A Parallel Learning Strategy for Adaptive Kriging-based Reliability Analysis Methods

Zeyu Wang
PhD Candidate, the Ohio State University, Columbus, USA

Abdollah Shafieezadeh
Associate Professor, the Ohio State University, Columbus, USA

ABSTRACT: Adaptive Kriging-based reliability analysis methods have shown great advantages over conventional methods for their computational efficiency and accuracy. However, the widely accepted learning strategies such as Expected Feasibility function and U function can select one training point for each iteration, and therefore are not suitable for parallel processing. To address this limitation, the uncertainty of the failure probability is estimated through Adaptive Kriging with probabilistic classification-based Monte Carlo simulation based on the fact that the total number of failure points follows a Poisson Binomial distribution. By maximally reducing the uncertainty of the estimated failure probability, the theoretically optimal learning strategy is derived in this paper. Due to the computational difficulty in implementing the optimal learning strategy, a pseudo optimal parallel learning strategy is proposed to closely reach the optimal solution. The efficiency of the proposed parallel learning strategy is investigated here by implementing two benchmark reliability problems. Results indicate that the total number of evaluations to the performance function through the proposed parallel learning strategy can be even close to the approach based on single training point enriching.

1. INTRODUCTION

Structures and infrastructure systems are vulnerable against gradual and shock-type disturbances such as aging, and natural and manmade hazards. Probabilities of occurrence of the failure of these systems are key for quantifying and managing risks. These probabilities are typically characterized by conducting reliability analysis, where the target is to estimate the probability of failure, denoted as P_f. In the analysis of reliability, P_f can be calculated as:

$$P_f = \int_{g(\boldsymbol{x})\leq 0} \rho(\boldsymbol{x})d\boldsymbol{x} = \int_{\Omega} I_g(\boldsymbol{x})\rho(\boldsymbol{x})d\boldsymbol{x}, \quad (1)$$

where $\boldsymbol{x}$ is the vector of random variables, $g(\boldsymbol{x})$ is the so-called performance or limit state function, $\rho(\boldsymbol{x})$ is the joint probability density function (PDF) of $\boldsymbol{x}$ and $I_g(\boldsymbol{x})$ is a failure indicator function. Note that $I_g(\boldsymbol{x}) = 1$ when $g(\boldsymbol{x}) \leq 0$, and $I_g(\boldsymbol{x}) = 0$ when $g(\boldsymbol{x}) > 0$. Modern reliability analysis methods aim to estimate P_f with as few as possible number of evaluations to the computationally demanding numerical models. Those well-developed techniques primarily include simulation-based sampling techniques (e.g., the crude Monte-Carlo simulation (Rubinstein and Kroese 2016), importance sampling (Echard et al. 2013), and subset simulation (Huang, Chen, and Zhu 2016)) and approximation-based approaches (e.g. first- or second-order reliability analysis methods) (Ditlevsen and Madsen 1996; Lemaire 2013)), which estimate P_f by searching for the most probable point in the probabilistic space. Although the first group can produce desirable estimates of P_f, they are computationally expensive. On the contrary, approximation-based methods are often computationally fast, but they lack the accuracy for problems with non-linear responses near the limit state. To address these

limitations, surrogate model-based reliability analysis has emerged and shown a great potential. These surrogate models typically include polynomial chaos expansion (Blatman and Sudret 2010) and Kriging (Echard, Gayton, and Lemaire 2011; Bichon et al. 2008). Among these techniques, Kriging-based methods have been developing fast (Echard, Gayton, and Lemaire 2011; Wang and Shafieezadeh 2019).

Different from other deterministic regression models, outputs from the Kriging follow Gaussian distribution. The well-trained surrogate model can substitute the sophisticated model for the process of Monte Carlo simulations. Two widely used active learning-based reliability analysis methods are: Efficient Global Reliability Analysis (EGRA) proposed by Bichon et al. (2008) and Adaptive Kriging-based Monte Carlo simulation (AK-MCS) proposed by Echard, Gayton, and Lemaire (2011). It is known that the learning strategy for properly selecting training points plays a decisive role in Kriging-based reliability analysis methods. Aside from EFF (Expected Feasibility Function) and U learning functions, a lot of creative developments and variations of the aforementioned two methods have also been proposed. *LIF* (Least Improvement Function) takes advantage of the probability density of each point to prioritize training points with large probability densities (Sun et al. 2017). $\boldsymbol{\psi}_d$ and $\boldsymbol{\psi}_\sigma$ pick the best training point that is far away from the existing training points to avoid the ill-conditioning problem, and is close to the limit state with high variance. It is shown that these functions are efficient in strategically picking design points (Echard, Gayton, and Lemaire 2011). However, they can select only one training point upon each learning iteration, which is not flexible and does not allow parallel processing. Hence, a number of parallel training point-enriching strategies, such as k-means clustering, have been proposed by Lelièvre et al. (2018). However, these parallel learning strategies are not optimal in selection of multiple training points.

