c}_n}{\hat{c}_n} < \epsilon_{\text{RE}}$  is satisfied for two successive iterations, go to **Step 8**; Otherwise, go to **Step 7**.

**Step 7: Enrich the training dataset**

Identify the next point by solving  $\mathbf{x}^{(n+1)} = \arg \max_{\mathbf{x} \in D_{\delta_1}} \mathcal{L}F(\mathbf{x})$  using any appropriate optimization algorithm such as, e.g. starfish optimization algorithm (SFOA) [50]. Evaluate the log-likelihood function at  $\mathbf{x}^{(n+1)}$  to obtain  $\mathcal{Z}^{(n+1)} = \mathcal{L}(\mathbf{x}^{(n+1)})$ , and update the training set:  $\mathcal{D} = \mathcal{D} \cup \{\mathbf{x}^{(n+1)}, \mathcal{Z}^{(n+1)}\}$ . Set  $n = n + 1$  and return to **Step 4**.

**Step 8: Check the stopping criterion #2**

Calculate the CoV of plug-in model evidence estimator  $\hat{c}_n$ , denoted as  $\text{CoV}[\hat{c}_n]$ . If  $\text{CoV}[\hat{c}_n] \leq \eta$ , proceed to **Step 10**; otherwise, go to **Step 9**.

**Step 9: Enrich the sample pool**

To reduce the sampling variability in the plug-in model evidence estimator, enlarge the current sample set. Generate an additional set of  $N$  samples, denoted as  $\mathbb{S}^+$ , from the prior distribution as described in **Step 2**. Update the existing sample pool with the new sample, i.e.,  $\mathbb{S} = \mathbb{S} \cup \mathbb{S}^+$ , and proceed to **Step 5**.

**Step 10: Approximate the posterior PDF**

Once the plug-in model evidence estimator  $\hat{c}_n$  and the plug-in estimator  $e^{m_{\mathcal{L}_n}(\mathbf{x})}$  for the likelihood function have been adequately estimated, the posterior PDF can be constructed. The approximation of the posterior PDF is given by:

$$\hat{f}(\mathbf{x}|\mathbf{Y}_{\text{obs}}) = \frac{e^{m_{\mathcal{L}_n}(\mathbf{x})} f(\mathbf{x})}{\hat{c}_n}. \quad (33)$$

With the posterior PDF available, the mean and standard deviation of the model parameters can be directly evaluated via numerical integration; in this study, the MCS is adopted. Furthermore, if posterior samples are of interest, they can be generated using sampling strategies based on the posterior PDF, sampling-importance-resampling (SIR) [51], sequential importance sampling, or acceptance-rejection sampling. In this study, the SIR method is used.

**4. Numerical examples**

In this section, four numerical examples are presented to illustrate the performance of the proposed method. In each of these examples, the parameters in the proposed SBALC method are set as:  $n_0 = 4 \sim 10$ ,  $N_M = 2 \times 10^4$ ,  $N = 2 \times 10^4$ ,  $b = 1$ ,  $\epsilon_{\text{RE}} = 0.1$ ,  $\eta = 0.02$ ,  $\delta_0 = 1 \times 10^{-2}$ , and  $\delta_1 = 1 \times 10^{-5}$ . For comparison, several state-of-the-art methods, including TMCMC [21,52], BUS [24,53], and ABQ-BU [45] are evaluated. Among these methods, TMCMC and BUS are direct simulation-based approaches that do not employ surrogate models, whereas ABQ-BU is a surrogate-assisted active learning approach based on Gaussian process regression and adaptive Bayesian quadrature. Since both ABQ-BU and SBALC belong to the class of surrogate-based Bayesian updating methods, ABQ-BU serves as the most relevant benchmark for evaluating the effectiveness of the proposed active learning strategy. To ensure the robustness of the results, both the proposed method and the comparison methods are repeated 20 times, and the corresponding CoV is calculated. The reported numbers of model evaluations and computational time are averaged over repeated runs. For the last two examples, however, the TMCMC and BUS are only conducted once due to infeasibility. All simulations are performed on a Windows 10 system equipped with an Intel(R) Core(TM) i9-14900 K processor (2.19 GHz), 60 GB RAM, and a 64-bit architecture.

**4.1. Example 1: An illustrative example**

The first example considers a sigmoid function, which has an analytical solution for illustrative purposes. This example involves a single random variable  $X$ , with a prior distribution following the Gaussian distribution  $\mathcal{N}(1.5, 2^2)$ . The corresponding response function is expressed as follows:

$$\mathcal{R}(X) = \frac{10}{1 + e^{-1.2(X-1)}}. \quad (34)$$

In this example, the observation data is set to  $Y_{\text{obs}} = \mathcal{R}(1) = 5$ . To construct the likelihood function, we assumed that the prediction error follows a zero-mean Gaussian distribution with a standard deviation of  $\sigma_e = 0.2$ . According to Eq. (2), the likelihood function is expressed as:

$$L(Y_{\text{obs}}|x) = \exp\left(-\frac{(\mathcal{R}(1) - \mathcal{R}(x))^2}{2 \times 0.2^2}\right). \quad (35)$$

Since the likelihood function is a deterministic function of the random variable  $X$ , it enables a semi-analytical solution for the posterior PDF, which serves as a reference for comparison.

Fig. 2 shows the computational details of the proposed method with an initial sample size of  $n_0 = 4$ . Each column corresponds to a different number of training samples ( $n = 4, 5, 8$ ), and includes three subplots: (1) the GPR model plot with 95% prediction interval, (2) the learning function  $\mathcal{LF}(x)$  with the optimal next training point, and (3) the approximate posterior PDF compared to the analytical solution. At  $n = n_0 = 4$ , where only the initial training points is used, the uncertainty of GPR model is large, as indicated by the wide prediction interval, and the estimated posterior PDF significantly deviates from the analytical result. The learning function effectively captures the most informative region in the GPR model and guides the selection of an optimal next training point. When one new point is added ( $n = 5$ ), the uncertainty is greatly reduced, leading to a posterior PDF that more closely approximates the analytical one. By  $n = 8$ , the prediction interval becomes narrow, and the estimated posterior PDF aligns very well with the analytical solution, demonstrating the efficiency and effectiveness of the proposed active learning method.

To further assess the performance of the proposed method, comparisons are made with BUS, TMCMC, and ABQ-BU. The comparison results, including computational time, the number of model evaluations ( $N_{\text{call}}$ ), model evidence estimation ( $\hat{c}$ ), and the posterior mean ( $\hat{\mu}_X$ ) and standard deviation ( $\hat{\sigma}_X$ ) of  $X$ , are presented in Table 1. As shown in Table 1, the model evidence estimation, posterior mean and standard deviation of  $X$  obtained from the proposed method closely match the reference values of the analytical PDF, while maintaining a low CoV that indicates high robustness. Moreover, compared to BUS, TMCMC, and ABQ-BU, the proposed method requires fewer model evaluations and attains the second-lowest computational time.

