

An Incremental Kriging Method for Sequential Optimal Experimental Design

Yaohui Li^{1,2}, Yizhong Wu^{1,3} and Zhengdong Huang¹

Abstract: Kriging model, which provides an exact interpolation and minimizes the error estimates, is a highly-precise global approximation model in contrast with other traditional response surfaces. Therefore, sequential exploratory experimental design (SEED) with Kriging model is crucial for globally approximating a complex black-box function. However, the more sampling points are, the longer time it would take to update the Kriging model during sequential exploratory design. This paper, therefore, proposes a new construction method called incremental Kriging method (IKM) to improve the constructing efficiency with just a little and controllable loss of accuracy for Kriging model. The IKM, based on the matrix segmentation theory, is under the premise that the correlated parameter θ remains unchanged. Meantime, it utilizes the original model and incremental sampling data to quickly obtain an updated Kriging model. Fortunately, a large number of numerical tests showed that the stability of parameter θ would become better and better with the continual increase of the new sampling points in most instances. Even if there is a slight change for value θ , there is not obvious effect on the accuracy of Kriging model. Then, a new sequential incremental experimental design (SIED) algorithm based on IKM is presented to construct Kriging model steadily and effectively. At each sampling step, the SIED method finds an optimal sampling point which maximizes the mean square error of the current model. Meantime, it judges whether the θ should be changed according to the updating criterion. Kriging model will be updated by IKM when θ remains unchanged, or recreates the model with all the sampling points, otherwise. Finally, seven numerical tests and three engineering examples are given to illustrate the applicability, effectiveness and superiority of the proposed methods.

¹ National CAD Supported Software Engineering Centre in Huazhong University of Science and Technology, Wuhan, 430074, Hubei, P.R. China

² School of Electrical Engineering & Mechano-Electronic Engineering in Xuchang University, Xuchang, 461000, Henan, P.R. China.

³ Corresponding author. Email: cad.wyz@mail.hust.edu.cn

Keywords: Global approximation, Black-box function, Kriging model, Sequential experimental design, Incremental Kriging method, Computational cost.

1 Introduction

Global approximation has been widely using in many engineering problems. Firstly, it is very effective and applicable to approximately substitute a complex “black-box” function under some special occasions such as parametric experiments design, real-time simulation and hardware in the loop. On the other hand, global approximation can greatly improve the design efficiency of complex analytical models. Furthermore, some complex engineering problems can be simplified by experiments design [Wang and Shan (2007)], which is helpful for engineers and designers to gradually understand the characteristics of original model by parametric tests, sensibility analysis and visualization of functional relationship. Last but not least, different from the optimization process of some average response surface methods, a global approximation process needs more sampling points so as to achieve global search through the entire design space.

Kriging model, which can provide an exact interpolation and minimize the error estimates (best linear unbiased predictor) in the spatial distribution [Van Beers and Kleijnen (2004)], is one of the widely used global approximation models. Kriging was firstly presented by [Krige (1951)]. And it was further developed by Matheron (1963); Cressie (1992) and Stein (1999). Kriging approximation for deterministic computer models were researched and used by Jones, Schonlau, Welch (1998); Sacks, Welch, Mitchell, Wynn (1989); Santner, Williams, Notz, (2003); Martin, Simpson (2005) and Shahsavani, Grimvall (2009). And real-time simulation models were studied and applied by Panda, Manohar (2008); De Munck, Moens, Desmet, Vandepitte (2009); Han, Görtz, Zimmermann (2012) and Jia, Taflanidis (2013). However, creating Kriging model will be expensive for a large number of experiment data. Hence, fast construction should be researched so as to generate an approximation model at a low cost. Sakata, Ashida, Zako (2004) proposed a fast Kriging algorithm, which efficiently realizes Kriging approximation and optimization with large-scale sampling data. Hartman, Hössjer (2008) presented a Gaussian Markov random field, which can accelerate calculations and decrease memory overhead for Kriging model with large data set.

In addition, a good experimental design method is also essential for fast Kriging modeling. One-off experimental design is relatively straightforward, but it is inappropriate for the construction of Kriging model. It is time-consuming due to large number of simulation evaluations, and it may result in the abnormal use of the model when large-scale data needs to be sampled. Furthermore, it is difficult to determine an appropriate sampling method because information or characteristic of

original model can't be obtained. Therefore, sequential experimental design may be a good choice.

Li, Luo, Rong, Zhang (2013) established an adaptive Kriging model to improve the computational efficiency and numerical accuracy in the approximation of design functions. Zhao, Liu, Li, Yang, Chen (2013) employed moving Kriging interpolation to improve the accuracy of the reliability analysis with a reasonable computational cost. Boukouvala, Ierapetritou (2012) used Kriging model to analyze a black-box process and developed an adaptive sampling strategy. Jin, Chen, Sudjianto, (2005) stated the sequential sampling, which allows engineers to control the sampling process and is more robust and efficient than one-off sampling. Beers, Kleijnen (2005) presented the customized sequential design for random simulation experiments. Lin (2004) proposed a sequential exploratory experiment design (SEED) method, which can generate new sampling data for Kriging model at each step. The simplified and improved flowchart of SEED method is shown in Fig. 1.

For SEED method, Space-filling sampling (such as Latin Hypercube designs (LHD)) is used to generate initial sampling data and corresponding function evaluation. The initial Kriging model will be created by DACE (Design and Analysis of Computer Experiments) method [Søren N Lophaven, Hans Bruun Nielsen, Jacob Søndergaard (2002)]. With the increase of sample obtained by error analysis, Kriging model will be continuously updated till the model validation condition (sample size or model accuracy) is met.

SEED is a stable and exact sequential experimental design. However, too much time may be consumed in the sequential model construction process because Kriging model should be rebuilt by all the sampling points in each update.

This paper, therefore, first puts forward a new method named IKM to fast update a new Kriging model for global approximation. Furthermore, the SIED algorithm based on IKM is proposed to ensure the stability and effectiveness of Kriging model. In the SIED algorithm, the optimal sampling data can be found by the valid information of prior model. In the meantime, whether the parameter θ should be recalculated can be determined in accordance with the update criterion. Once parameter θ remains unchanged, Kriging model will be updated by the IKM, or else, the model should be recreated by DACE with all sampling data. Therefore, the key idea of our proposed methods is that the IKM and the sequential experiment design method are used to improve modeling efficiency and robustness at the cost of a small loss of precision. These methods will be discussed in detail in the following sections.

In this paper, at first, Kriging method and the derivation process of IKM are given in Section 2. Next, the SIED method and some key issues are described in Section

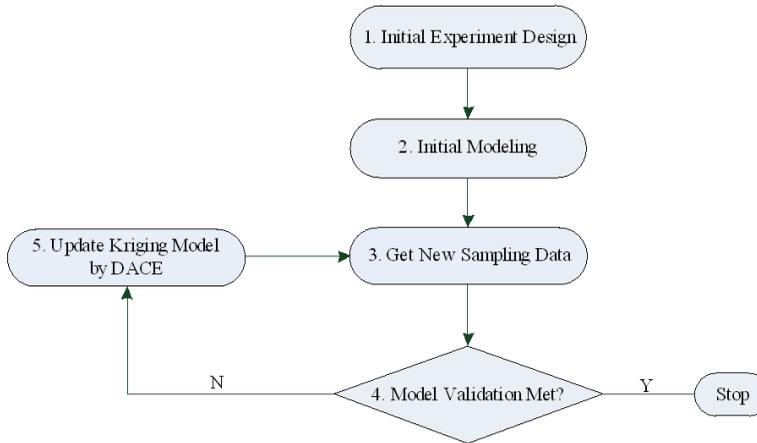


Figure 1: The flowchart of the SEED method.

3. Additionally, seven numerical tests and three engineering applications are used to discuss the feasibility analysis of IKM, modeling efficiency and model accuracy of the SIED method in Section 4 and 5. At last, conclusions are given in Section 6.

2 Incremental Kriging method (IKM)

2.1 Kriging model

Given a set of m design points $X = [x_1, \dots, x_m]^T$ with $\mathbf{X} \in \mathfrak{R}^{m \times n}$ and responses $Y = [y_1, \dots, y_m]^T$ with $Y \in \mathfrak{R}^{m \times 1}$, Kriging model is a combination of a polynomial model and random process [Sasena (2002); Martin (2009)]:

$$Y(\mathbf{x}) = \mathbf{F}\boldsymbol{\beta} + Z(\mathbf{x}), \quad (1)$$

where vector $Y(\mathbf{x})$ is the estimated function of interest. The matrix \mathbf{F} ($\mathbf{F} \in \mathfrak{R}^{m \times p}$) is composed of regression function $f(\mathbf{x})$ at the m known observations. Weighting coefficient $\boldsymbol{\beta}$ of the regression functions is a p -dimension vector. Characteristic of the random process $Z(\mathbf{x})$ can be described by:

$$E[Z(\mathbf{x})] = 0 \quad \text{and} \quad (2)$$

$$E[Z(\mathbf{x})Z(\mathbf{w})] = \sigma^2 R(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}), \quad (3)$$

where σ^2 is a process variance, $\boldsymbol{\theta}$ is a correlation coefficient vector. $R(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x})$ is a spatial correlation function model, which can be obtained by calculating Eq. (4)

$$R(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}) = \prod_{i=1}^n R_i(\theta_i, \omega_i - x_i). \quad (4)$$

The smoothness of Kriging model, the differentiability of the surface and the impact of nearby points can be controlled by $R(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x})$ after determining the correlation of two observations. Søren N Lophaven et al. (2002) offer seven correlation function models. Among them, the Gaussian correlation model has been widely used because it can yield a continuous and smooth curve for two or more design variables.