In this article, the theoretically optimal learning strategy is derived and a new learning strategy for enriching multiple training points is proposed to overcome aforementioned drawbacks. To explore such optimal learning strategy, uncertainty of the failure probability estimated through adaptive Kriging-based reliability methods is derived based on the fact that the total number of failure points estimated through probabilistic classification-based Monte Carlo simulation follows a Poisson Binomial distribution. This estimate of the uncertainty is used to derive the optimal learning strategy that identifies training points that maximally reduce the uncertainty in failure probability estimates.

This article is organized as follows. Section 2 briefly introduces the Kriging model and probabilistic classification-based Monte Carlo simulation. In Section 3, the proposed parallel learning strategy is elaborated. In Section 4, an example is investigated to evaluate the proposed strategy. Conclusions are drawn in Section 5.

2. KRIGING MODEL

In Kriging surrogate model-based reliability analysis, the performance or limit state function, $g(\boldsymbol{x})$ is substituted by Kriging surrogate model $\hat{g}(\boldsymbol{x})$ that is adaptively trained ("UQLab Kriging (Gaussian Process Modelling) Manual" n.d.). In this section, Kriging model with uncorrelated and correlated responses are briefly introduced (Echard, Gayton, and Lemaire 2011; Wang et al. 2017). Generally, Kriging model $\hat{g}(\boldsymbol{x})$ can be represented as:

$$\hat{g}(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{\beta}) + Z(\boldsymbol{x}) = \boldsymbol{f}^T(\boldsymbol{x})\boldsymbol{\beta} + Z(\boldsymbol{x}), \quad (2)$$

where $F(\boldsymbol{x}, \boldsymbol{\beta})$ is the regression base representing the Kriging trend, which can be a constant or a polynomial. $\boldsymbol{f}(\boldsymbol{x})$ is the Kriging basis and $\boldsymbol{\beta}$ is the regression coefficients. $\boldsymbol{f}^T(\boldsymbol{x})\boldsymbol{\beta}$ usually have ordinary (β_0), linear ($\beta_0+\sum_{n=1}^{N}\beta_n x_n$) or quadratic ($\beta_0 + \sum_{n=1}^{N}\beta_n x_n + \sum_{n=1}^{N}\sum_{k=n}^{N}\beta_{nk} x_n x_k$) forms, whereas n is the dimension of the random input vector, $\boldsymbol{x}$. Note that ordinary Kriging is used entirely in this paper. $Z(\boldsymbol{x})$ is the Kriging interpolation following a stationary normal Gaussian process with zero mean and a

covariance matrix between two points, $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$, as defined below:

$$\mathrm{COV}\left(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)\right) = \sigma^2 R(\boldsymbol{x_i}, \boldsymbol{x_j}; \boldsymbol{\theta}), \quad (3)$$

where σ^2 is the process variance or the generalized mean square error from the regression part and $R(\boldsymbol{x}_i, \boldsymbol{x}_j; \boldsymbol{\theta})$ is the correlation function or the kernel function representing the correlation function of the process with hyper-parameter $\boldsymbol{\theta}$. Multiple types of correlation functions are available in Kriging model including linear, exponential, Gaussian, Matérn models, among others ("UQLab Kriging (Gaussian Process Modelling) Manual" n.d.). In this paper, the Gaussian kernel function is implemented:

$$R(\boldsymbol{x}_i, \boldsymbol{x}_j; \boldsymbol{\theta}) = \prod_{n=1}^{N} \exp\left(-\theta^n \left(x_i^n - x_j^n\right)^2\right), \quad (4)$$