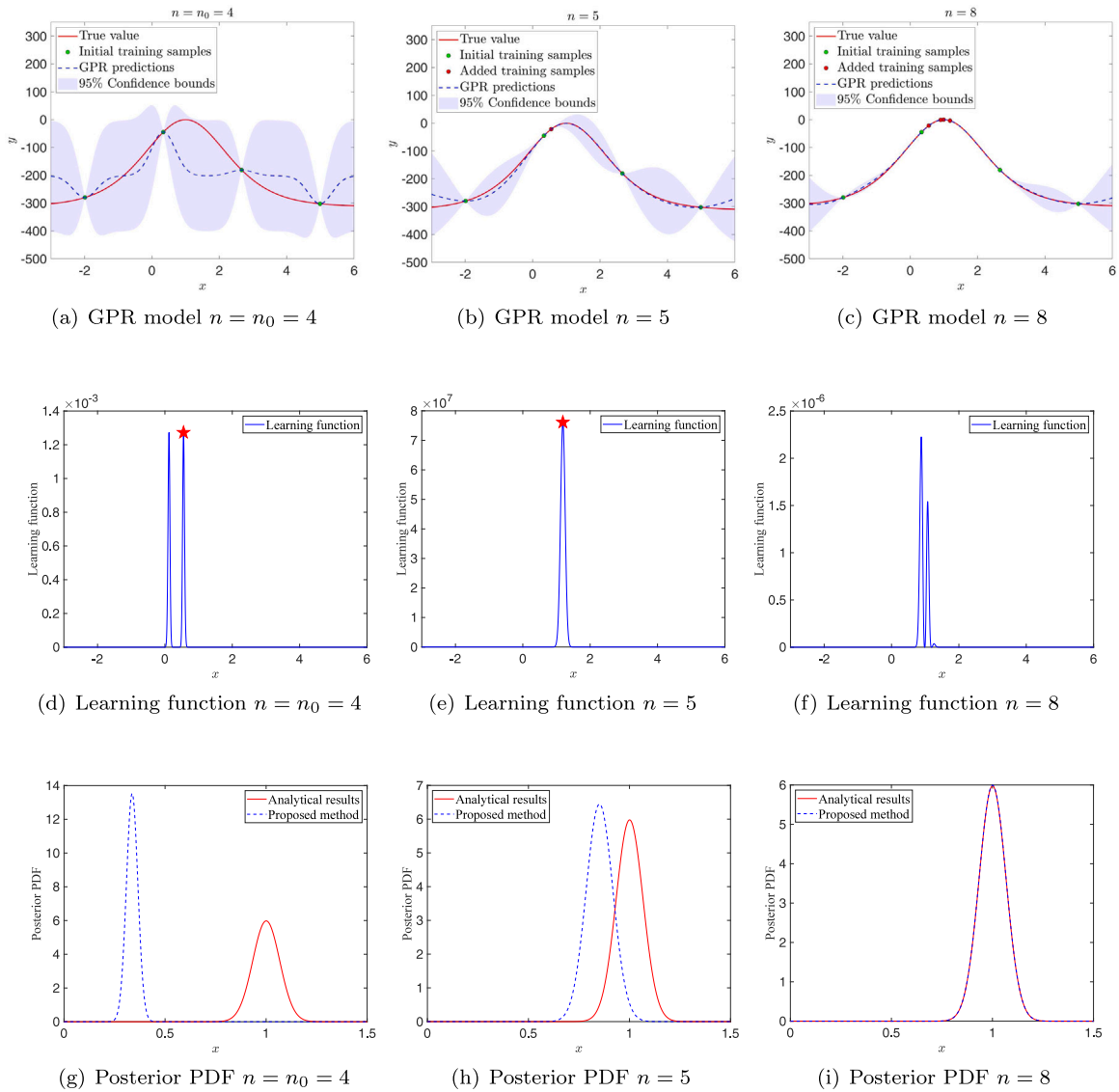


**Fig. 2.** Illustration of the computational details for Example 1.

**Table 1**

Model updating results for Example 1.

| Method          | Time (s) | $N_{call}$ |       | $\hat{\epsilon}$ |       | $\hat{\mu}_X$ |       | $\hat{\sigma}_X$ |       |
|-----------------|----------|------------|-------|------------------|-------|---------------|-------|------------------|-------|
|                 |          | Mean       | CoV   | Mean             | CoV   | Mean          | CoV   | Mean             | CoV   |
| Reference       |          |            |       | 0.0323           | 0.00% | 1.0006        | 0.00% | 0.0668           | 0.00% |
| BUS             | 11.42    | 10 000     | 0.00% | 0.0324           | 4.45% | 1.0010        | 0.87% | 0.0667           | 1.19% |
| TMCMC           | 95.20    | 10 000     | 0.00% | 0.0322           | 2.26% | 1.0002        | 0.83% | 0.0665           | 0.91% |
| ABQ-BU          | 636.24   | 8.6        | 6.00% | 0.0317           | 4.71% | 0.9998        | 0.44% | 0.0694           | 7.79% |
| Proposed method | 21.34    | 8.0        | 0.00% | 0.0323           | 1.55% | 1.0001        | 0.06% | 0.0669           | 0.72% |

#### 4.2. Example 2: Three degree-of-freedom (DOF) mass–spring system

The second example considers a three DOF mass–spring system, as shown in Fig. 3. This system has been widely investigated in previous studies [54–57]. In this system, the nominal mass values are set to  $m_1 = m_2 = m_3 = 1.0$  kg, and the nominal stiffness values are specified as  $k_1 = k_2 = k_5 = 1.5$  N/m,  $k_3 = k_4 = 1.0$  N/m, and  $k_6 = 3.0$  N/m. In this study, the stiffness parameters  $k_1$ ,  $k_2$ , and  $k_5$  are treated as uncertain variables to be updated. Their prior distributions are assumed as Gaussian distribution with mean values

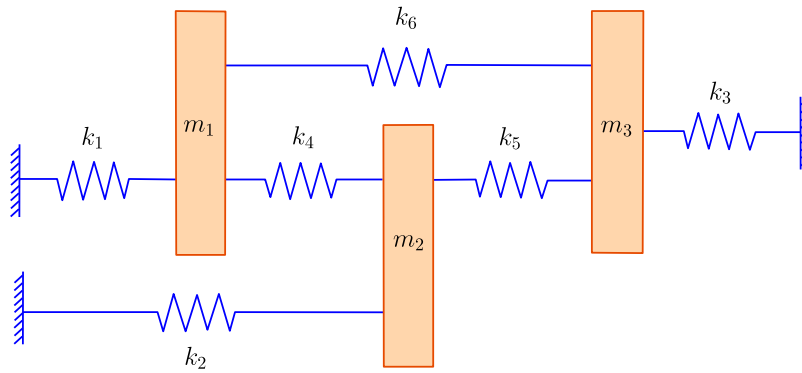


Fig. 3. Three degree-of-freedom spring mass system.

Table 2

Model updating results for Example 2.

| Method                     | BUS                               | TMCMC                             | ABQ-BU                           | Proposed method                  |
|----------------------------|-----------------------------------|-----------------------------------|----------------------------------|----------------------------------|
| Time (s)                   | 4.92                              | 37.06                             | 27 970.15                        | 9.65                             |
| $N_{\text{call}}$ (CoV)    | 10 000                            | 10 000                            | 144.15 (43.66%)                  | 23.3 (3.71%)                     |
| $\hat{c}$ (CoV)            | $3.1478 \times 10^{-14}$ (11.22%) | $3.0598 \times 10^{-14}$ (10.95%) | $3.2542 \times 10^{-14}$ (46.2%) | $3.2107 \times 10^{-14}$ (9.02%) |
| $\hat{\mu}_{k_1}$ (CoV)    | 1.5342 (0.30%)                    | 1.5333 (0.15%)                    | 1.5505 (2.51%)                   | 1.5331 (0.39%)                   |
| $\hat{\mu}_{k_2}$ (CoV)    | 1.4459 (0.39%)                    | 1.4463 (0.19%)                    | 1.4296 (2.62%)                   | 1.4550 (0.46%)                   |
| $\hat{\mu}_{k_5}$ (CoV)    | 1.4439 (0.15%)                    | 1.4441 (1.11%)                    | 1.4467 (1.21%)                   | 1.4550 (0.19%)                   |
| $\hat{\sigma}_{k_1}$ (CoV) | 0.0989 (2.52%)                    | 0.0988 (1.24%)                    | 0.0816 (29.38%)                  | 0.0970 (2.77%)                   |
| $\hat{\sigma}_{k_2}$ (CoV) | 0.1049 (2.33%)                    | 0.1035 (1.35%)                    | 0.0890 (28.26%)                  | 0.1027 (2.77%)                   |
| $\hat{\sigma}_{k_5}$ (CoV) | 0.0501 (1.91%)                    | 0.0496 (1.49%)                    | 0.0416 (26.01%)                  | 0.0501 (2.54%)                   |

Note: Values in parentheses denote CoV obtained from 20 independent runs.

$\mu_{k_1} = \mu_{k_2} = \mu_{k_5} = 2.0$  N/m and standard deviations  $\sigma_{k_1} = \sigma_{k_2} = \sigma_{k_5} = 0.3$  N/m. The uncertainty in  $k_1$ ,  $k_2$ , and  $k_5$  propagates through the system and influences the observed natural frequencies  $f_{\text{obs},1}$ ,  $f_{\text{obs},2}$ ,  $f_{\text{obs},3}$ .