According to the analysis of Eqs. (1) ~ (4), correlation matrix \mathbf{R} (symmetric and positive definite) and the design matrix \mathbf{F} can be formulated by:

$$\mathbf{R}_{m \times m} = \begin{bmatrix} R(x_1, x_1), \dots, R(x_1, x_m) \\ \vdots \quad \ddots \quad \vdots \\ R(x_m, x_1), \dots, R(x_m, x_m) \end{bmatrix} \text{ and } \mathbf{F}_{m \times p} = \begin{bmatrix} f_1(x_1), \dots, f_p(x_1) \\ \vdots \quad \ddots \quad \vdots \\ f_1(x_m), \dots, f_p(x_m) \end{bmatrix} \quad (5)$$

On the basis of the unbiased estimator theory, the regression problem $\mathbf{F}\boldsymbol{\beta} \approx \mathbf{Y}$ has a generalized least squares solution [Cressie (1992)]

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y} \quad (6)$$

and a variance estimation

$$\sigma^2 = \frac{1}{m} (\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}}). \quad (7)$$

In the light of Eq. (4), the matrix \mathbf{R} and thereby $\hat{\boldsymbol{\beta}}$ and σ^2 are all dependent on parameter $\boldsymbol{\theta}$, which is obtained by an unconstrained optimization [Søren Nymand Lophaven, Hans Bruun Nielsen, Jacob Søndergaard (2002)] problem based on maximum likelihood estimation (MLE) theory, i.e., optimal value $\boldsymbol{\theta}$ can be determined by the maximization of

$$-(m \ln \sigma^2 + \ln |\mathbf{R}|) / 2. \quad (8)$$

As a matter of fact, the best approximate curve may not be generated by the optimal $\boldsymbol{\theta}$, in other words, as long as $\boldsymbol{\theta}$ approaches the minimum, the approximation will be well behaved. This is the reason why Hooke-Jeeves method (a local exploratory algorithm) instead of those global exploratory algorithms is used for the optimal process of $\boldsymbol{\theta}$.

Since the correlation matrix \mathbf{R} is symmetric and positive definite, Cholesky factorization of \mathbf{R} should be introduced by

$$\mathbf{R} = \mathbf{C}\mathbf{C}^T, \quad (9)$$

where \mathbf{C} is a Cholesky factor. Let

$$\tilde{\mathbf{F}} = \mathbf{C}^{-1} \mathbf{F}, \quad \tilde{\mathbf{Y}} = \mathbf{C}^{-1} \mathbf{Y}, \quad (10)$$

then Eq. (6) can be transformed into

$$\hat{\boldsymbol{\beta}} = (\tilde{\mathbf{F}}^T \mathbf{F})^{-1} \tilde{\mathbf{F}}^T \tilde{\mathbf{Y}}. \tag{11}$$

Furthermore, in order to prevent an ill-condition \mathbf{R} , “thin” QR factorization [Golub and Van Loan (2012)] of $\tilde{\mathbf{F}}$ is introduced by

$$\tilde{\mathbf{F}} = \mathbf{Q}\mathbf{G}^T. \tag{12}$$

According to Eqs. (7), (11) and (12), one can finally get

$$\mathbf{G}^T \hat{\boldsymbol{\beta}} = \mathbf{Q}^T \tilde{\mathbf{Y}} \tag{13}$$

and the variance estimator

$$\sigma^2 = \frac{1}{m} \left\| \tilde{\mathbf{Y}} - \tilde{\mathbf{F}} \hat{\boldsymbol{\beta}} \right\|^2. \tag{14}$$

2.2 Incremental updating for Kriging model

Parameter set of Kriging model is mainly composed of parameters $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and σ^2 [Søren Nymand Lophaven et al. (2002)]. In general cases, Kriging model can be constructed as long as value $\boldsymbol{\theta}$ and observations can be obtained by optimization and sampling. However, computational cost of the optimization process should be considered for Kriging modeling. Thus, the IKM is used to efficiently reduce the frequency of using the optimization process by employing constant parameter $\boldsymbol{\theta}$. Advantages of the method will be better showed with the continuous increase of new data points. The specific derivation process of IKM can be stated as follows.

The data sets of new Kriging model are given by

$$\mathbf{X} = [\mathbf{X}_0 \Delta \mathbf{X}]^T, \quad \mathbf{Y} = [Y_0 \Delta Y]^T, \quad \mathbf{F} = [\mathbf{F}_0 \Delta \mathbf{F}]^T, \tag{15}$$

where, \mathbf{X}_0 , Y_0 and \mathbf{F}_0 are the data sets of prior Kriging model, and $\Delta \mathbf{X}$, ΔY and $\Delta \mathbf{F}$ are the new ones obtained by adding new design points. It is necessary to ensure that the increased data sets only have a slight impact on the changing of parameter $\boldsymbol{\theta}$ (it is deemed to be a fixed value, or else, IKM is invalid).

Furthermore, Cholesky factorization of new correlation matrix \mathbf{R} is given by

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_0 & \mathbf{L} \\ \mathbf{L}^T & \Delta \mathbf{R} \end{bmatrix} = \mathbf{C}\mathbf{C}^T = \begin{bmatrix} \mathbf{C}_1 & 0 \\ \mathbf{C}_2 & \mathbf{C}_3 \end{bmatrix} \begin{bmatrix} \mathbf{C}_1^T & \mathbf{C}_2^T \\ 0 & \mathbf{C}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \mathbf{C}_1^T & \mathbf{C}_1 \mathbf{C}_2^T \\ \mathbf{C}_2 \mathbf{C}_1^T & \mathbf{C}_2 \mathbf{C}_2^T + \mathbf{C}_3 \mathbf{C}_3^T \end{bmatrix}, \tag{16}$$

where correlation matrix \mathbf{L} and $\Delta \mathbf{R}$ can be calculated by Eq. (4). $\mathbf{R}_0 = \mathbf{C}_0 \mathbf{C}_0^T$ and Eq. (16) are used to determine \mathbf{C}_1 , \mathbf{C}_2 and \mathbf{C}_3 of \mathbf{R} :

$$\begin{cases} \mathbf{C}_1 = \mathbf{C}_0 \\ \mathbf{C}_2 = \mathbf{L}^T \mathbf{C}_0^{-T}, \\ \mathbf{C}_3 = \Delta \mathbf{C} \end{cases} \tag{17}$$

where $\Delta\mathbf{C}$ is Cholesky factor of $(\Delta\mathbf{R} - \mathbf{C}_2\mathbf{C}_2^T)$, then the inverse of lower triangular matrix \mathbf{C} for Eq. (16) is easily obtained by

$$\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}_0^{-1} & 0 \\ \mathbf{N} & \Delta\mathbf{C}^{-1} \end{bmatrix}, \quad \mathbf{N} = -\Delta\mathbf{C}^{-1}\mathbf{C}_2\mathbf{C}_0^{-1}. \quad (18)$$

In addition, by using Eq. (10), the matrix $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{Y}}$ are expressed as

$$\tilde{\mathbf{Y}} = \mathbf{C}^{-1}\mathbf{Y} = \begin{bmatrix} \mathbf{C}_0^{-1} & 0 \\ \mathbf{N} & \Delta\mathbf{C}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_0 \\ \Delta\mathbf{Y} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{Y}}_0 \\ \Delta\tilde{\mathbf{Y}} \end{bmatrix}, \quad \Delta\tilde{\mathbf{Y}} = \mathbf{N}\mathbf{Y}_0 + \Delta\mathbf{C}^{-1}\Delta\mathbf{Y} \quad (19)$$

and

$$\tilde{\mathbf{F}} = \mathbf{C}^{-1}\mathbf{F} = \begin{bmatrix} \mathbf{C}_0^{-1} & 0 \\ \mathbf{N} & \Delta\mathbf{C}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{F}_0 \\ \Delta\mathbf{F} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{F}}_0 \\ \Delta\tilde{\mathbf{F}} \end{bmatrix}, \quad \Delta\tilde{\mathbf{F}} = \mathbf{N}\mathbf{F}_0 + \Delta\mathbf{C}^{-1}\Delta\mathbf{F}. \quad (20)$$

The “thin” (or “economy size”) QR factorization of $\tilde{\mathbf{F}}$ can be computed by

$$\tilde{\mathbf{F}} = \begin{bmatrix} \tilde{\mathbf{F}}_0 \\ \Delta\tilde{\mathbf{F}} \end{bmatrix} = \mathbf{Q}\mathbf{G}^T = \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix} \mathbf{G}^T. \quad (21)$$

In accordance with Eq. (21), $\tilde{\mathbf{F}}_0$ and $\Delta\tilde{\mathbf{F}}$ are expressed as

$$\tilde{\mathbf{F}}_0 = \mathbf{Q}_1\mathbf{G}^T, \quad \Delta\tilde{\mathbf{F}} = \mathbf{Q}_2\mathbf{G}^T, \quad (22)$$

one can get $\mathbf{G}^T = \mathbf{G}_0^T$ and $\mathbf{Q}_1 = \mathbf{Q}_0$ because of the uniqueness of QR factorization. The corresponding \mathbf{Q} can be described as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_0 \\ \Delta\mathbf{Q} \end{bmatrix}, \quad (23)$$

where $\Delta\mathbf{Q}$ is QR factorization of $\Delta\tilde{\mathbf{F}}$:

$$\Delta\tilde{\mathbf{F}} = \Delta\mathbf{Q}\mathbf{G}_0^T. \quad (24)$$

Thus, the estimator of parameter $\boldsymbol{\beta}$ is equivalent to the solution of

$$\mathbf{G}_0^T\hat{\boldsymbol{\beta}} = \mathbf{Q}^T\tilde{\mathbf{Y}} = [\mathbf{Q}_0^T \ \Delta\mathbf{Q}^T] \begin{bmatrix} \tilde{\mathbf{Y}}_0 \\ \Delta\tilde{\mathbf{Y}} \end{bmatrix} = \mathbf{Q}_0^T\tilde{\mathbf{Y}}_0 + \Delta\mathbf{Q}^T\Delta\tilde{\mathbf{Y}} = \mathbf{G}_0^T\hat{\boldsymbol{\beta}}_0 + \mathbf{G}_0^T\Delta\hat{\boldsymbol{\beta}}. \quad (25)$$

Arrangement for Eq. (25) is shown as

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_0 + \Delta\hat{\boldsymbol{\beta}}. \quad (26)$$

In Eq. (26), $\hat{\boldsymbol{\beta}}$ is generalized least squares estimate of regression coefficient for a new model. Meanwhile, MLE of the variance σ^2 is calculated from Eq. (14) and Eq. (26):

$$\sigma^2 = \frac{1}{m} \left\| \tilde{Y} - \tilde{\mathbf{F}}(\hat{\boldsymbol{\beta}}_0 + \Delta\hat{\boldsymbol{\beta}}) \right\|^2. \tag{27}$$

There is no need for IKM to obtain a new Cholesky factor by decomposing new correlation matrix \mathbf{R} because it can be gained from Eq. (16). When we only introduce a new sampling point each time, Eq. (16) should be simplified into

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_0 & \mathbf{r}(\boldsymbol{\theta}, \omega, \mathbf{x}) \\ \mathbf{r}^T(\boldsymbol{\theta}, \omega, \mathbf{x}) & 1 \end{bmatrix}, \tag{28}$$

where vector $\mathbf{r}(\boldsymbol{\theta}, \omega, \mathbf{x})$ is acquired from Eq. (4). According to Eq. (16) and Eq. (28), \mathbf{C}_1 , \mathbf{C}_2 and \mathbf{C}_3 of \mathbf{R} can be easily expressed as

$$\begin{cases} \mathbf{C}_1 = \mathbf{C}_0 \\ \mathbf{C}_2 = \mathbf{r}^T \mathbf{C}_0^{-T} \\ \mathbf{C}_3 = 1 - \mathbf{r}^T \mathbf{R}_0 \mathbf{r} \end{cases}. \tag{29}$$

Constant parameter $\boldsymbol{\theta}$ is a prerequisite for us to finish the solution of relative parameters. Fortunately, computational time of IKM can be almost ignored when we only add one point to the last sampling data. How to obtain an optimal sampling point and when to use IKM will be discussed in Section 3.