where N is the dimension of the random input vector, $\boldsymbol{x}_i$ or $\boldsymbol{x}_j$. The hyper-parameter $\boldsymbol{\theta}$ can be estimated via maximum likelihood estimation (MLE) or cross-validation ("UQLab Kriging (Gaussian Process Modelling) Manual" n.d.). It has been shown that the Kriging prediction is very sensitive to the value of $\boldsymbol{\theta}$ (Kaymaz 2005; Wang et al. 2017). In this article, the well-known DACE toolbox is implemented (Lophaven, Nielsen, and Søndergaard 2002b), (Lophaven, Nielsen, and Søndergaard 2002a), where θ_i is searched in $(0,10)$. MLE is aimed to search for:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \left(\left| \boldsymbol{R}(\boldsymbol{x}_i, \boldsymbol{x}_j; \boldsymbol{\theta}) \right|^{\frac{1}{m}} \sigma^2 \right), \quad (5)$$

where m is the number of known training points or design-of-experiment (DoE) points. The Kriging model with correlated responses can be represented as (Wang et al. 2017):

$$\boldsymbol{Y}_U \sim N(\boldsymbol{\mu}_U, \boldsymbol{\Sigma}_U), \quad (6)$$

where the Kriging mean can be expressed in a matrix form:

$$\boldsymbol{\mu}_U = \boldsymbol{F}_U \boldsymbol{\beta} + \boldsymbol{r}_U^T \boldsymbol{R}^{-1} (\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}), \quad (7)$$

and the corresponding covariance matrix can be read as:

$$\boldsymbol{\Sigma} = \sigma^2 (\boldsymbol{R}_U + \boldsymbol{u}_U^T (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{u}_U - \boldsymbol{r}_U^T \boldsymbol{R}^{-1} \boldsymbol{r}_U), \quad (8)$$

where

$$\boldsymbol{R}_U = \left(\boldsymbol{R}(\boldsymbol{x}_p^u, \boldsymbol{x}_q^u; \boldsymbol{\theta}) \right)_{N_u \times N_u}^T, \boldsymbol{x}_p^u, \boldsymbol{x}_q^u \in S_U,$$

$$\boldsymbol{r}_U = \left(\boldsymbol{R}(\boldsymbol{x}_l, \boldsymbol{x}_p^u; \boldsymbol{\theta}) \right)_{N_u \times m}^T, \boldsymbol{x}_l \in S_{DoE},$$

$$\boldsymbol{u}_U = \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r}_U - \boldsymbol{F}_U^{\boldsymbol{T}},$$

$$\boldsymbol{F}_U(\boldsymbol{x}) = \left[f(\boldsymbol{x}_1^u), f(\boldsymbol{x}_2^u), \ldots f\left(\boldsymbol{x}_{N_u}^u\right) \right]^T. \quad (9)$$

Different from the Kriging model with uncorrelated responses, Kriging model with correlated responses follow the assumption in building the Kriging model that all the known and untried responses follow mutually correlated normal distribution. It is shown that Kriging model with correlated responses works better than that with uncorrelated responses in terms of estimating the confidence interval of failure probability (Wang and Shafieezadeh, 2019). Steps for adaptive Kriging-based reliability analysis are available in a number of studies including (Echard, Gayton, and Lemaire 2011; Wang and Shafieezadeh 2019, 2018). Note that several learning functions (e.g. Expected Feasibility Function (EFF) and U function) have been proposed to select the next best training points. The point that maximizes the EFF response is chosen as the next best training point to refine the Kriging model. On the other hand, U learning function measures the uncertainty of wrong sign (+/-) estimation of $\hat{g}(\boldsymbol{x})$. The point that minimizes the response of U is selected in the learning iteration.

3. MULTIPLE TRAINING POINTS ENRICHMENT STRATEGY

It is shown that both the EFF and U learning functions are efficient in strategically picking

design points (Echard, Gayton, and Lemaire 2011). However, they can select only one training point upon each learning iteration, which is not flexible and impractical for implementation. To select multiple training points at one time and simultaneously keep the total number of evaluation of the performance function minimum, parallel learning strategies need to be explored. In this section, a new learning strategy for selecting multiple training points is proposed to overcome aforementioned drawbacks. Before proposing this strategy, the uncertainty of the estimated failure probability is represented (Wang and Shafieezadeh, 2019), which facilitates the derivation of the optimal learning strategy.