A total of 30 sets of observations are generated using nominal mass values and stiffness parameters  $k_1$ ,  $k_2$ , and  $k_5$ , where the means are set to their nominal stiffness values and the standard deviations to  $\sigma_{k_1} = \sigma_{k_2} = \sigma_{k_5} = 0.2$  N/m. With the above setting, the likelihood function can be expressed as:

$$L(\mathbf{Y}_{\text{obs}} | \mathbf{x}) = \prod_{r=1}^{30} \exp \left( -\frac{1}{2\sigma_e^2} \sum_{i=1}^3 (f_{\text{obs},i}^{(r)} - f_i(\mathbf{x}))^2 \right) = \exp \left( -\frac{1}{2\sigma_e^2} \sum_{r=1}^{30} \sum_{i=1}^3 (f_{\text{obs},i}^{(r)} - f_i(\mathbf{x}))^2 \right), \quad (36)$$

where  $\sigma_e = 0.01$ ,  $N_{\text{obs}} = 30$  for the baseline case,  $f_{\text{obs},i}^{(r)}$  denotes observed value of the  $i$ th natural frequency in the  $r$ th observation set, and  $f_i(\mathbf{x})$  denotes the  $i$ th natural frequency, obtained as the system response for a given realization of  $k_1$ ,  $k_2$ , and  $k_5$ .

To evaluate the performance of the proposed method, commonly used updating techniques, including BUS, TMCMC, and ABQ-BU, are employed to estimate the posterior distribution of  $k_1$ ,  $k_2$ ,  $k_5$ . The comparative results, including computational time, the number of model evaluations, the estimated model evidence  $\hat{c}$ , and the posterior mean and standard deviation of  $k_1$ ,  $k_2$ ,  $k_5$  are summarized in Table 2. From Table 2, it can be seen that the results obtained using the proposed method ( $n_0 = 4$ ) are close to those obtained from the other methods and exhibit a low CoV, indicating the robustness of the proposed method. Notably, the proposed method requires significantly fewer model evaluations (an average of only 23.3) than the other methods and achieves a much shorter computational time (9.65 s), second only to BUS.

After determining the posterior distributions of  $k_1$ ,  $k_2$ ,  $k_5$ , the corresponding samples of the model outputs  $f_1$ ,  $f_2$ ,  $f_3$  can be easily obtained. Fig. 4 illustrates the initial model outputs, the model outputs after updating using the proposed method, and the observed values. As shown in Fig. 4, the outputs of the prior model (in red) display a significantly wider spread and greater uncertainty in all frequency pair spaces compared to the posterior samples (in blue) obtained using the proposed method. After updating, the posterior samples are clustered around the observed natural frequencies (green triangles) and align closely with the nominal value obtained from the corresponding nominal values of the stiffness (orange cross), demonstrating the effectiveness of the proposed method.

To investigate the influence of increasingly informative datasets, an additional sensitivity study is conducted by varying the number of observation sets  $N_{\text{obs}}$  from 1 to 300. Each observation set contains the three natural frequencies  $f_1$ ,  $f_2$ , and  $f_3$ . As shown in Fig. 5, the relative error remains below 0.07 over the entire range of observation sizes, indicating that the proposed method consistently provides accurate Bayesian updating results. Fig. 6 further illustrates the computational performance of the proposed method. The number of model evaluations  $N_{\text{call}}$  remains nearly constant for small observation sizes and increases only when the posterior distribution becomes highly concentrated. Meanwhile, the CoV of the estimated model evidence  $\text{CoV}(\hat{c})$  exhibits an overall increasing trend with moderate fluctuations as  $N_{\text{obs}}$  increases, reflecting the growing difficulty of estimating the model evidence

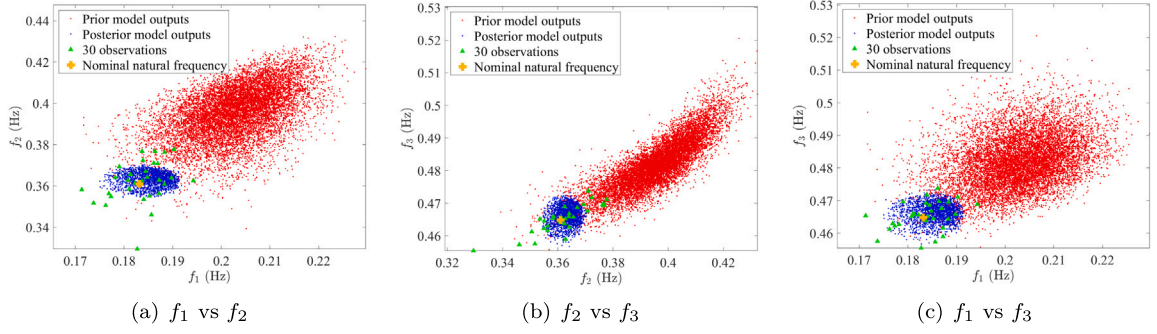


Fig. 4. The observations and the output samples before and after updating for Example 2.

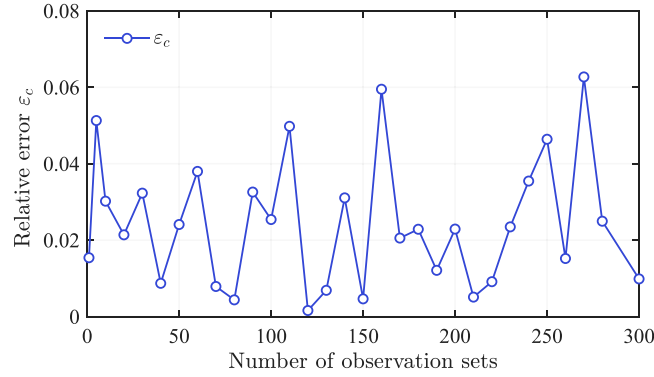


Fig. 5. Relative error of the proposed method.

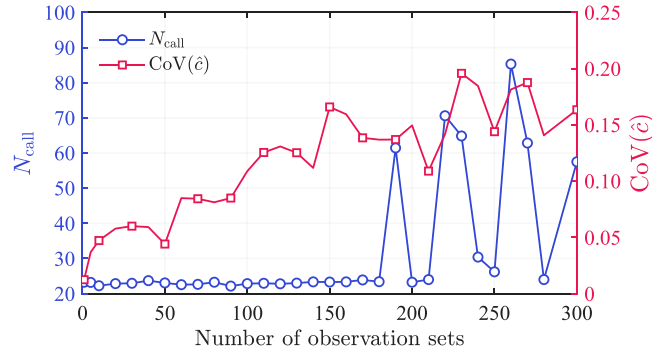


Fig. 6.  $N_{call}$  and  $CoV(\hat{c})$  for the proposed method.

for sharper likelihood functions. Nevertheless, the proposed method maintains high estimation accuracy with a relatively limited number of model evaluations, demonstrating its robustness and computational efficiency under increasingly informative datasets.

#### 4.3. Example 3: Modal analysis for the shoulder link of a lightweight robotic arm

The third example involves the shoulder link of a lightweight robotic arm, was previously investigated by Song et al. [45]. The configuration of this shoulder link is illustrated in Fig. 7. The material of the shoulder link is assumed to be homogeneous, three material properties including Young's modulus ( $E$ ), Poisson's ratio ( $\nu$ ), and mass density ( $\rho$ ) are considered as uncertain parameters. These uncertain parameters influence the natural frequencies and mode shapes of the structure and are subsequently calibrated using Bayesian updating. The model analysis is performed using the PDE toolbox in Matlab. The corresponding finite element mesh and the first five modes of the robotic arm's shoulder link (for nominal values of  $E$ ,  $\nu$  and  $\rho$ ) are shown in Fig. 8.