3 Sequential Incremental Experimental Design (SIED) algorithm

3.1 SIED flow

We present a new method named sequential increment experimental design algorithm (SIED), which mainly integrates the IKM, sampling criterion and update criterion. The flowchart of SIED method is shown in Fig. 2. Specific steps are described as follows.

Step 1: Initial experiment design. LHD (a space-filling experiment design) is adopted to obtain initial sampling points, and the corresponding responses are evaluated by some function or real-time simulation. In order to ensure the stability and uniformity of the spatial distribution, we will employ a reasonable number of sampling points to create an initial Kriging model.

Step 2: Initial modeling. Initial Kriging model will be created by the DACE method.

Step 3: Model validation. If the validation condition is met, the modeling process will be terminated, otherwise, the next step will be executed. In general, we hope the model accuracy can be chosen as a validation method. However, in order to reasonably account for the effect and advantage of the presented method, a suitable sample number will be selected as the stopping criterion through synthetically considering model dimensionality, accuracy, running time and computer power.

Step 4: Optimal sampling. One of the aims in this paper is to employ an efficient optimal sampling criterion, in which MSE information from the prior Kriging model is used as a guide to identify a new sampling point. Specific description will be discussed in Section 3.2.1.

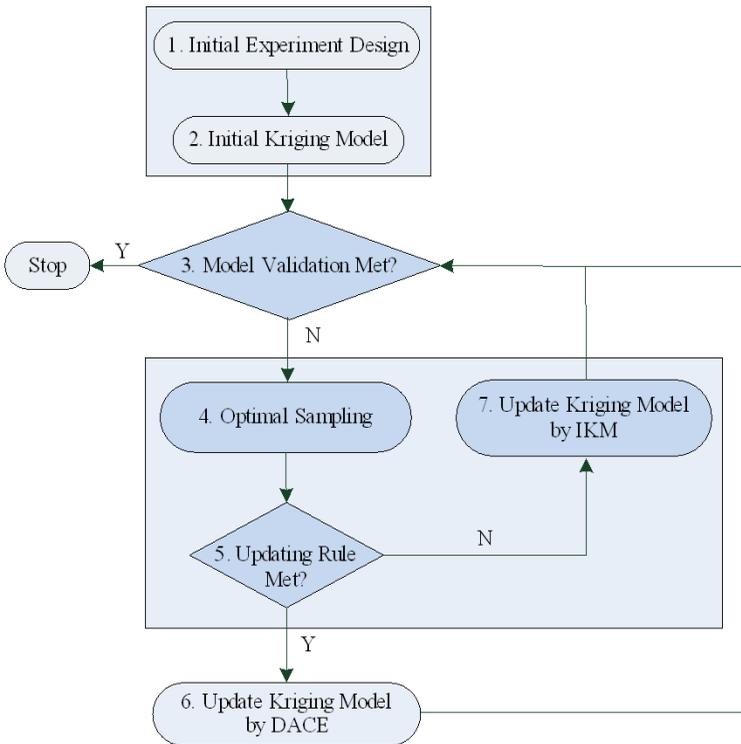


Figure 2: The flowchart of the SIED algorithm.

Step 5: Update criterion. Update criterion can decide which method (the IKM or DACE method) will be used to recreate Kriging model. Specific definition and expression will be shown in Section 3.2.2.

Step 6: DACE modeling. If the update criterion is met, the DACE method will be

used to rebuild Kriging model, and then go back Step 3.

Step 7: IKM Modeling. If the rule can't be satisfied, the IKM will be used to reconstruct Kriging model with the current parameter θ . Specific process has been discussed in Section 2.2. And then go back Step 3.

Next, some key points of the SIED algorithm will be discussed in details.

3.2 Key issues of the SIED algorithm

3.2.1 Optimal sampling criterion based DIRECT

In optimal sampling process, the allocation of new data points is affected by two factors: one is to make the points "spread over" the design space as evenly as possible, and the other is to locate the points in regions with large prediction errors. Therefore, three methods may be adopted to obtain new sampling data for Kriging model in sequential optimal sampling. Maximum entropy principle proposed by Currin, Mitchell, Morris and Ylvisaker (1991) is used to develop computer experiment designs. Similarly, Maximum MSE design and integral MSE (IMSE) design are applied to some deterministic computer experiments by Sacks et al. (1989).

For maximum entropy sampling, the designers tend to add new points which are as far away from current points as possible, but information of the response values is not taken into account in the decision-making process. To this extent, the method is not flexible because it does not affiliate to specific simulations or function. We expect a sampling method may locate one or more new data that are believed to yield maximum potential information from prior observations or Kriging model. Maybe, both maximum MSE and minimum IMSE are good choices. Their optimization procedures are similar except that IMSE increases a weight function and an integral process. Therefore, the maximum MSE design is finally chosen as optimal sampling criterion by overall consideration.

For Kriging model, definition of the estimated MSE [Martin and Simpson (2005)] is written as

$$\varphi(\mathbf{x}) = MSE[Y(\mathbf{x})] = \sigma^2 \left\{ 1 - [\mathbf{f}(\mathbf{x})^T \mathbf{r}(\mathbf{x})^T] \begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix} \right\}. \quad (30)$$

The corresponding sampling criterion can be expressed by

$$\begin{aligned} &\text{find} && \mathbf{x} \\ &\text{maximize} && \varphi(\mathbf{x}) \end{aligned} \quad (31)$$

where, $\varphi(\mathbf{x})$ is the estimated MSE at point \mathbf{x} . More potential and useful information can be gathered from the prior Kriging model by seeking Maximum MSE. Furthermore, correlation function $\mathbf{r}(\mathbf{x})$ of Eq. (30) is relevant with the distance between

the new sampling point and the prior sampling data. Therefore, the new sampling point will be uniformly distributed in the whole design space.

Optimal sampling process is a sequential exploring process until the optimal sampling point is found. To make the MSE have a global and rapid convergence, DIRECT algorithm (an efficient global optimization algorithm) discussed by Jones (2001) will be used to obtain global optimal value. DIRECT method will be terminated once the constraint $|\varphi^{(i)} - \varphi^{(i-1)}| < 0.001$ is met in the process of five consecutive iterations.

There are two merits associated with this sampling strategy:

- (1) Each of new sampling data may be the best choice for model accuracy;
- (2) With the continuous increase of new sampling points, the accuracy will go downhill quickly.

3.2.2 Update criterion

In initial phase of modeling, the addition of a new sampling point with maximum MSE may have a big impact on parameter θ . However, with the continuous increase of the new sampling points, the stability of parameter θ will become better and better. When the number of samples is up to a certain extent, parameter θ tends to keep a relatively steady state. In this case, even if there is a slight change for the value of θ , there is only a small effect on Kriging model accuracy. Therefore, how to introduce a suitable update criterion in the SIED algorithm to determine whether parameter θ need an updating is the problem we are going to solve.

The six sigma principle is considered as an evaluation strategy that employs continuous probabilistic methods and statistical techniques to judge and improve production quality. Jones et al. (1998) used the criteria to implement leave-one-out cross-validation for Kriging. It is believed that the Kriging model was approximately 99.7% confident if the target value lies in the confidence interval using the mean prediction to add and subtract triple standard error. Therefore, the evaluation principle will be adopted as the update criteria of SIED algorithm.

The purpose of using the update criterion is to use the current Kriging model to effectively guide or judge whether the next real response value will lie in the confidence interval (using the mean prediction to add and subtract triple standard error) or not. If the response value lies in the interval, keep θ unchanged, otherwise, update θ .

Given the current Kriging model obtained by k sampled points, a new sampling point (x_{k+1}, y_{k+1}) with maximum MSE $S^2(x_{k+1})$ gained by the optimal sampling criterion, and the current model prediction \hat{y}_k calculated by Eq. (1), then the update

criterion is described as following according to the six sigma evaluation criteria:

$$UR_{k+1} = \left| \frac{y_{k+1} - \hat{y}_k}{S(x_{k+1})} \right| > 3. \quad (32)$$

If the inequality (32) is met, it means that UR_{k+1} does not lie in the interval $[-3, +3]$, i.e., (x_{k+1}, y_{k+1}) is invalid and unsuitable for the current model, then we should reconstruct Kriging model by DACE method. Or else, we only need rebuild this model by IKM.

This update criterion has some randomness. After a great deal of experiments, the probability that the update criterion lying in the interval $[-3, +3]$ is quite low in the initial stage. But there still exists some sampling points which can meet the required criterion. When total sample achieves a certain number, UR_{k+1} probability lied in $[-3, +3]$ is almost one hundred percent, but the sampling data which can't be met still exist in some cases. In summary, the influence of parameter θ is nearly ignored with increasing the new sampling data, which will save more modeling time because of greater use of IKM.

4 Numerical tests

In section 3, The SIED method has been discussed. But its feasibility, time spent and accuracy need be further checked and tested for different dimensions of problems. The validation and comparison process for every test function or engineering application and can be simply stated as follow.

1. Initial sampling points obtained by LHD (Latin Hypercube Design) are used to build initial Kriging model.
2. New sampling point will introduced to sampling data set by optimal sampling criterion and updated rule of SIED or SEED one by one till total sample reaches a certain amount.
3. Visualization of functional relationship between the number of new sampling points and the norm of parameter θ obtained by SEED method is built to illustrate the feasibility of IKM in SIED.
4. Accumulated time spent (unit of time: second) on SIED and SEED can be got from executive program and will be compared to illustrate the advantage of SIED method.
5. Accuracy (RMSE) of the approximate models obtained by SIED method will be compared with those of LHD (only for engineering application) and the SEED method under the same conditions.