Denote that the failure probability estimated through Probabilistic Classification-based Monte Carlo Simulation (PC-MCS) can be estimated as:

$$\hat{P}_f^{pc} = \frac{\hat{N}_f^{pc}}{N_{MCS}}, \tag{10}$$

where $\hat{N}_f^{pc}$ is the expected number of failure points in S. In this approach, for each candidate design sample, $\boldsymbol{x}_i$, the outcome of the indicator function follows a Bernoulli distribution:

$$I_{\hat{g}}(\boldsymbol{x}_i) \sim B\left(\mu_b(\boldsymbol{x}_i), \sigma_b^2(\boldsymbol{x}_i)\right), \boldsymbol{x}_i \in S, \tag{11}$$

where B denotes the Bernoulli distribution, $\mu_b(\boldsymbol{x}_i)$ is the Bernoulli mean with $\mu_b(\boldsymbol{x}_i) = \Phi\left(\frac{-\mu_{\hat{g}}(\boldsymbol{x}_i)}{\sigma_{\hat{g}}(\boldsymbol{x}_i)}\right)$ and $\sigma_b^2(\boldsymbol{x}_i)$ is the variance of the Bernoulli distribution, $\sigma_b^2(\boldsymbol{x}_i) = \mu_b(\boldsymbol{x}_i)\left(1 - \mu_b(\boldsymbol{x}_i)\right)$. As $\hat{N}_f^{pc}$ can be derived as the expected value of the sum of $I_{\hat{g}}(\boldsymbol{x}_i)$, $\boldsymbol{x}_i \in S$, it follows that $\hat{N}_f^{pc}$ has Poisson binomial distribution (PBD):

$$\hat{N}_f^{pc} \sim PB\left(\mu_{\hat{N}_f^{pc}}, \sigma_{\hat{N}_f^{pc}}^2\right), \boldsymbol{x}_i \in S, \tag{12}$$

where $\mu_{\hat{N}_f^{pc}}$ and $\sigma_{\hat{N}_f^{pc}}^2$ are the mean value and variance of $\hat{N}_f^{pc}$. Therefore, the Confidence Interval of $\hat{N}_f^{pc}$, with confidence level α, can be calculated as:

$$\hat{N}_f^{pc} \in \left(\boldsymbol{\Theta}_{\hat{N}_f^{pc}}^{-1}\left(\frac{\alpha}{2}\right), \boldsymbol{\Theta}_{\hat{N}_f^{pc}}^{-1}\left(1 - \frac{\alpha}{2}\right)\right), \tag{13}$$

where $\boldsymbol{\Theta}_{\hat{N}_f^{pc}}^{-1}(\cdot)$ is the inverse CDF of PBD with mean $\mu_{\hat{N}_f^{pc}}$ and variance $\sigma_{\hat{N}_f^{pc}}^2$. According to(Wang and Shafieezadeh, 2019), the distribution of $\hat{N}_f^{pc}$ in Eq.(12) can be represented as:

$$\hat{N}_f^{pc} \sim N\left(\mu_{\hat{N}_f^{pc}}, \sigma_{\hat{N}_f^{pc}}^2\right),$$

$$\mu_{\hat{N}_f^{pc}} = \sum_{i=1}^{N_{MCS}} \Phi\left(\frac{-\mu_{\hat{g}}(\boldsymbol{x}_i)}{\sigma_{\hat{g}}(\boldsymbol{x}_i)}\right),$$

$$\sigma_{\hat{N}_f^{pc}}^2 = \sum_{i=1}^{N_{MCS}} \sum_{j=1}^{N_{MCS}} \boldsymbol{\Sigma}_{pb_{i,j}}, \boldsymbol{x}_i \in S \tag{14}$$

where $N(\cdot)$ denotes the normal distribution with corresponding CDF, $\Phi(\cdot)$, and $\boldsymbol{\Sigma}_{pb_{i,j}}$ is the element of the covariance matrix, $\boldsymbol{\Sigma}_{pb}$, which can be represented as (Wang and Shafieezadeh, 2019),

$$\boldsymbol{\Sigma}_{pb} = \begin{bmatrix} {\sigma_b}^2(\boldsymbol{x}_1) & \cdots & \rho_{1,N_{MCS}}\sigma_b(\boldsymbol{x}_1)\sigma_b(\boldsymbol{x}_{N_{MCS}}) \\ \vdots & \ddots & \vdots \\ \rho_{N_{MCS},1}\sigma_b(\boldsymbol{x}_{N_{MCS}})\sigma_b(\boldsymbol{x}_1) & \cdots & {\sigma_b}^2(\boldsymbol{x}_{N_{MCS}}) \end{bmatrix}, \tag{15}$$