The prior distribution of uncertain parameters is summarized in Table 3. The first five natural frequencies were observed, namely  $f_{obs,1} = 1947.2$  Hz,  $f_{obs,2} = 2662.0$  Hz,  $f_{obs,3} = 4982.3$  Hz,  $f_{obs,4} = 5112.6$  Hz, and  $f_{obs,5} = 7819.5$  Hz, are used to update the distribution

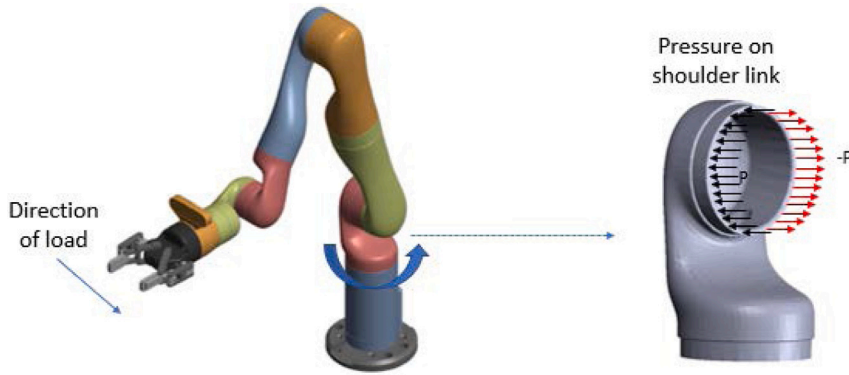


Fig. 7. The shoulder link of a robotic arm.

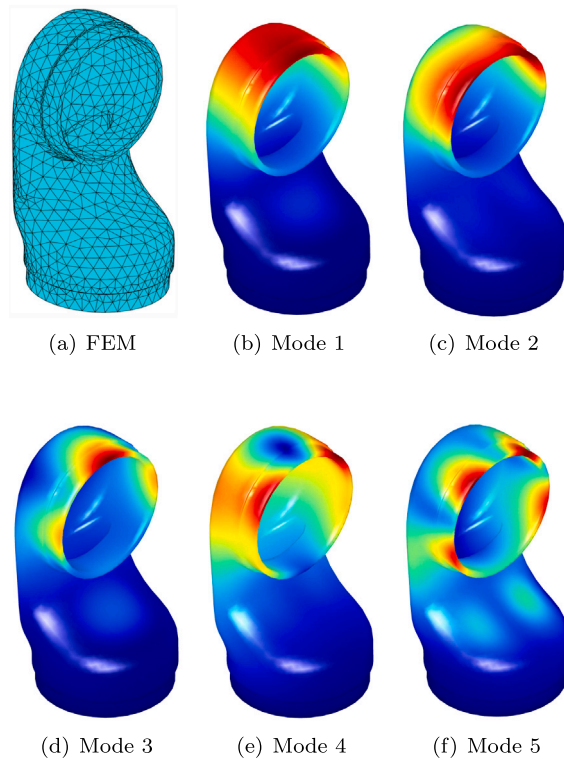


Fig. 8. The finite element mesh and first five modes of the robotic arm's shoulder link.

of  $E$ ,  $\nu$ , and  $\rho$ . Under this setup, the likelihood function is defined as:

$$L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) = \exp\left(-\frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^5 \left(\frac{f_i^2(\mathbf{x})}{f_{\text{obs},i}^2} - 1\right)^2\right), \quad (37)$$

where  $\sigma_\epsilon = 0.2$ , and  $f_i(\mathbf{x})$  is the  $i$ th natural frequency obtained from the model.

Table 4 presents a comparison of the results obtained using BUS, TMCMC, ABQ-BU, and the proposed method. Since this example involves computationally intensive finite element analysis (FEA), BUS and TMCMC were executed once, taking 3.21 and 2.67 days, respectively. In contrast, both ABQ-BU and the proposed method were executed 20 times, with an initial sample size  $n_0 = 5$ . As shown in Table 4, the estimated model evidence, the posterior mean and standard deviation of  $E$ ,  $\nu$ , and  $\rho$  obtained from the proposed method are consistent with those obtained from the other methods. Regarding computational efficiency, although the proposed method requires 10 model evaluations, slightly more than ABQ-BU (which averages 7.10 model evaluations), its total

**Table 3**

Prior distribution of uncertain parameters for Example 3.

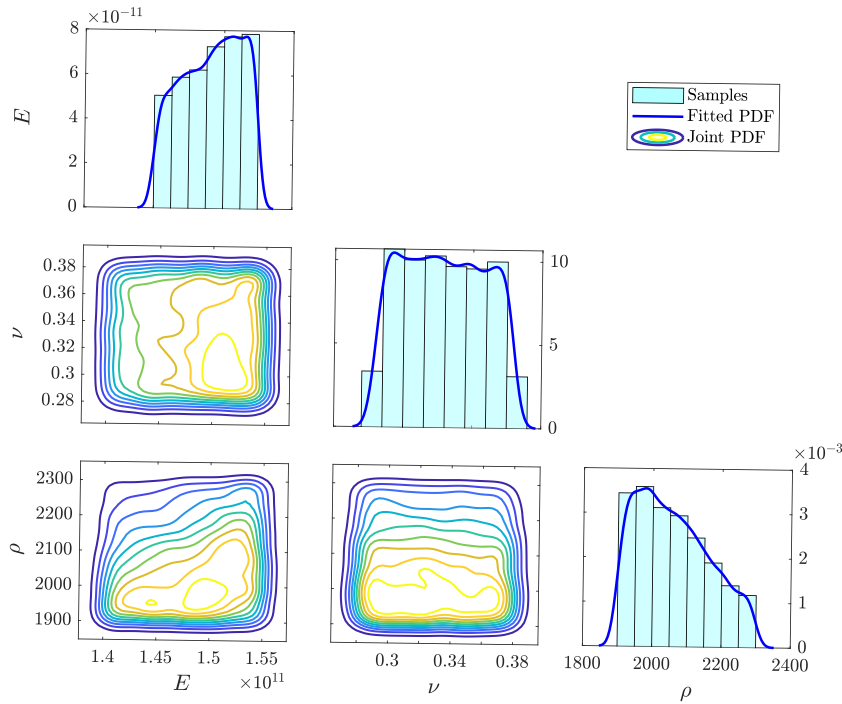
| Parameter | Physical meaning                  | Distribution type | Lower bound | Upper bound |
|-----------|-----------------------------------|-------------------|-------------|-------------|
| $E$       | Young's modulus (GPa)             | Uniform           | 140         | 155         |
| $\nu$     | Poisson's ratio                   | Uniform           | 0.28        | 0.38        |
| $\rho$    | Mass density (kg/m <sup>3</sup> ) | Uniform           | 1900        | 2300        |

**Table 4**

Model updating results for Example 3.

| Method                    | BUS                     | TMCMC                   | ABQ-BU                          | Proposed method                 |
|---------------------------|-------------------------|-------------------------|---------------------------------|---------------------------------|
| Time (s)                  |                         |                         | 2939.62                         | 101.44                          |
| $N_{cal}$ (CoV)           | 10 000                  | 10 000                  | 7.10 (4.34%)                    | 10 (0.00%)                      |
| $\hat{c}$ (CoV)           | 0.6971                  | 0.6358                  | 0.6814 (6.44%)                  | 0.6860 (0.24%)                  |
| $\hat{\mu}_E$ (CoV)       | $1.4811 \times 10^{11}$ | $1.4824 \times 10^{11}$ | $1.4811 \times 10^{11}$ (0.01%) | $1.4812 \times 10^{11}$ (0.03%) |
| $\hat{\mu}_\nu$ (CoV)     | 0.3298                  | 0.3291                  | 0.3293 (0.27%)                  | 0.3294 (0.10%)                  |
| $\hat{\mu}_\rho$ (CoV)    | $2.0605 \times 10^3$    | $2.0591 \times 10^3$    | $2.0639 \times 10^3$ (0.20%)    | $2.0627 \times 10^3$ (0.03%)    |
| $\hat{\sigma}_E$ (CoV)    | $4.2313 \times 10^9$    | $3.9543 \times 10^9$    | $4.2313 \times 10^9$ (0.79%)    | $4.270 \times 10^9$ (0.51%)     |
| $\hat{\sigma}_\nu$ (CoV)  | 0.0288                  | 0.0271                  | 0.0287 (2.52%)                  | 0.0289 (0.43%)                  |
| $\hat{\sigma}_\rho$ (CoV) | 106.0985                | 98.4536                 | 105.1449 (2.43%)                | 105.6342 (0.55%)                |

Note: Values in parentheses denote CoV obtained from 20 independent runs.

**Fig. 9.** Posterior distributions of  $E$ ,  $\nu$ , and  $\rho$  in Example 3.

computational time is significantly lower (only 101.44 s). This improvement is mainly attributed to the fact that the proposed method avoids the need to generate a large number of sample paths from the posterior GP. Furthermore, the proposed method achieves a smaller CoV, which demonstrates its robustness.