All tests are performed in Matlab 2011a by a Dell machine with i3-2120 3.3GHz CPU and 2GB RAM. It is noted that changing trends of each element in θ is almost the same as that of its norm after a great deal of experiments. So we will no longer discuss the trend of each element in θ .

4.1 One-dimensional problems

As a simplest test problem, the function which is multi-peaked, continuous and smooth is chosen as follow:

$$f(x) = \sin(0.02x) \times \cos(0.2(x + 25)) \times \cos(0.1(x + 50)), x \in [0, 100]$$

Four points are sampled by LHD to build initial Kriging model, and new sampling points will be continually added till total sample number reaches 1,000 for SIED and SEED.

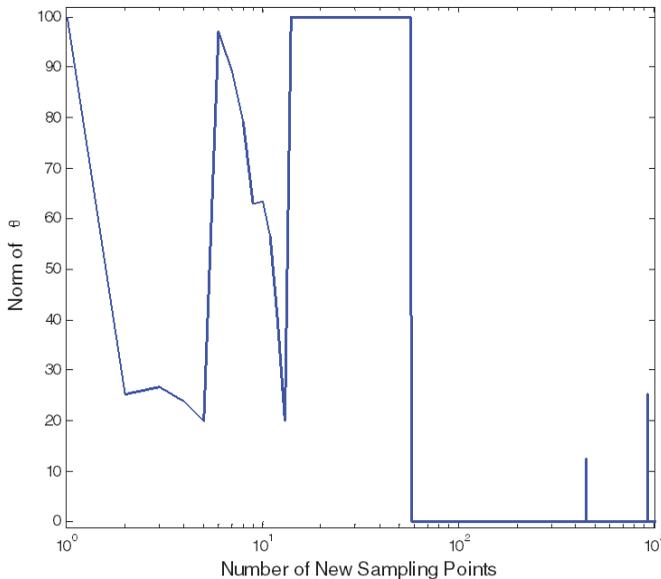


Figure 3: Functional relationship between the number of new sampling points and the norm of θ for the one-dimensional problem in SEED.

Using SEED method, Fig. 3 shows the functional relationship between the number of new sampling points (excluding initial sampling points) and norm of θ . In the beginning, change of the norm is drastic, but it gradually becomes slow and appears many constant values in big sampling intervals. It reasonably describes the feasibility of IKM in SIED.

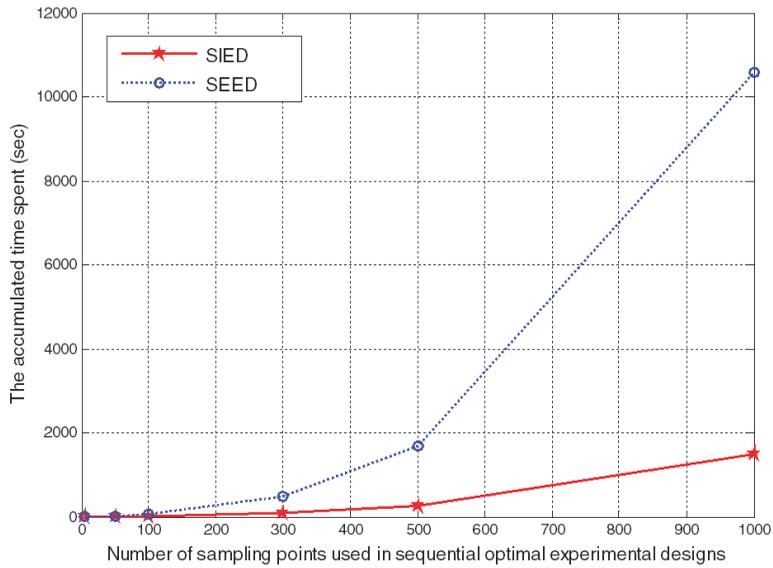


Figure 4: Results on accumulated time spent of the SIED algorithm and the SEED method for the one-dimensional problem.

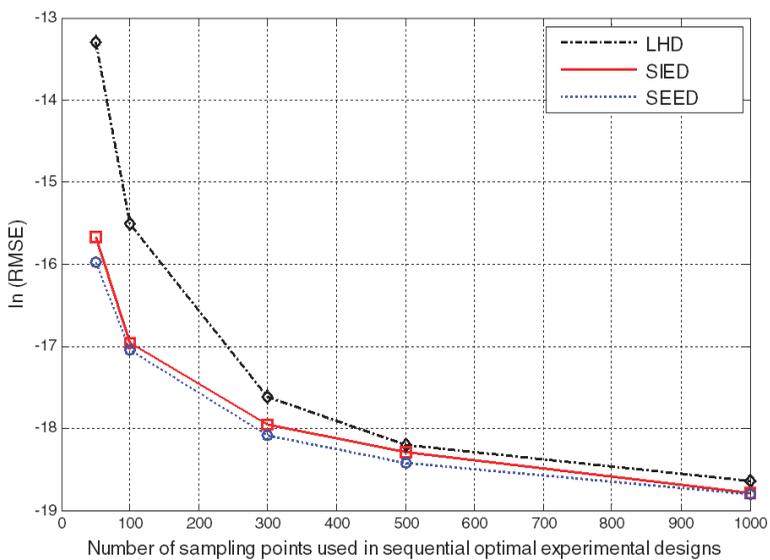


Figure 5: Results on model accuracy of the sequential sampling based on LHD, the SIED algorithm and the SEED method for the one-dimensional problem.

Computational time of the SIED algorithm and the SEED method is shown in Fig. 4. Firstly, initial time spent based on initial LHD is equal for SIED and SEED. Next, accumulated time spent of SIED is compared with that of SEED when accumulated total number of sampling points (including initial sampling point) respectively reaches 50, 100, 300, 500 and 1,000. It is obvious that computing time of SIED is far less than that of SEED, especially for the later stage, time spent of the SIED has been significantly reduced by approximately ninety percent. It has commendably validated the modeling efficiency of SIED.

Results of models accuracy (natural logarithm of RMSE) on the sequential sampling based on LHD, the SEED method and the SIED algorithm are showed in Fig. 5. It is unnecessary to describe initial model accuracy based on initial LHD because they are equal for SIED, SEED and the sequential sampling based on LHD. Therefore, the same accumulated sampling-point numbers (i.e., 50, 100, 300, 500 and 1,000) are chosen to have on accuracy comparison. It is clear that the changing trends of precision are gradually decreasing and slowing. At the 1,000th sampling point, the accuracy (6.8000e-009) of the SEED method is the highest, and that (8.0496e-009) of LHD is the lowest. But the model accuracy (6.8206e-009) of the SIED algorithm is quite close to that of the SEED method, which means that it only has a little accuracy loss for the SIED algorithm. It is worth noting that different modeling method will not produce too much precision variation for one-dimensional problem with a large number of observations.

4.2 Two-dimensional problems

Four test functions: Goldstein and Price function, Schaffer function, Six-hump camel-back function and Himmelblau function are chosen as follows.

(1) Goldstein and Price Function

$$f(x_1, x_2) = (1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)) \times ((30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)), \quad x_{1,2} \in [-2, 2]$$

(2) Schaffer Function

$$f(x_1, x_2) = 0.5 + \frac{\sin^2\left(\sqrt{x_1^2 + x_2^2}\right) - 0.5}{\left(1 + 0.001(x_1^2 + x_2^2)\right)}, \quad x_{1,2} \in [-2, 2]$$

(3) Six-hump Camel-Back Function

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4, \quad x_{1,2} \in [-2, 2]$$

(4)Himmelblau function

$$f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2, \quad x_{1,2} \in [-6, 6]$$

To illustrate these test problems, ten sampling points respectively obtained by LHD will be employed to build initial Kriging model. SIED and SEED will be terminated when the total number of sample arrive at 2,000 points. Hence, 1,990 vector θ will be obtained by the SEED method, the specific changing conditions of their norm for each of the test functions have been shown in Fig. 6 and Fig. 7. The oscillation processes are terminated soon and quickly converge to relatively fixed values. So the use of IKM in SIED will be adaptive to enhance modeling efficiency since there are many constant intervals for θ .

Accumulated time consumptions of the SIED algorithm and the SEED method for the four functions are plotted in Fig. 8. Schaffer function spends the most time (about 5.2453e+4 s), and other three functions (Goldstein and Price function, Six-hump camel-back function and Himmelblau function) cost relatively less time (about 2.5561e+04 s, 2.6046e+04 s and 2.4116e+04 s) in the SEED. But in the SIED, Schaffer function and other three functions only spend about 8034.6e s, 3431.5 s, 3389.5 s and 2691.3, respectively. Percentage rates for time spent are respectively 84.75%, 86.58%, 86.99% and 88.84% descent. And the rates will be further increased with the argument of sampling data. There is no doubt that the SIED algorithm economizes more modeling time in contrast with the SEED method.

Results of model accuracy on the SEED method, the SIED algorithm and LHD based on the sequential sampling are shown in Fig. 9. Model accuracy based on the LHD is visibly less than these of the SEED and the SIED. Average decline rates of accuracy for the SIED are respectively 4.87%, 3.22%, 3.11% and 8.99% in comparison with that of the SEED. It illustrates that loss of accuracy in the SIED algorithm is acceptable.

4.3 Multi-dimension problems

A slice function and a nine-dimensional function are shown as follows.

$$f(x_1, x_2, x_3) = x_1 e^{-(x_1^2 + x_2^2 + x_3^2)}, \quad x_1, x_2, x_3 \in [-2, 2]$$

$$f(\mathbf{x}) = -0.5(x_1 x_4 - x_2 x_3 + x_3 x_9 - x_5 x_9 + x_5 x_8 - x_6 x_7) \\ x_i \in [-10, 10] \quad (i = 1, \dots, 8) \text{ and } x_9 \in [0, 20]$$

Twenty-one and sixty-five sampling points are obtained to build initial Kriging models for slice and the 9-D function. For SIED and SEED, new sampling points

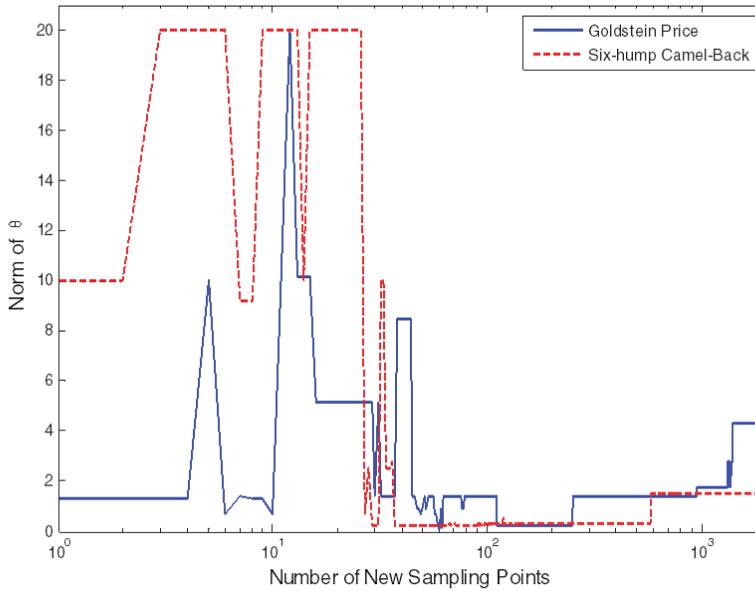


Figure 6: Functional relationship between the number of new sampling points and the norm of θ for the four the Goldstein Price and Six-hump Camel-Back function in SEED.