According to the normal distribution, the Confidence Interval (CI) of $\hat{N}_f^{pc}$ can be calculated as:

$$\hat{N}_f^{pc} \in \left[\mu_{\hat{N}_f^{pc}} - \gamma_{ci}\sigma_{\hat{N}_f^{pc}},\ \mu_{\hat{N}_f^{pc}} + \gamma_{ci}\sigma_{\hat{N}_f^{pc}}\right], \tag{16}$$

where $\gamma_{ci} = 1.96$ in this paper, which corresponds to $\alpha = 0.05$. According to Eq. (10) and (16), the CI of $\hat{P}_f^{pc}$ can be obtained by:

$$\hat{P}_f^{pc} \sim \left[\frac{\mu_{\hat{N}_f^{pc}} - \gamma_{ci}\sigma_{\hat{N}_f^{pc}}}{N_{MCS}}, \frac{\mu_{\hat{N}_f^{pc}} + \gamma_{ci}\sigma_{\hat{N}_f^{pc}}}{N_{MCS}}\right], \quad \alpha = 0.05,\ \gamma_{ci} = 1.96. \tag{17}$$

Note that the uncertainty of the estimated failure probability $\hat{P}_f^{pc}$ can be reflected in the variance $\sigma^2_{\hat{N}_f^{pc}}$ in Eq. (14), which is the summation of each elements of the covariance matrix in Eq. (15). The standard deviations $\sigma_b(\boldsymbol{x}_i)$and $\sigma_b(\boldsymbol{x}_j)$ in Eq. (15) will be reduced when new training points are added, which indicates that the variance $\sigma^2_{\hat{N}_f^{pc}}$ will also decrease. Hence, to maximally reduce the uncertainty of the estimated failure probability $\hat{P}_f^{pc}$, the optimal learning strategy can be defined as:

$$\boldsymbol{x}_{tr}^* = \underset{[\boldsymbol{x}^{\{1\}},\boldsymbol{x}^{\{2\}},\dots\boldsymbol{x}^{\{pa\}}]\in S}{\arg\max} \left\{\sigma_{\hat{N}_f^{pc}}^{current} - \sigma_{\hat{N}_f^{pc}}^{next}\right\} \tag{18}$$

where $\boldsymbol{x}_{tr}^* = \left[\boldsymbol{x}_{tr}^{\{1\}}, \boldsymbol{x}_{tr}^{\{2\}}, \dots \boldsymbol{x}_{tr}^{\{pa\}}\right]$ are the selected best training samples, pa denotes the number of $\boldsymbol{x}_{tr}^*$, $\sigma_{\hat{N}_f^{pc}}^{current}$ and $\sigma_{\hat{N}_f^{pc}}^{next}$ denote the standard deviation, $\sigma_{\hat{N}_f^{pc}}$, at the current and next learning iteration, respectively. Method for the proposed parallel learning strategy is summarized in Algorithm 1.

4. NUMERICAL STUDIES

In this section, the performance of the proposed parallel learning strategy along with EFF and U learning function are investigated. The stopping criterion for learning is set as:

$$\hat{P}_f^{pc} \cong P_f^{MCS} \ when\ \frac{\sigma_{\hat{N}_f^{pc}}}{\mu_{\hat{N}_f^{pc}}} \leq 10^{-3} \tag{19}$$

where $\sigma_{\hat{N}_f^{pc}}$ and $\mu_{\hat{N}_f^{pc}}$ are two parameters in Eq. (14). A non-linear four-branch series system problem has been investigated in many studies (Echard, Gayton, and Lemaire 2011; Wang and Shafieezadeh 2019). Random variables x_1 and x_2 for this problem all follow mutually independent standard normal distributions (e.g. mean of 0 and standard deviation of 1). The performance function, $g(\boldsymbol{x})$ is defined as:

$$g(x_1, x_2) = \min \begin{cases} 3 + 0.1(x_1 - x_2)^2 - \frac{(x_1 + x_2)}{\sqrt{2}} \\ 3 + 0.1(x_1 - x_2)^2 + \frac{(x_1 + x_2)}{\sqrt{2}} \\ (x_1 - x_2) + \frac{6}{\sqrt{2}} \\ -(x_1 - x_2) + \frac{6}{\sqrt{2}} \end{cases} \tag{20}$$