Fig. 9 shows the posterior PDFs of  $E$ ,  $\nu$ , and  $\rho$  obtained using the proposed method. The diagonal plots display the marginal posterior PDFs of  $E$ ,  $\nu$ , and  $\rho$ . These PDFs are estimated using the Kernel Density Estimation method [58]. Compared with the corresponding uniform prior distributions, the posterior PDFs become noticeably narrower, indicating that the measured natural frequencies provide informative data for updating the uncertain model parameters and reducing their uncertainties. The degree of uncertainty reduction varies among the three parameters. The posterior distribution of  $\rho$  is the most concentrated, while that of  $E$

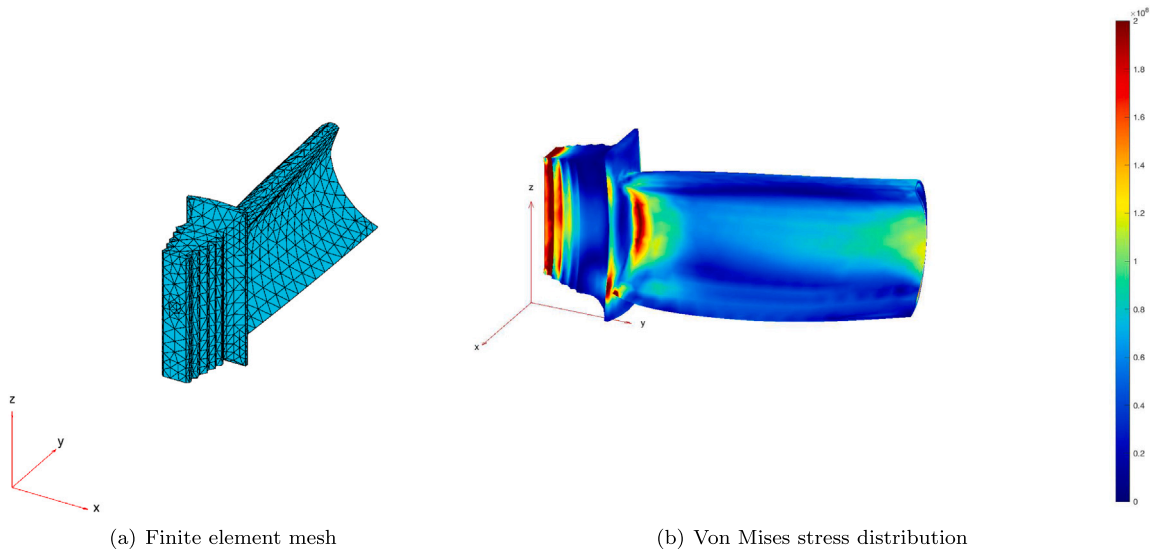


Fig. 10. The finite element mesh and von Mises stress distribution of a jet engine turbine blade.

also exhibits a noticeable reduction in spread. In contrast, the posterior distribution of  $\nu$  remains relatively broad. This behavior suggests that the measured natural frequencies are more sensitive to  $\rho$  and  $E$  than to  $\nu$ . Such a result is physically reasonable because the natural frequencies are primarily governed by the structural mass and stiffness properties, which are directly affected by  $\rho$  and  $E$ , whereas the influence of  $\nu$  on the modal characteristics is comparatively weaker. The lower-triangular contour plots depict the joint posterior PDFs of the parameter pairs. The contours reveal a clear positive posterior dependence between  $E$  and  $\rho$ . This dependence reflects their compensating effects on the natural frequencies: an increase in stiffness associated with a larger  $E$  can be balanced by an increase in mass associated with a larger  $\rho$ . Overall, the posterior probability mass is concentrated within a substantially smaller region than the prior parameter space, demonstrating the effectiveness of the proposed method in reducing parameter uncertainty.

#### 4.4. Example 4: Thermal stress analysis of jet engine turbine blade

The last example investigates a turbine blade structure, a critical component in jet engines. This blade has internal cooling channels through which cool air flows to regulate the blade temperature within the range allowed by the material. Turbine blades are typically made of nickel-based alloys for their excellent resistance to extreme thermal environments. However, prolonged exposure to high temperatures causes thermal expansion, which generates significant mechanical stresses at joints and leads to deformations on the order of millimeters. To accurately capture these effects, thermal analysis of the blade is crucial. The blades were modeled and analyzed using the Partial Differential Equation (PDE) Toolbox in MATLAB. The corresponding finite element mesh and von Mises stress distribution are shown in Fig. 10.

The primary physical quantity of interest in this study is the maximum Von Mises stress of the blade, which is affected by several parameters, including Young's modulus ( $E$ ), the coefficient of thermal expansion ( $\gamma$ ), Poisson's ratio ( $\nu$ ), pressure load on the pressure side ( $P_1$ ), pressure load on the suction side ( $P_2$ ), thermal conductivity ( $K$ ) and the reference temperature ( $T$ ). The objective is to update the prior distributions of these parameters by considering the observed measurements of the blade's maximum displacement. The prior distributions for all parameters are summarized in Table 5. Three displacement observations are used:  $u_{\text{obs},1} = 0.0016$  m,  $u_{\text{obs},2} = 0.0018$  m,  $u_{\text{obs},3} = 0.0020$  m. Based on this setup, the likelihood function is defined as:

$$L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) = \exp \left( -\frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^3 \left( \frac{u^2(\mathbf{x})}{u_{\text{obs},i}^2} - 1 \right)^2 \right), \quad (38)$$

where  $\sigma_\epsilon = 0.2$ , and  $u(\mathbf{x})$  denotes the model-predicted maximum Von Mises stress given the parameter vector  $\mathbf{x}$ .

Table 6 presents a comparative analysis of the results obtained from BUS, TMCMC, and the proposed method. Due to the computational cost of FEA in this example, both BUS and TMCMC were executed only once, requiring 3.15 and 1.83 days, respectively. These results serve as benchmark references. In contrast, the proposed method ( $n_0 = 10$ ) is significantly more efficient, taking only 0.18 h per run. As shown in Table 6, the results produced by the proposed method are consistent with those from BUS and TMCMC and demonstrate low CoV. Moreover, the proposed method requires significantly fewer model evaluations (only 88.25), demonstrating notable computational efficiency.

Fig. 11 displays the posterior distributions of the updated parameters obtained using the proposed method. The diagonal plots show the marginal posterior PDFs, while the lower triangular subplots present contour plots of the joint posterior PDF. Compared

**Table 5**  
Prior distribution of uncertain parameters for Example 4.

| Parameter | Physical meaning                             | Distribution type | Lower bound             | Upper bound             |
|-----------|----------------------------------------------|-------------------|-------------------------|-------------------------|
| $E$       | Young's modulus (Pa)                         | Uniform           | $1.9295 \times 10^{11}$ | $2.6105 \times 10^{11}$ |
| $\gamma$  | Coefficient of thermal expansion (1/K)       | Uniform           | $1.0795 \times 10^{-5}$ | $1.4605 \times 10^{-5}$ |
| $\nu$     | Poisson's ratio                              | Uniform           | 0.2295                  | 0.3105                  |
| $P_1$     | Pressure load on the pressure side (Pa)      | Uniform           | 9.7550                  | 13.2250                 |
| $P_2$     | Pressure load on the suction side (Pa)       | Uniform           | $4.2500 \times 10^5$    | $5.7500 \times 10^5$    |
| $K$       | Thermal conductivity (W/m/K)                 | Uniform           | $3.8250 \times 10^5$    | $5.1750 \times 10^5$    |
| $T$       | Reference temperature ( $^{\circ}\text{C}$ ) | Uniform           | 255                     | 345                     |