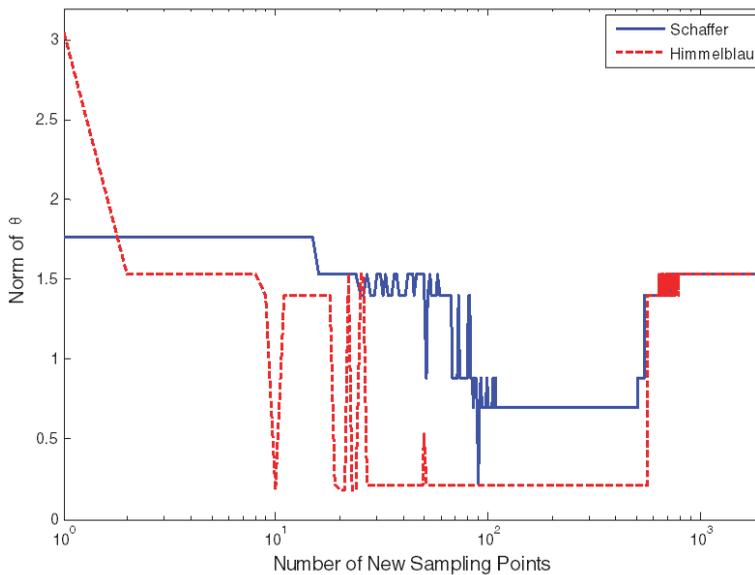


Figure 7: Functional relationship between the number of new sampling points and the norm of θ for Schaffer function and Himmelblau function in SEED.

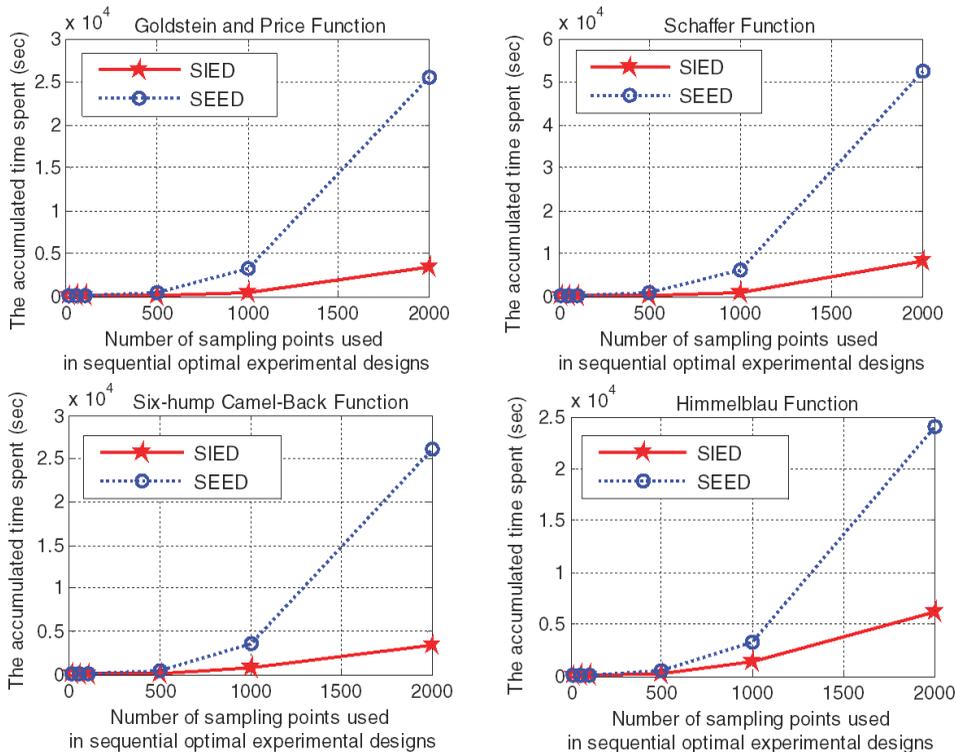


Figure 8: Results on accumulated time spent of the SIED algorithm and the SEED method for the four functions.

will be introduced till total numbers respectively reach 2,000 and 1,000. In this case, variation tendency for norm of θ in SEED are displayed in Fig. 10 and Fig.11. The changing trends in Fig. 10 are from quick to slow, and finally converge to a fixed value. There is only a drastic change for θ between 550 and 610 new sampling points in Fig. 11. So employing the IKM of SIED for multi-dimension problem is more suitable for global approximation.

For the two functions, results on accumulated time spent and model accuracy of the SIED, the SEED and LHD are shown in Fig. 12, Fig. 13, Fig. 14 and Fig. 15. It is clear that computational time and model accuracy are both applicable for SIED, i.e., the accuracy obtained by the SIED algorithm is close to the one obtained by the SEED method, but its computing time is far less than that of SEED. Unfortunately, convergence speeds of the accuracy for the multi-dimension functions, especially for the 9-D function, are comparatively slow.

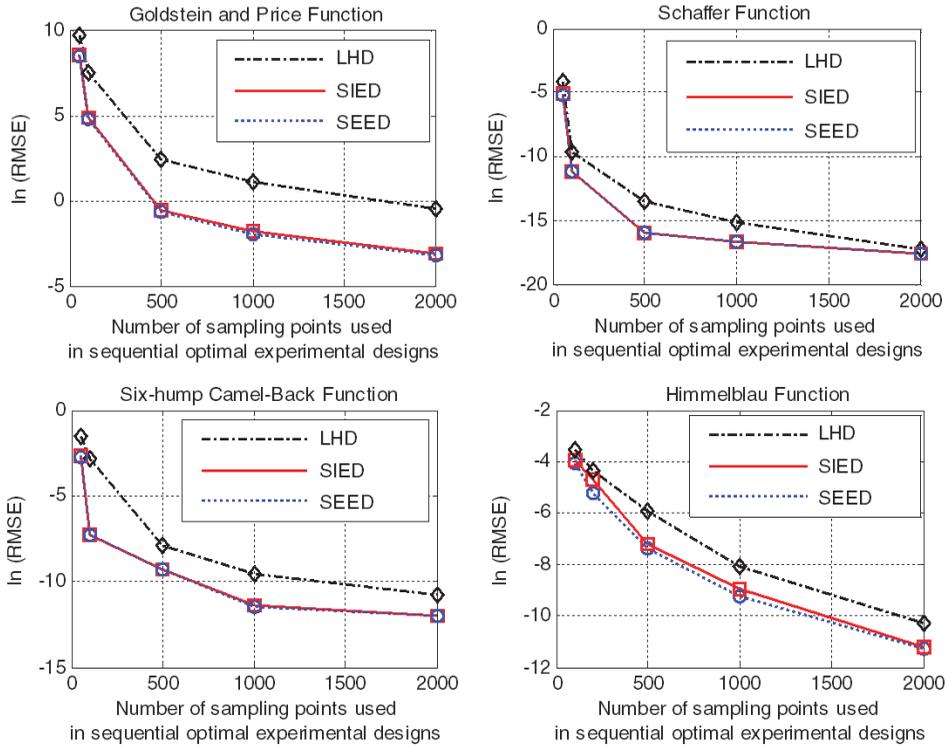


Figure 9: Results on model accuracy of the sequential sampling based on LHD, the SIED algorithm and the SEED method for the four test functions.

5 Engineering applications

5.1 two-bar truss

Model of the symmetric two-bar truss shown in Figure 16 has been approximately built by Balling and Clark (1992). The aim is to minimize the weight of truss system. Here we mainly consider the model approximation problem based some constraints. Given the load $P = 33,000$ lbs, the distance $B = 30$ in between the supports, the cross section wall thickness $t = 0.1$ in, Young’s modulus $E = 3e7$ psi, density $\rho = 0.3$ lbs/in³, and yield stress $\sigma_y = 60,000$ psi. There are two design variables—mean tube diameter x_1 (D) and height x_2 (H) of the truss.

The problem can be stated as

$$f(x_1, x_2) = 0.1885x_1\sqrt{900 + x_2^2},$$

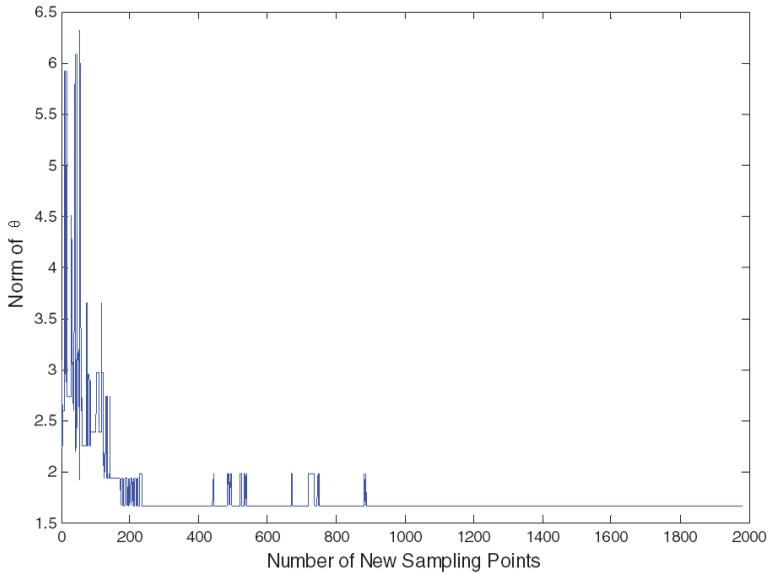


Figure 10: Functional relationship between the number of new sampling points and the norm of θ for slice function in SEED.