As suggested in (Echard, Gayton, and Lemaire 2011), $\frac{(n_d+1)(n_d+2)}{2}$ initial training points are sufficient, where n_d is the number of dimensions (e.g. $n_d = 2$). The performance of these strategies is compared in terms of the number of calls to the performance function, N_{call}, the number of iterations, N_{ite}, estimated probability of failure, $\hat{P}_f$ and the true error, ϵ. Table 1 shows reliability analysis results with different learning strategies (e.g. single training point selection via EFF and U and multiple training point selection with different numbers of parallel trainings). For this example, the number of candidate design samples is set to 1×10^5 so that the coefficient of variation of failure probability, COV_{P_f}, is smaller than 0.05. In this comparison, the same set of candidate design samples and initial training points are used for all methods in order to remain consistent in failure probability estimation.

As shown in Table 1, the number of iterations with parallel learning strategy is significantly reduced compared to the approach via single training point-based enrichment. For details, N_{ite} for parallel learning strategies with $pa = 3$, $pa = 5$, $pa = 8$ and $pa = 10$ training points-based enrichment only need 24, 15, 11 and 13 iterations. With 6 initial training samples, N_{call} through the proposed method ($pa = 3$, $pa = 5$, $pa = 8$ and $pa = 10$) are 78, 81, 94 and 136, respectively. Furthermore, N_{call} through single training point-based enrichment with EFF and U learning function are 76 and 65. N_{call} through the proposed parallel learning strategy with $pa = 3$ and $pa = 5$ training point-based enrichment (e.g.

$N_{call} = 78$ and $N_{call} = 81$) are very close to the approach estimated through single training point-based learning strategy with EFF learning function. However, N_{ite} is only 24 and 15 for parallel learning strategies with 3 and 5 training points enrichment. The proposed parallel learning strategy estimates failure probability through adaptive Kriging more efficiently. For examples, if 8 computational units are available, the total number of iterations are only 11. Importantly, more computational units cannot guarantee fewer iterations. For example, N_{ite} is 13 for parallel learning strategy with 10 training points enrichment, which, however, is even more than the approach with 8 training points. It can be inferred that large number of training points for each iteration is unnecessary since extra training points make insignificant contributions to the construction of the Kriging model. Hence, the appropriate number of multiple training point enrichment for parallel learning strategies is about 5 for this example. The process of parallel learning strategy with 5 training point enrichment is illustrated in Fig. 1 for different periods of training with $N_{call} = 36$, $N_{call} = 51$, $N_{call} = 66$ and $N_{call} = 81$. Fig. 1 indicates that new training points $\boldsymbol{x}_{tr}^{*}$ (blue cross dots) are all located in the vicinity of the limit state and are distant from each other in different regions.

Algorithm 1. Adaptive Kriging-based Reliability Analysis with Parallel Learning Strategy

1. Generate initial candidate design samples *S* with Latin Hypercube Sampling (LHS)
2. Randomly select initial training samples $\boldsymbol{x}_{tr}$ from *S* and evaluate their responses $g(\boldsymbol{x}_{tr})$
3. Construct the Kriging model $\hat{g}(\boldsymbol{x})$ based on $\boldsymbol{x}_{tr}$ and $g(\boldsymbol{x}_{tr})$
4. Estimate the mean $\sigma_{\hat{g}}(\boldsymbol{x})$, standard deviation $\sigma_{\hat{g}}(\boldsymbol{x})$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_{pb}$ $\hat{P}_f^{MCS}$ for *S* with $\hat{g}(\boldsymbol{x})$
5. Use the parallel learning strategy to search for pa multiple training points:
   (a). Search for the training points $\left[\boldsymbol{x}_{tr}^{\{2\}}, \ldots \boldsymbol{x}_{tr}^{\{pa\}}\right]$ according to Eq. (18)
6. Check if the stopping criterion (e.g. Eq.(19)) is satisfied or not:
   (a). If satisfied, go to step 7.
   (b). If not satisfied, estimate the response $g(\boldsymbol{x}_{tr}^{*})$ for $\boldsymbol{x}_{tr}^{*}$ and go back to Step 3.
7. Output $\hat{P}_f^{MCS}$

Table 1. Reliability analysis results with different learning strategies for the example.