**Table 6**  
Model updating results for Example 4.

| Method                      | BUS                     | TMCMC                   | Proposed method                 |
|-----------------------------|-------------------------|-------------------------|---------------------------------|
| Time (s)                    |                         |                         | 13297.08                        |
| $N_{call}$ (CoV)            | 10 000                  | 10 000                  | 88.25 (9.34%)                   |
| $\hat{c}$ (CoV)             | 0.0763                  | 0.0744                  | 0.0737 (1.26%)                  |
| $\hat{\mu}_E$ (CoV)         | $2.267 \times 10^{11}$  | $2.2625 \times 10^{11}$ | $2.2664 \times 10^{11}$ (0.15%) |
| $\hat{\mu}_\gamma$ (CoV)    | $1.3657 \times 10^{-5}$ | $1.3616 \times 10^{-5}$ | $1.3641 \times 10^{-5}$ (0.06%) |
| $\hat{\mu}_\nu$ (CoV)       | 0.2703                  | 0.2700                  | 0.2699 (0.18%)                  |
| $\hat{\mu}_{P_1}$ (CoV)     | 11.3610                 | 11.3752                 | 11.3666 (0.15%)                 |
| $\hat{\mu}_{P_2}$ (CoV)     | $5.0702 \times 10^5$    | $5.0618 \times 10^5$    | $5.0673 \times 10^5$ (0.13%)    |
| $\hat{\mu}_K$ (CoV)         | $4.4474 \times 10^5$    | $4.4403 \times 10^5$    | $4.4477 \times 10^5$ (0.17%)    |
| $\hat{\mu}_T$ (CoV)         | 283.6655                | 285.5074                | 284.4551 (0.13%)                |
| $\hat{\sigma}_E$ (CoV)      | $1.9631 \times 10^{10}$ | $1.8672 \times 10^{10}$ | $1.9566 \times 10^{10}$ (0.75%) |
| $\hat{\sigma}_\gamma$ (CoV) | $6.9606 \times 10^{-7}$ | $6.7455 \times 10^{-7}$ | $7.0857 \times 10^{-7}$ (0.71%) |
| $\hat{\sigma}_\nu$ (CoV)    | 0.0234                  | 0.0220                  | 0.0234 (0.62%)                  |
| $\hat{\sigma}_{P_1}$ (CoV)  | 0.9868                  | 0.9349                  | 0.9930 (0.76%)                  |
| $\hat{\sigma}_{P_2}$ (CoV)  | $4.3234 \times 10^4$    | $4.0860 \times 10^4$    | $4.3241 \times 10^4$ (0.55%)    |
| $\hat{\sigma}_K$ (CoV)      | $3.9265 \times 10^5$    | $3.6431 \times 10^4$    | $3.8866 \times 10^5$ (0.94%)    |
| $\hat{\sigma}_T$ (CoV)      | 20.9581                 | 19.9603                 | 21.3262 (1.25%)                 |

Note: Values in parentheses denote CoV obtained from 20 independent runs.

to their uniform priors, the posteriors of parameters such as  $\gamma$  and  $T$  are significantly more concentrated, indicating that these parameters are more sensitive to the observed data. In contrast, the posterior of  $E$ ,  $\nu$  remains relatively flat, suggesting that they are less sensitive to the observations. The joint distributions further reveal correlations between certain parameters, highlighting their statistical dependencies.

## 5. Summary and conclusions

In this study, a streamlined Bayesian active learning cubature (SBALC) is proposed for Bayesian model updating. In this method, the log-likelihood function is approximated using Gaussian process (GP) regression in a streamlined Bayesian active learning way. Based on the posterior mean function of the log-likelihood function, a plug-in estimator for model evidence is introduced. By deriving an upper bound on the absolute error between the posterior model evidence and the plug-in estimator, the epistemic uncertainty of the plug-in estimator can be quantified only through its upper and lower bounds, thereby eliminating the need for costly posterior GP sampling. Building on this result, a novel stopping criterion is introduced to determine the termination of the iterative process, together with a new learning function for selecting the next most informative evaluation point. As a result, SBALC requires only the posterior mean and standard deviation functions of the GP to obtain a reliable estimate of model evidence, while simultaneously producing an approximate posterior distribution as a by-product. The performance of the proposed method is demonstrated by four numerical examples, in comparison with several existing methods. The results indicate that SBALC substantially reduces both the number of model evaluations and computational time, while maintaining high estimation accuracy and robustness. This advantage is particularly evident in applications involving computationally intensive model analyses.

It is acknowledged that SBALC is currently limited to solving the likelihood function of moderate complexity. For highly non-linear or non-smooth likelihood function, particularly in high-dimensional settings, both the computational cost and the estimation error tend to increase significantly. Future work will focus on enhancing SBALC by adopting more efficient stochastic simulation techniques to better address highly non-linear or non-smooth likelihood function, and by integrating advanced dimension-reduction strategies to cope with high-dimensional problems. Moreover, efficiency can be further improved by evaluating multiple points per iteration, thereby enabling effective use of parallel computing resources.

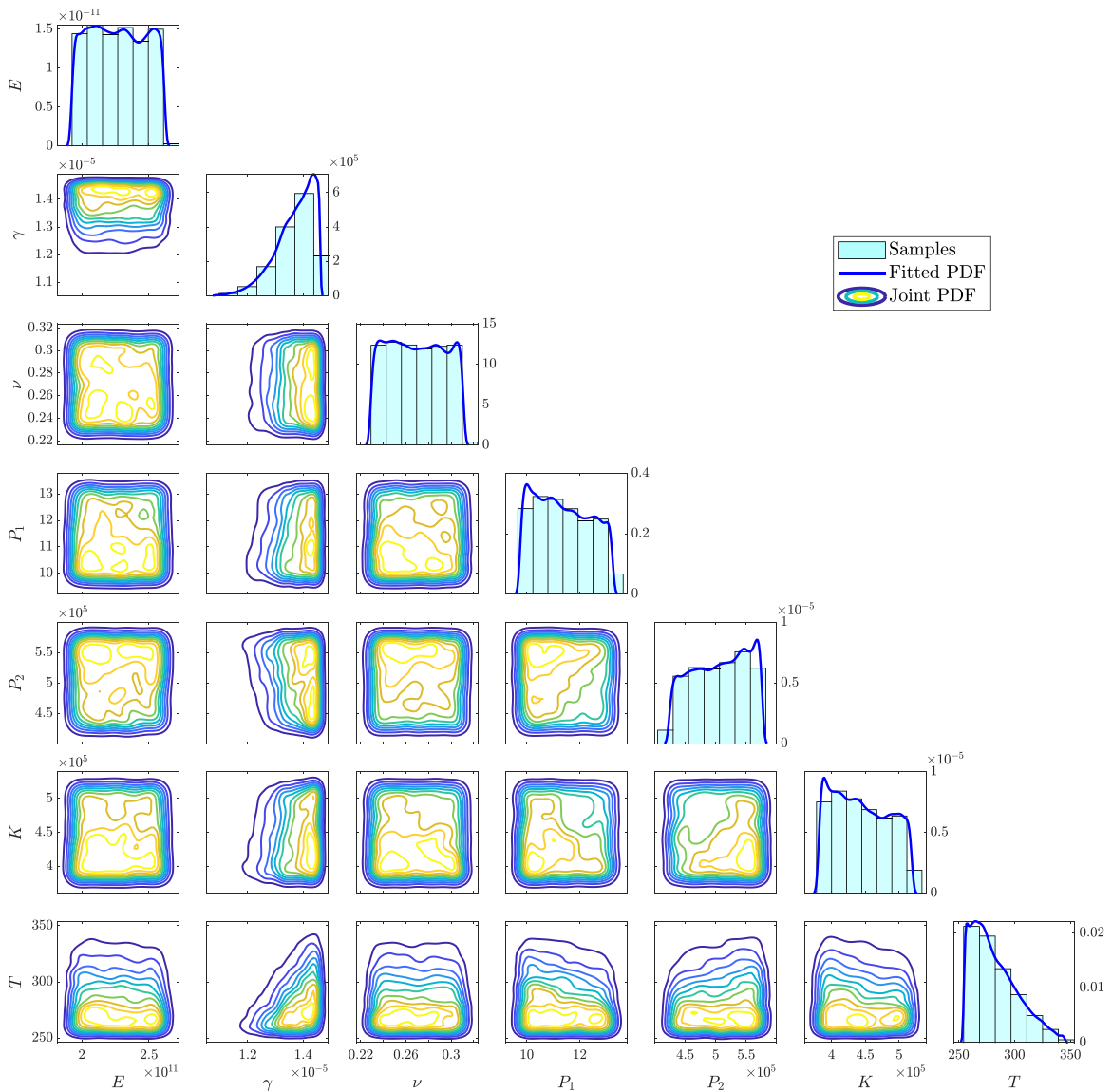


Fig. 11. Posterior distributions of updated parameters in Example 4.