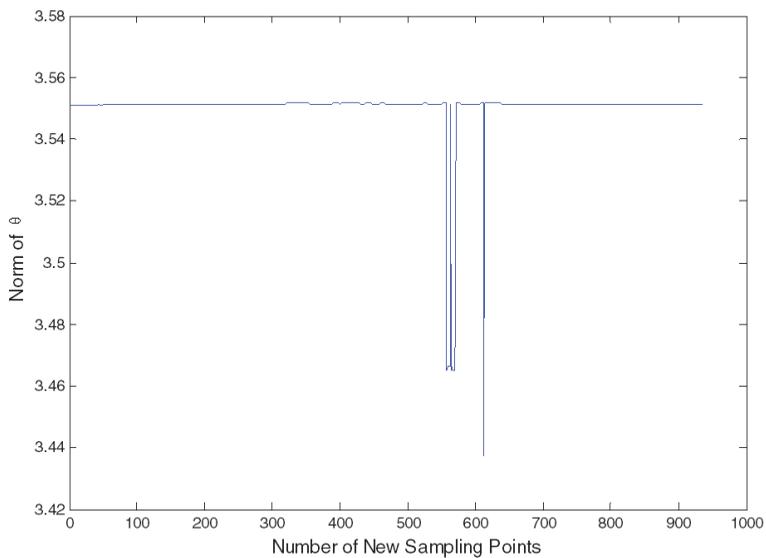


Figure 11: Functional relationship between the number of new sampling points and the norm of θ for the nine-dimensional function in SEED.

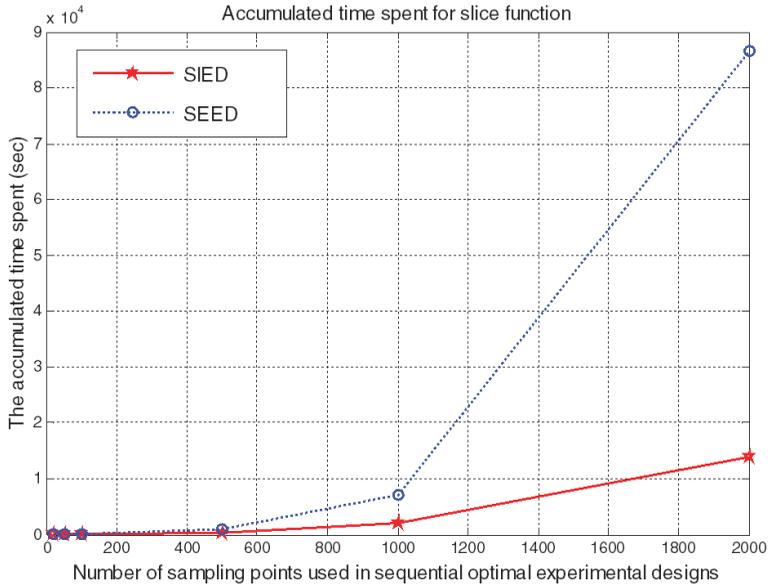


Figure 12: Results on accumulated time spent of the SIED algorithm and the SEED method for slice function.

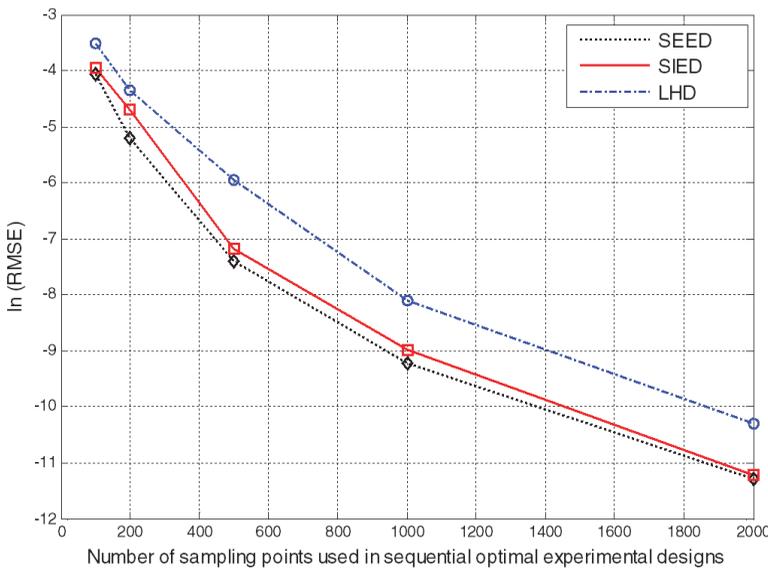


Figure 13: Results on model accuracy of the sequential sampling based on LHD, the SIED algorithm and the SEED method for slice function.

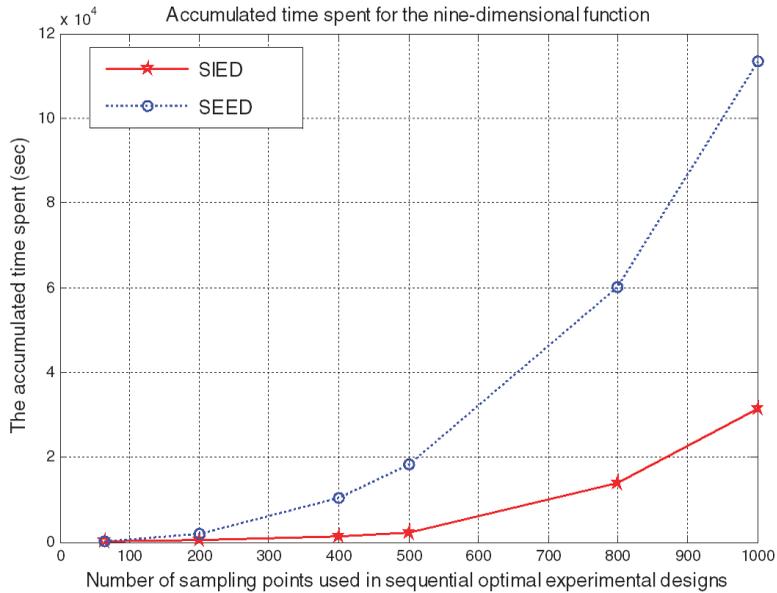


Figure 14: Results on accumulated time spent of the SIED algorithm and the SEED method for the nine-dimensional function.

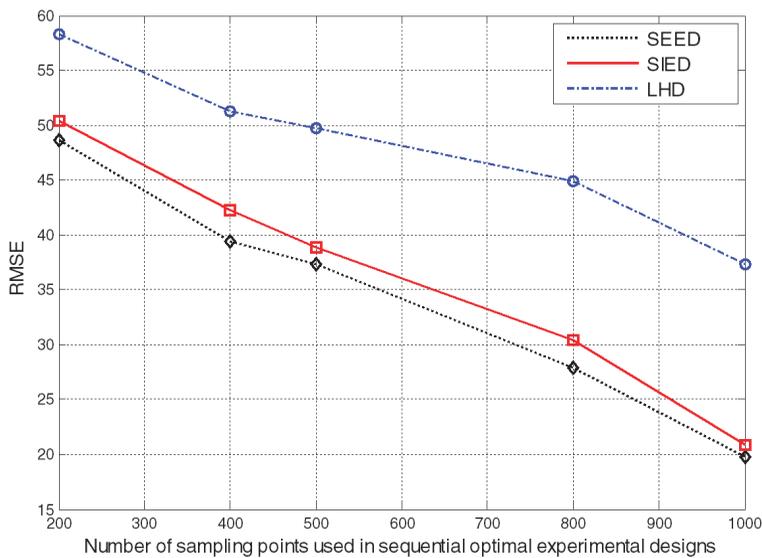


Figure 15: Results on model accuracy of the sequential sampling based on LHD, the SIED algorithm and the SEED method for the nine-dimensional function.

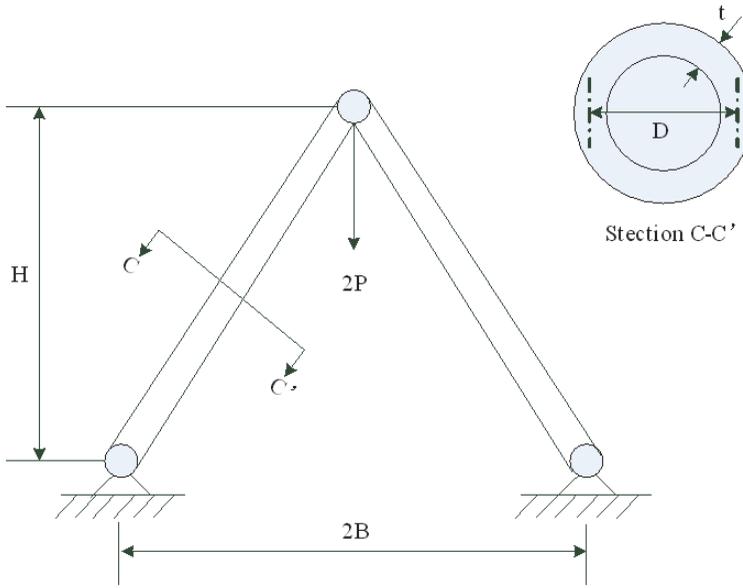


Figure 16: A two-bar truss.

subject to

$$0.5 \text{ in} \leq x_1 \leq 5 \text{ in}$$

$$5 \text{ in} \leq x_2 \leq 50 \text{ in}$$

$$g_1(x_1, x_2) = \frac{2.9609 \times 10^8 (x_1^2 + 0.01)}{7200 + 8x_2^2} - \frac{33000 \sqrt{900 + x_2^2}}{0.3142x_1x_2} \leq 0.$$

$$g_2(x_1, x_2) = 60000 - \frac{33000 \sqrt{900 + x_2^2}}{0.3142x_1x_2} \leq 0$$

Determination of initial sampling data is the same as Section 4.2. Total number of sample is terminated at 2,000 observations for SIED and SEED. Fig. 17 shows parameter θ has a big fluctuation only between 187th and 211th sampling points, and a constant value is always kept in other sampling intervals in SEED. The accumulated computing time and RSME of the models are respectively compared in Fig.18 and Fig. 19. All results prove the SIED algorithm is efficient and robust for the global approximation of two-bar truss.