Learning strategy	N_{call}	N_{ite}	$\hat{P}_f^{pc}$	ϵ
MCS	1×10^5	-	4.4×10^{-3}	-
EFF	$6 + 70$	70	4.4×10^{-3}	0
U	$6 + 59$	59	4.4×10^{-3}	0
Proposed method ($pa = 3$)	$6 + 72$	24	4.4×10^{-3}	0
Proposed method ($pa = 5$)	$6 + 75$	15	4.4×10^{-3}	0
Proposed method ($pa = 8$)	$6 + 88$	11	4.4×10^{-3}	0
Proposed method ($pa = 10$)	$6 + 130$	13	4.4×10^{-3}	0

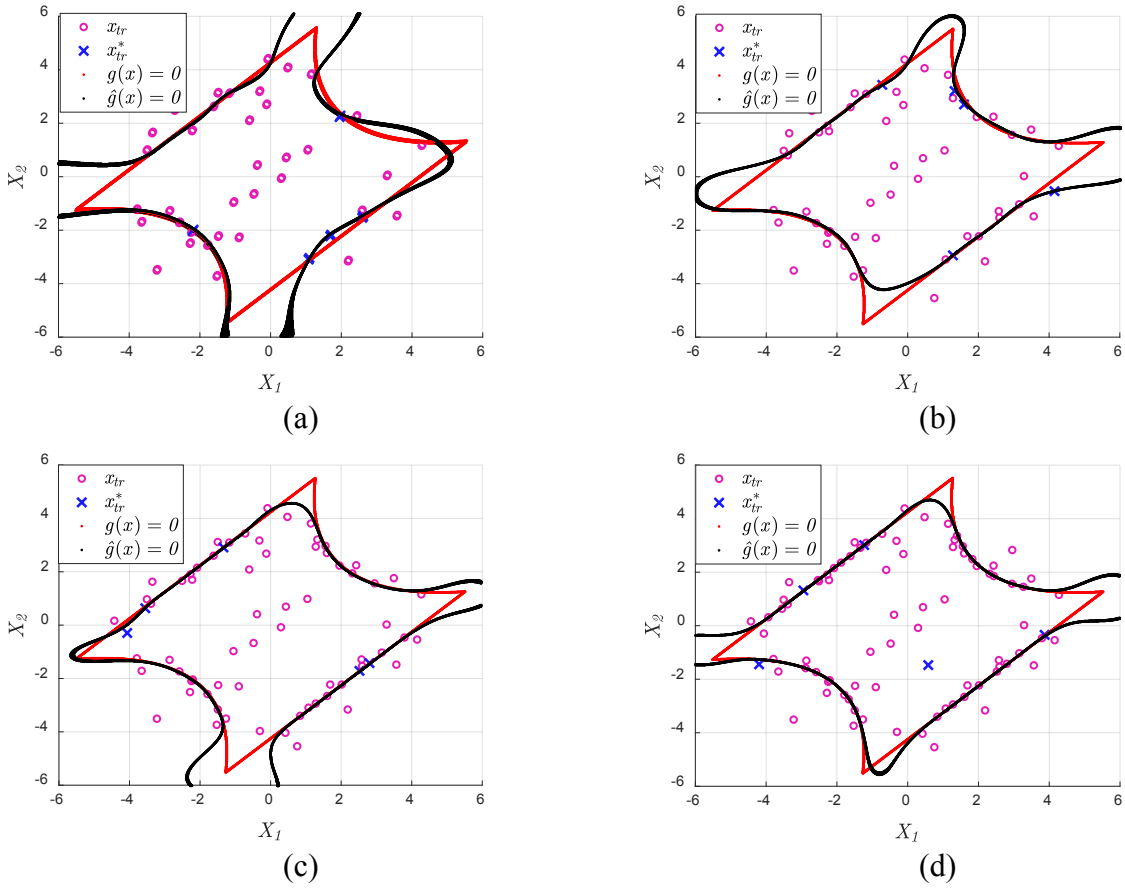

Fig. 1. Illustration of parallel learning strategy ($pa = 5$) for the example in section 4 with (a) $N_{call} = 36$, (b) $N_{call} = 51$, (c) $N_{call} = 66$, and (d) $N_{call} = 81$.