### CRediT authorship contribution statement

**Pei-Pei Li:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Funding acquisition. **Chao Dang:** Writing – review & editing, Software, Methodology, Investigation, Conceptualization. **Cristóbal H. Acevedo:** Writing – review & editing, Methodology, Funding acquisition. **Marcos A. Valdebenito:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Matthias G.R. Faes:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

The authors gratefully acknowledge the support of the Alexander von Humboldt Foundation for the postdoctoral grant of Pei-Pei Li, the Henriette Herz Scouting program (Matthias G.R. Faes), the State Key Laboratory of Disaster Reduction in Civil Engineering,

Tongji University (Grant number SLDRCE24-02), the financial support of the German Research Foundation (DFG) for Chao Dang (Grant number 530326817), and the financial support provided through the project “Digitization of Infrastructure Structures for Structural Health Monitoring (SHM-Digitalisierung und Überwachung von Infrastrukturbauwerken)” within dtcc.bw-Digitalization and Technology Research Center of the Bundeswehr, which is funded by the European Union-NextGenerationEU.

## Appendix. Technical details in the proof of Proposition 2

This appendix provides the detailed derivation of the expectation term  $\mathbb{E}_Z[|e^{aZ} - 1|]$  and the proof of the boundedness of the auxiliary function  $h_b(\cdot)$ , which are used in Proposition 2.

### A.1. Evaluation of $\mathbb{E}_Z[|e^{aZ} - 1|]$

The remaining expectation in Eq. (20) can be evaluated as:

$$\mathbb{E}_Z[|e^{aZ} - 1|] = \int_0^\infty (e^{az} - 1)\phi(z) dz + \int_{-\infty}^0 (1 - e^{az})\phi(z) dz, \quad (39)$$

where  $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$  denotes the PDF of  $Z$ .

Applying the change of variable  $y = -z$  to the second integral in Eq. (39), it follows that

$$\int_{-\infty}^0 (1 - e^{az})\phi(z) dz = \int_0^\infty (1 - e^{-ay})\phi(y) dy. \quad (40)$$

Substituting Eq. (40) into Eq. (39) and using the relation  $\int_0^\infty \phi(z) dz = \frac{1}{2}$ , we obtain:

$$\mathbb{E}_Z[|e^{aZ} - 1|] = \underbrace{\int_0^\infty e^{az}\phi(z) dz}_{=: I_+(a)} - \underbrace{\int_0^\infty e^{-az}\phi(z) dz}_{=: I_-(a)}, \quad (41)$$

where the integrals  $I_+(a)$  and  $I_-(a)$  can be evaluated explicitly as follows:

$$I_+(a) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(z^2/2)+az} dz = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{(z-a)^2}{2} + \frac{a^2}{2}} dz = e^{a^2/2} \Phi(a), \quad (42)$$

and

$$I_-(a) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(z^2/2)-az} dz = e^{a^2/2} \Phi(-a), \quad (43)$$

respectively.

Substituting Eqs. (42) and (43) into Eq. (41) yields

$$\mathbb{E}_Z[|e^{aZ} - 1|] = e^{a^2/2} (\Phi(a) - \Phi(-a)) = e^{a^2/2} (2\Phi(a) - 1). \quad (44)$$

This result is used in Eq. (21) of the main text.

### A.2. Boundedness of $h_b(\cdot)$

Recall that  $h_b(\sigma_{\mathcal{L}_n}(\mathbf{x}))$  is defined in Eq. (26) of the main text. For the Gaussian kernel, there exists a constant  $\sigma_0 > 0$  such that  $\sigma_{\mathcal{L}_n}(\mathbf{x}) \in [0, \sigma_0]$  for all  $\mathbf{x} \in D_{\mathbf{X}}$ . To establish the boundedness of  $h_b(\sigma_{\mathcal{L}_n}(\mathbf{x}))$  on  $[0, \sigma_0]$ , we first examine its behavior near the origin. Using the first-order expansions  $\Phi(t) = \frac{1}{2} + \frac{t}{\sqrt{2\pi}} + o(t)$  and  $\sinh(z) = z + o(z)$  as  $t, z \rightarrow 0$ , we obtain

$$\lim_{\sigma_{\mathcal{L}_n}(\mathbf{x}) \rightarrow 0} h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) = \lim_{s \rightarrow 0} \frac{e^{s^2/2} (2\Phi(s) - 1)}{2 \sinh(bs)} = \frac{\sqrt{2/\pi}}{2b}. \quad (45)$$

Therefore,  $h_b$  admits a continuous extension at  $\sigma_{\mathcal{L}_n}(\mathbf{x}) = 0$ . Next, for  $\sigma_{\mathcal{L}_n}(\mathbf{x}) > 0$ , applying the inequality  $2\Phi(t) - 1 \leq \min\{\sqrt{2/\pi} t, 1\}$  for  $t \geq 0$  together with  $\sinh(z) \geq z$  ( $z \geq 0$ ), yields

$$h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) \leq \begin{cases} \frac{e^{1/2}}{2b} \sqrt{\frac{2}{\pi}}, & 0 \leq \sigma_{\mathcal{L}_n}(\mathbf{x}) \leq 1, \\ \frac{e^{\sigma_0^2/2}}{2 \sinh(b)}, & 1 \leq \sigma_{\mathcal{L}_n}(\mathbf{x}) \leq \sigma_0. \end{cases} \quad (46)$$

Hence,

$$h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) \leq \max \left\{ \frac{e^{1/2}}{2b} \sqrt{\frac{2}{\pi}}, \frac{e^{\sigma_0^2/2}}{2 \sinh(b)} \right\}, \quad \text{for all } s \in (0, \sigma_0]. \quad (47)$$

Combining this bound with the continuity of  $h_b$  at  $s = 0$ , we conclude that  $h_b$  is bounded on the interval  $[0, \sigma_0]$ . Consequently,

$$M_b := \sup_{\sigma_{\mathcal{L}_n}(\mathbf{x}) \in [0, \sigma_0]} h_b(\sigma_{\mathcal{L}_n}(\mathbf{x})) \leq \max \left\{ \frac{e^{1/2}}{2b} \sqrt{\frac{2}{\pi}}, \frac{e^{\sigma_0^2/2}}{2 \sinh(b)} \right\}. \quad (48)$$

This establishes the boundedness required in Proposition 2.

## Data availability

No data was used for the research described in the article.