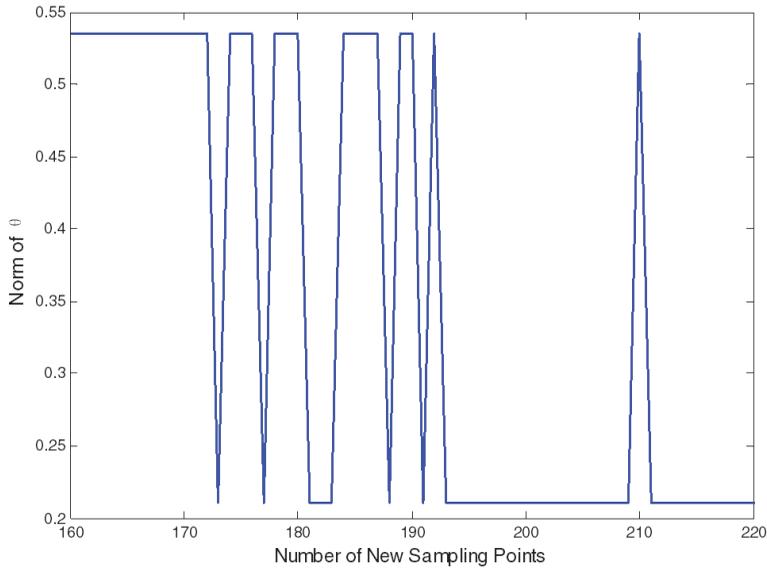


Figure 17: Functional relationship between the number of new sampling points and the norm of θ for the two-bar truss problem in SEED.

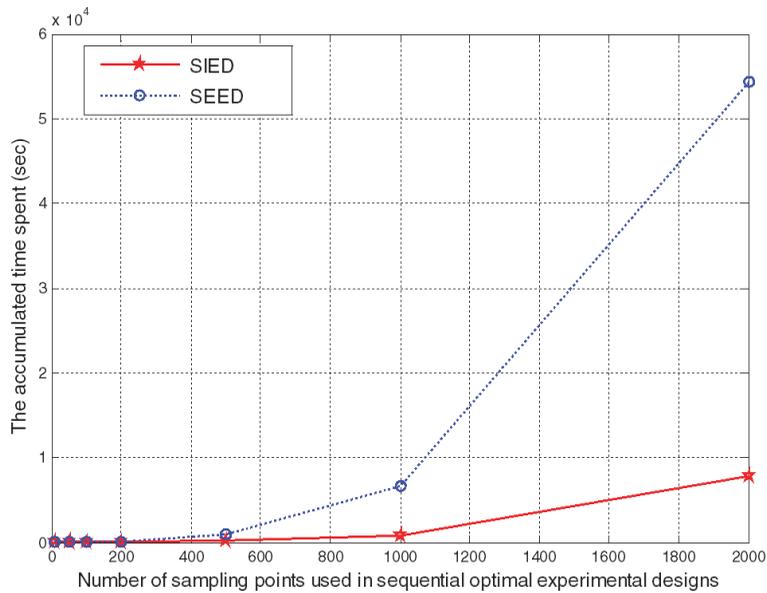


Figure 18: Results on accumulated time spent of the SIED algorithm and the SEED method for the two-bar truss problem.

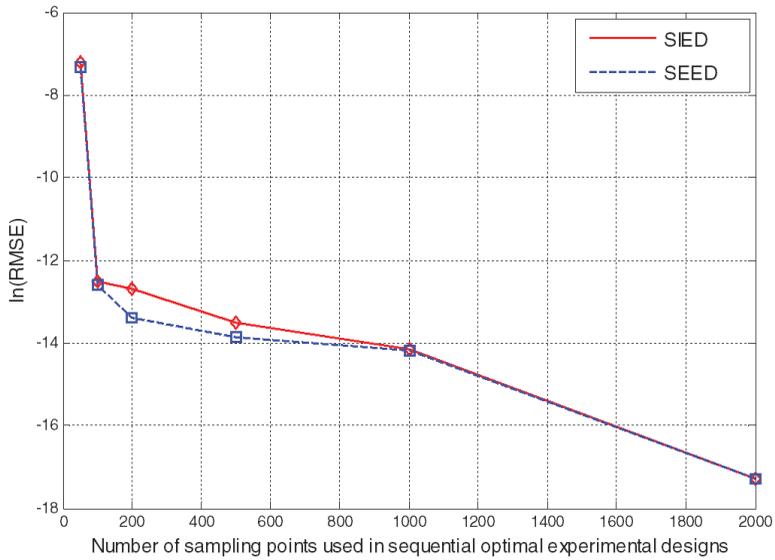


Figure 19: Results on model accuracy of the SIED algorithm and the SEED method for the two-bar truss problem.

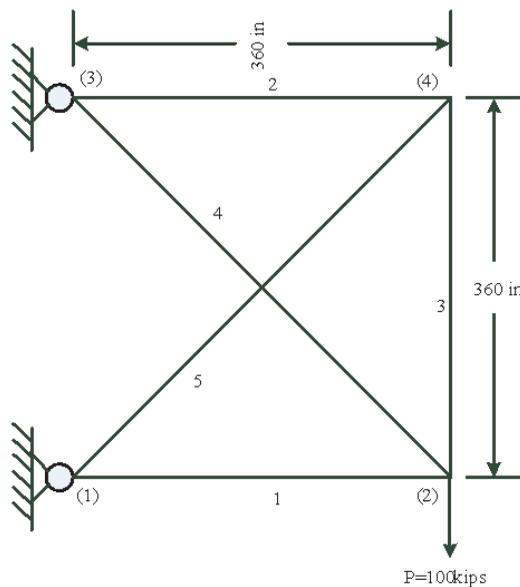


Figure 20: A five-bar planar truss.

5.2 Five-bar planar truss

A five-bar planar truss shown in Figure 20 is a scaled-down version of the ten-bar planar truss problem. It has been studied by Sobieszczanski-Sobieski, Barthelemy and Riley (1982). The aim is to minimize the weight of truss system subject to constraints that the stress in each bar be less than σ_a . A load ($P = 100$ kips) is applied at node (2). Five-bar lengths can be seen or calculated by $L = 360$ in. Set Young's Modulus $E = 1.0E7$ psi, density $\rho = 0.1$ lbs/in³, and allowable stress $\sigma_a = \pm 25,000$ psi. Five design variables (x_1, x_2, x_3, x_4, x_5) are the cross sectional areas of each bar.

The problem is expressed as

$$f(\mathbf{x}) = 36(x_1 + x_2 + x_3 + \sqrt{2}(x_4 + x_5))$$

subject to

$$0.1 \text{ in}^2 \leq x_{1,2,3,4,5} \leq 10 \text{ in}^2 \text{ and } -25000 \text{ psi} \leq \sigma_i \leq 25000 \text{ psi}, i = 1, \dots, 5$$

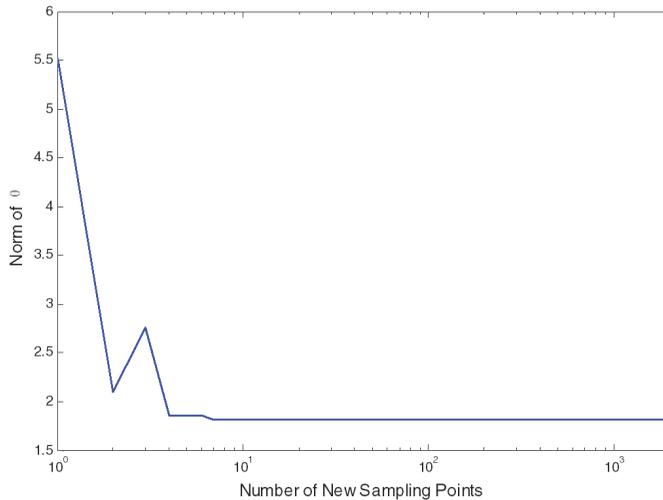


Figure 21: Functional relationship between the number of new sampling points and the norm of θ for the five-bar planar truss problem in SEED.

Thirty-three sampling data is chosen for initial model in SIED and SEED. Sequential construction of model is terminated at 2,000 observations. Fig. 21 shows that a big fluctuation for θ is only appeared in initial 50 new data points in SEED. Results on accumulated time consumption and RMSE are respectively appeared in Fig. 22 and Fig. 23. It turns out that the SIED algorithm is also applicable for five-bar planar truss.

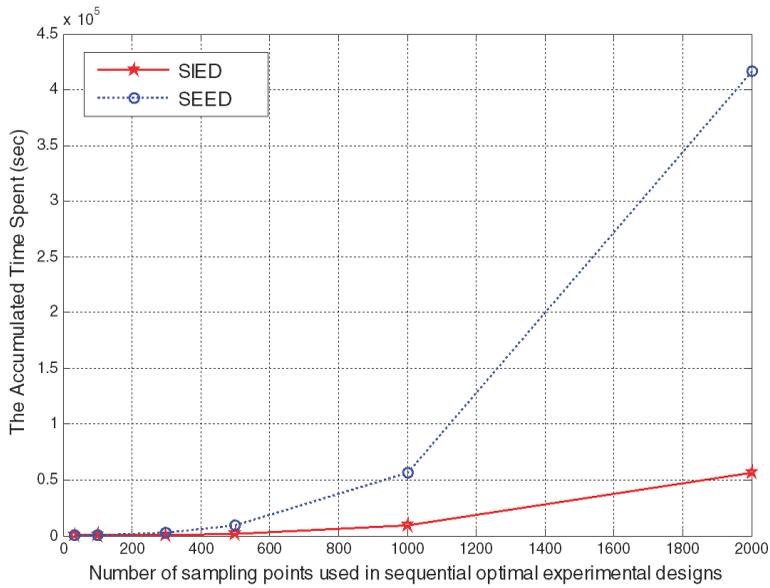


Figure 22: Results on accumulated time spent of the SIED algorithm and the SEED method for the five-bar planar truss problem.

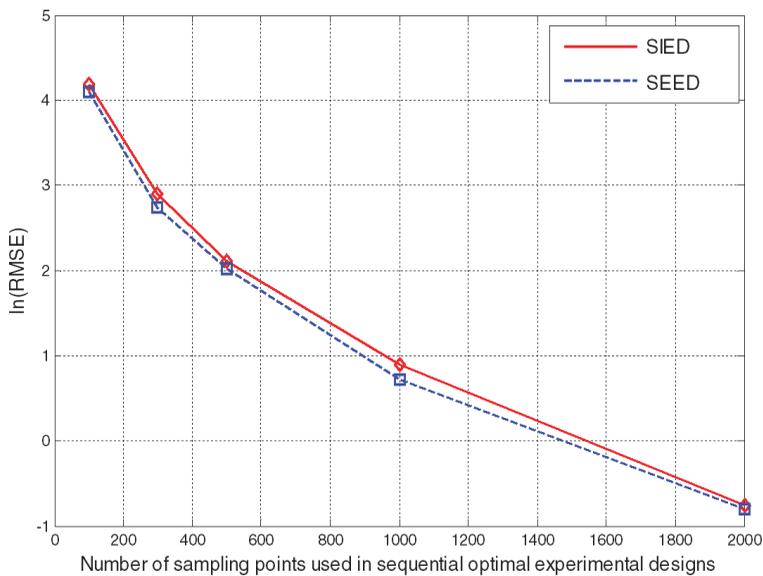


Figure 23: Results on model accuracy of the SIED algorithm and the SEED method for the five-bar planar truss problem.