5. CONCLUSIONS

This paper proposes a systematic parallel learning strategy to enable the process of enriching multiple training points in each iteration for adaptive Kriging-based reliability analysis. To explore the optimal learning strategy, uncertainty of the failure probability estimated through adaptive Kriging-based reliability methods is derived by the fact that the total number of failure points estimated through probabilistic classification-based Monte Carlo simulation follows a Poisson Binomial distribution. Afterwards, by realizing that the best learning strategy is to maximally reduce the uncertainty of the estimated failure probability, optimal learning strategy in theory is implicitly represented in this article. To overcome the computational inefficiency in implementing the optimal learning strategy, a pseudo optimal learning strategy is proposed. The efficiencies and in general the advancements offered by the proposed parallel learning strategy are explored by investigating two benchmark reliability problems. Results showcase that the total number of evaluations to the performance function through the proposed parallel learning strategy can be close to the nonparallel approach.

ACKNOWLEDGMENTS

This research has been partly funded by the U.S. National Science Foundation (NSF) through awards CMMI-1462183, 1563372, and 1635569. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Bichon, Barron J., Michael S. Eldred, Laura Painton Swiler, Sandaran Mahadevan, and John M. McFarland. 2008. "Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions." *AIAA Journal* 46 (10): 2459–2468.

Blatman, Géraud, and Bruno Sudret. 2010. "An Adaptive Algorithm to Build up Sparse Polynomial Chaos Expansions for Stochastic Finite Element Analysis." *Probabilistic Engineering Mechanics* 25 (2): 183–197.

Ditlevsen, Ove, and Henrik O. Madsen. 1996. *Structural Reliability Methods*. Vol. 178. Wiley New York.

Echard, B., N. Gayton, and M. Lemaire. 2011. "AK-MCS: An Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation." *Structural Safety* 33 (2): 145–154.

Echard, B., Nicolas Gayton, Maurice Lemaire, and N. Relun. 2013. "A Combined Importance Sampling and Kriging Reliability Method for Small Failure Probabilities with Time-Demanding Numerical Models." *Reliability Engineering & System Safety* 111: 232–240.

Huang, Xiaoxu, Jianqiao Chen, and Hongping Zhu. 2016. "Assessing Small Failure Probabilities by AK–SS: An Active Learning Method Combining Kriging and Subset Simulation." *Structural Safety* 59: 86–95.

Kaymaz, Irfan. 2005. "Application of Kriging Method to Structural Reliability Problems." *Structural Safety* 27 (2): 133–151.

Lelièvre, Nicolas, Pierre Beaurepaire, Cécile Mattrand, and Nicolas Gayton. 2018. "AK-MCSi: A Kriging-Based Method to Deal with Small Failure Probabilities and Time-Consuming Models." *Structural Safety* 73 (July): 1–11.

Lemaire, Maurice. 2013. *Structural Reliability*. John Wiley & Sons.

Lophaven, Søren Nymand, Hans Bruun Nielsen, and Jacob Søndergaard. 2002a. "Aspects of the Matlab Toolbox DACE." Informatics and Mathematical Modelling, Technical University of Denmark, DTU.

"DACE-A Matlab Kriging Toolbox, Version 2.0."

Rubinstein, Reuven Y., and Dirk P. Kroese. 2016. *Simulation and the Monte Carlo Method*. John Wiley & Sons.

Sun, Zhili, Jian Wang, Rui Li, and Cao Tong. 2017. "LIF: A New Kriging Based Learning Function and Its Application to Structural Reliability Analysis." *Reliability Engineering & System Safety* 157: 152–165.

"UQLab Kriging (Gaussian Process Modelling) Manual." n.d. UQLab, the Framework for Uncertainty Quantification. Accessed May 13, 2017.

Wang, Jian, Zhili Sun, Qiang Yang, and Rui Li. 2017. "Two Accuracy Measures of the Kriging Model for Structural Reliability Analysis." *Reliability Engineering & System Safety*, 167 (November): 494–505.

Wang, Zeyu, and Abdollah Shafieezadeh. 2018. "ESC: An Efficient Error-Based Stopping Criterion for Kriging-Based Reliability Analysis Methods." *Structural and Multidisciplinary Optimization*, November.

Wang, Zeyu, and Abdollah Shafieezadeh. 2019. "REAK: Reliability Analysis through Error Rate-Based Adaptive Kriging." *Reliability Engineering & System Safety* 182 (February): 33–45.

Wang, Zeyu, and Abdollah Shafieezadeh. 2019. "Confidence Intervals for Failure Probability Estimates in Adaptive Kriging-Based Reliability Analysis." *Reliability Engineering & System Safety*. (Under review).