## References

- [1] M.I. Friswell, J.E. Mottershead, Finite element modelling, in: *Finite Element Model Updating in Structural Dynamics*, Springer, 1995, pp. 7–35.
- [2] J.D. Collins, G.C. Hart, T. Hasselman, B. Kennedy, Statistical identification of structures, *AIAA J.* 12 (2) (1974) 185–190.
- [3] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method in finite element model updating: A tutorial, *Mech. Syst. Signal Process.* 25 (7) (2011) 2275–2296.
- [4] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *J. Eng. Mech.* 124 (4) (1998) 455–461.
- [5] L.S. Katafygiotis, J.L. Beck, Updating models and their uncertainties. II: Model identifiability, *J. Eng. Mech.* 124 (4) (1998) 463–467.
- [6] N. Bleistein, R.A. Handelsman, *Asymptotic Expansions of Integrals*, Ardent Media, 1975.
- [7] R.E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.* 82 (1) (1960) 35–45.
- [8] R.E. Kalman, R.S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng.* 83 (1) (1961) 95–108.
- [9] A. Gelb, et al., *Applied Optimal Estimation*, MIT Press, 1974.
- [10] B.D. Anderson, J.B. Moore, *Optimal Filtering*, Courier Corporation, 2005.
- [11] S.J. Julier, J.K. Uhlmann, New extension of the Kalman filter to nonlinear systems, in: *Signal Processing, Sensor Fusion, and Target Recognition VI*, vol. 3068, Spie, 1997, pp. 182–193.
- [12] C.W. Fox, S.J. Roberts, A tutorial on variational Bayesian inference, *Artif. Intell. Rev.* 38 (2) (2012) 85–95.
- [13] P. Ni, J. Li, H. Hao, Q. Han, X. Du, Probabilistic model updating via variational Bayesian inference and adaptive Gaussian process modeling, *Comput. Methods Appl. Mech. Engrg.* 383 (2021) 113915.
- [14] X. Jia, W.-J. Yan, C. Papadimitriou, K.-V. Yuen, An analytically tractable solution for hierarchical Bayesian model updating with variational inference scheme, *Mech. Syst. Signal Process.* 189 (2023) 110060.
- [15] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* 21 (6) (1953) 1087–1092.
- [16] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* 57 (1) (1970) 97–109.
- [17] J.L. Beck, S.-K. Au, Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, *J. Eng. Mech.* 128 (4) (2002) 380–391.
- [18] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.* (6) (1984) 721–741.
- [19] R.M. Neal, et al., MCMC using Hamiltonian dynamics, *Handb. Markov Chain Monte Carlo* 2 (11) (2011) 2.
- [20] H.A. Jensen, D.J. Jerez, C. Figueroa, On the use of reliability methods and Hamiltonian Monte Carlo for complex identification problems in structural dynamics, *Mech. Syst. Signal Process.* 229 (2025) 112547.
- [21] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, *J. Eng. Mech.* 133 (7) (2007) 816–832.
- [22] N.J. Gordon, D.J. Salmond, A.F. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, in: *IEE Proceedings F (Radar and Signal Processing)*, vol. 140, (2) IET, 1993, pp. 107–113.
- [23] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamic systems, *J. Amer. Statist. Assoc.* 93 (443) (1998) 1032–1044.
- [24] D. Straub, I. Papaioannou, Bayesian updating with structural reliability methods, *J. Eng. Mech.* 141 (3) (2015) 04014134.
- [25] F. DiazDelaO, A. Garbuno-Inigo, S. Au, I. Yoshida, Bayesian updating and model class selection with subset simulation, *Comput. Methods Appl. Mech. Engrg.* 317 (2017) 1102–1121.
- [26] Q. Li, P. Ni, X. Du, Q. Han, K. Xu, Y. Bai, Bayesian finite element model updating with a variational autoencoder and polynomial chaos expansion, *Eng. Struct.* 316 (2024) 118606.
- [27] M.J. Atalla, D.J. Inman, On model updating using neural networks, *Mech. Syst. Signal Process.* 12 (1) (1998) 135–161.
- [28] H.-P. Wan, W.-X. Ren, Stochastic model updating utilizing Bayesian approach and Gaussian process model, *Mech. Syst. Signal Process.* 70 (2016) 245–268.
- [29] M. Seeger, Gaussian processes for machine learning, *Int. J. Neural Syst.* 14 (02) (2004) 69–106.
- [30] B. Echard, N. Gayton, M. Lemaire, AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation, *Struct. Saf.* 33 (2) (2011) 145–154.
- [31] C. Dang, P. Wei, J. Song, M. Beer, Estimation of failure probability function under imprecise probabilities by active learning-augmented probabilistic integration, *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A: Civ. Eng.* 7 (4) (2021) 04021054.
- [32] C. Dang, P. Wei, M.G. Faes, M.A. Valdebenito, M. Beer, Parallel adaptive Bayesian quadrature for rare event estimation, *Reliab. Eng. Syst. Saf.* 225 (2022) 108621.
- [33] Z. Wang, A. Shafieezadeh, Highly efficient Bayesian updating using metamodels: An adaptive kriging-based approach, *Struct. Saf.* 84 (2020) 101915.
- [34] Y. Liu, L. Li, S. Zhao, Efficient Bayesian updating with two-step adaptive kriging, *Struct. Saf.* 95 (2022) 102172.
- [35] M. Kitahara, C. Dang, M. Beer, Bayesian updating with two-step parallel Bayesian optimization and quadrature, *Comput. Methods Appl. Mech. Engrg.* 403 (2023) 115735.
- [36] Z.-H. Lu, W.-T. Pei, Z. Zhao, Z. Qian, Efficient Bayesian updating with single-loop kriging model for time-dependent model calibration, *Comput. Methods Appl. Mech. Engrg.* 444 (2025) 118150.
- [37] P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, X-TMCMC: Adaptive kriging for Bayesian inverse modeling, *Comput. Methods Appl. Mech. Engrg.* 289 (2015) 409–428.
- [38] I. Yoshida, T. Nakamura, S.-K. Au, Bayesian updating of model parameters using adaptive Gaussian process regression and particle filter, *Struct. Saf.* 102 (2023) 102328.
- [39] A.A. Taflanidis, B. Aakash, S.-r. Yi, J.P. Conte, Surrogate-aided Bayesian calibration with adaptive learning strategies, *Mech. Syst. Signal Process.* 237 (2025) 113014.

- [40] A. O'Hagan, Bayes–Hermite quadrature, *J. Statist. Plann. Inference* 29 (3) (1991) 245–260.
- [41] M. Osborne, D. Duvenaud, R. Garnett, C. Rasmussen, S. Roberts, Z. Ghahramani, Active learning of model evidence using Bayesian quadrature, *Adv. Neural Inf. Process. Syst.* 25 (2012) 46–54.
- [42] T. Gunter, M. Osborne, R. Garnett, P. Hennig, S. Roberts, Sampling for inference in probabilistic models with fast Bayesian quadrature, *Adv. Neural Inf. Process. Syst.* 27 (2014) 2789–2797.
- [43] H. Wang, J. Li, Adaptive Gaussian process approximation for Bayesian inference with expensive likelihood functions, *Neural Comput.* 30 (11) (2018) 3072–3094.
- [44] P. Wei, M. Kitahara, M.G. Faes, M. Beer, Probabilistic calibration of model parameters with approximate Bayesian quadrature and active machine learning, *J. Reliab. Sci. Eng.* 1 (2024) 015003.
- [45] J. Song, Z. Liang, P. Wei, M. Beer, Sampling-based adaptive Bayesian quadrature for probabilistic model updating, *Comput. Methods Appl. Mech. Engrg.* 433 (2025) 117467.
- [46] M. Kitahara, T. Kitahara, Sequential and adaptive probabilistic integration for Bayesian model updating, *Mech. Syst. Signal Process.* 223 (2025) 111825.
- [47] P. Wei, Bayesian model inference with complex posteriors: Exponential-impact-informed Bayesian quadrature, *Mech. Syst. Signal Process.* 239 (2025) 113333.
- [48] P.-P. Li, Y. Zhang, Y.-G. Zhao, Z. Zhao, E. Cai, An information reuse-based method for reliability updating, *Reliab. Eng. Syst. Saf.* 239 (2023) 109536.
- [49] G.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons, 1999.
- [50] C. Zhong, G. Li, Z. Meng, H. Li, A.R. Yildiz, S. Mirjalili, Starfish optimization algorithm (SFOA): a bio-inspired metaheuristic algorithm for global optimization compared with 100 optimizers, *Neural Comput. Appl.* 37 (5) (2025) 3641–3683.
- [51] A.F. Smith, A.E. Gelfand, Bayesian statistics without tears: a sampling–resampling perspective, *Amer. Statist.* 46 (2) (1992) 84–88.
- [52] W. Betz, I. Papaioannou, D. Straub, Transitional Markov chain Monte Carlo: observations and improvements, *J. Eng. Mech.* 142 (5) (2016) 04016016.
- [53] W. Betz, I. Papaioannou, J.L. Beck, D. Straub, Bayesian inference with subset simulation: strategies and improvements, *Comput. Methods Appl. Mech. Engrg.* 331 (2018) 72–93.
- [54] S. Bi, M. Broggi, M. Beer, The role of the bhattacharyya distance in stochastic model updating, *Mech. Syst. Signal Process.* 117 (2019) 437–452.
- [55] C. Mares, J. Mottershead, M. Friswell, Stochastic model updating: part 1—theory and simulated example, *Mech. Syst. Signal Process.* 20 (7) (2006) 1674–1695.
- [56] H.H. Khodaparast, J.E. Mottershead, M.I. Friswell, Perturbation methods for the estimation of parameter variability in stochastic model updating, *Mech. Syst. Signal Process.* 22 (8) (2008) 1751–1773.
- [57] E. Patelli, Y. Govers, M. Broggi, H.M. Gomes, M. Link, J.E. Mottershead, Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data, *Arch. Appl. Mech.* 87 (2017) 905–925.
- [58] M.P. Wand, M.C. Jones, *Kernel Smoothing*, CRC Press, 1994.