5.3 A black-box approximate model of cycloid gear pump

Cycloid gear pump [Fabiani et al. (1999)] is composed of the inner rotor and the outer rotor, the front cover, the back cover and the shell. The sketch of cycloid gear pump is shown in Fig. 24. W_1 and W_1' are two symmetrical meshing points between two sides of some tooth of the inner rotor and two teeth of the outer rotor when the inner and outer rotors have the minimum area. W_2 and W_2' are also two meshing ones between the two teeth of the inner rotor and the two sides of the outer rotor when the area is maximized. The four points are the theoretical reference positions, according to the tooth curves of cycloid gear pump and engagement theory, theoretical value [Mao, Li, Hu, Liu and Wang (2005)] of the closed-line angles can be calculated by

$$\alpha_0 = \cot^{-1} \left\{ \cot \frac{\pi}{z_2} + \frac{e}{R \sin(\frac{\pi}{z_2})} \left[1 - \frac{az_1}{\sqrt{R^2 + (ez_2)^2 + 2ez_2R \cos(\frac{\pi}{z_2})} - a} \right] \right\}$$

and

$$\beta_0 = \tan^{-1} \left\{ R \sin \frac{\pi}{z_2} / \left[R \cos \frac{\pi}{z_2} + e \left(\frac{az_1}{\sqrt{R^2 + (ez_2)^2 - 2ez_2R \cos(\frac{\pi}{z_2})} - 1 \right) \right] \right\}$$

where z_1 and z_2 are the number of teeth of inner and outer rotors; e is the center distance of the inner and outer rotors; R is the addendum circle radius of outer rotor; and a is tooth profile circle radius of the outer rotor. Theoretically, $\alpha_1 = \alpha_2 = \alpha_0$, $\beta_1 = \beta_2 = \beta_0$, but the actual situation is not so because the inertia of the fluid must be taken account.

In order to increase actual flow, for α_0 side (having a big oil cavity), we should turn off oil inlet cavity later so that more oil is entered (i.e., $\alpha_0 > \alpha_1$); considering that throttling, we should turn on the oil outlet cavity in advance ($\alpha_0 > \alpha_2$). Likewise, for β_0 side (having a small oil cavity), since the sealing zone is very small, the width of the sealing zone should be increased, therefore, we expect opening size (β_1) of the oil inlet cavity and closing size (β_2) of the oil outlet cavity are both slightly larger than β_0 . The actual flow will be directly affected by changing of the four meshing angles.

According to the above analysis, we should create such a model: $\alpha_1 - \alpha_0$, $\alpha_2 - \alpha_0$, $\beta_1 - \beta_0$ and $\beta_2 - \beta_0$ are defined as four input variables x_1, x_2, x_3, x_4 . Meantime, actual flow is defined as output variable y in the case of a fixed rotation speed (3000r/min). In actual simulation, the four input variables only have small changes. When x_1 ,

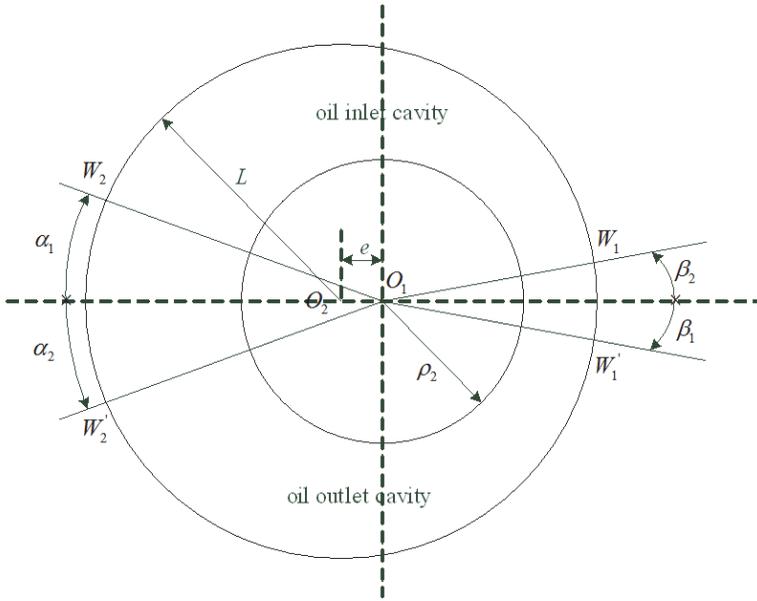


Figure 24: The sketch of cycloid gear pump.

$x_2 > 0^\circ$ and $x_3, x_4 < 0^\circ$, it is unrealistic for a counterclockwise rotation pump to improve volumetric efficiency; when $x_1, x_2 < -1.5^\circ$ and $x_3, x_4 > 1^\circ$, it is also inappropriate to enhance it because this will cause the oil chamber at α_0 side have a poor sealing, which makes oil from the high pressure zone directly leak to a low pressure area, or reduces the total amount of inlet and outlet oil at β_0 side. So domains of the four input variables are given: $x_1, x_2 \in [-1.5^\circ, 0^\circ], x_3, x_4 \in [0^\circ, 1^\circ]$.

For the SIED algorithm, twenty-nine initial input variables with LHD are firstly obtained; and then, according to the inner flow field model (Fig. 25) of a cycloid gear pump, objective evaluations (actual flow) are obtained by using PumpLinx simulation. Initial Kriging model in SIED and SEED will be constructed by DACE method. Finally, model global construction processes are terminated at the 2000th sampling point.

Fig. 26 shows that parameter θ is tending towards stability when the number of new sampling data reaches 875. The results on accumulated time spent and precision of the SIED and the SEED are showed in Fig. 27 and Fig. 28, respectively.

An untried design point $[x_1, x_2, x_3, x_4] = [-1.12^\circ, -0.63^\circ, 0.71^\circ, 0.34^\circ]$ is employed to obtain actual flow $y(y=2.6300 \text{ L})$ and estimate flow $\hat{y}(\hat{y}=2.6287 \text{ L})$. So the RMSE of y and \hat{y} is 0.08268, which basically meet the realistic requirements. At the same time, it illustrates the stability and effectiveness of the SIED algorithm. All results

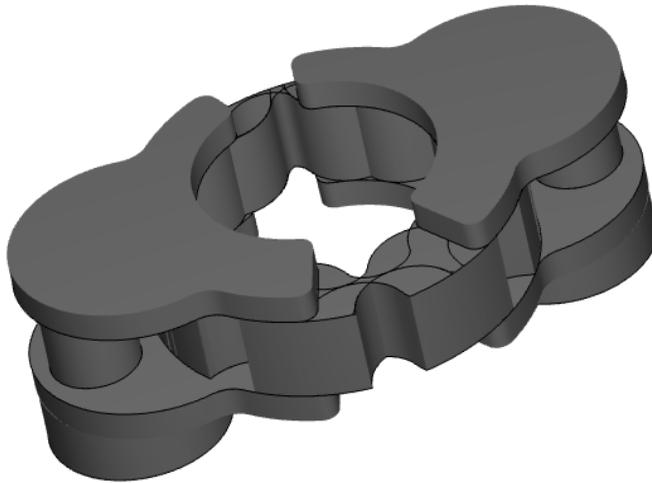


Figure 25: Inner flow field of a cycloid gear pump.

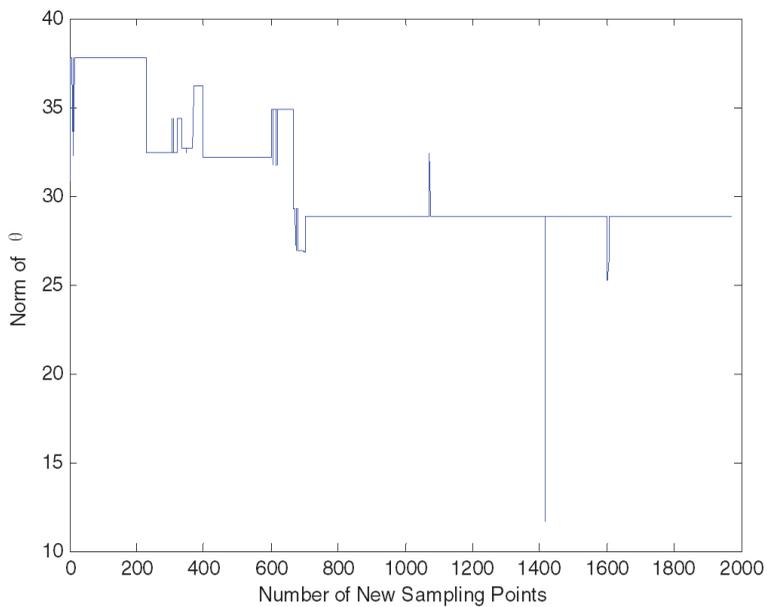


Figure 26: Functional relationship between the number of new sampling points and the norm of θ for inner flow field model of the cycloid gear pump in SEED.

showed it was appropriate to use SIED to improve actual flow of a cycloid gear pump by slightly reducing α_1 , α_2 or increasing β_1 , β_2 in the case of constant revolving speed.

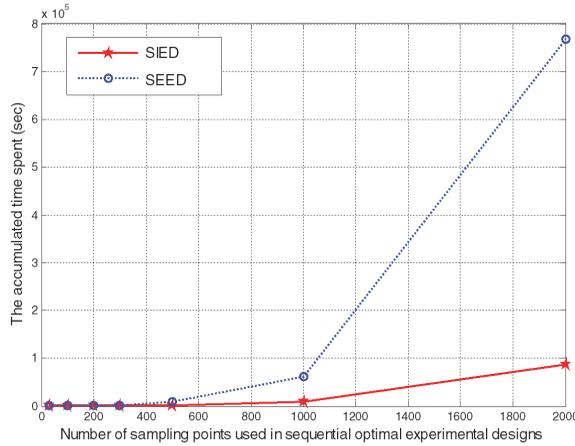


Figure 27: Results on accumulated time spent of the SIED algorithm and the SEED method for the inner flow field model of cycloid gear pump.

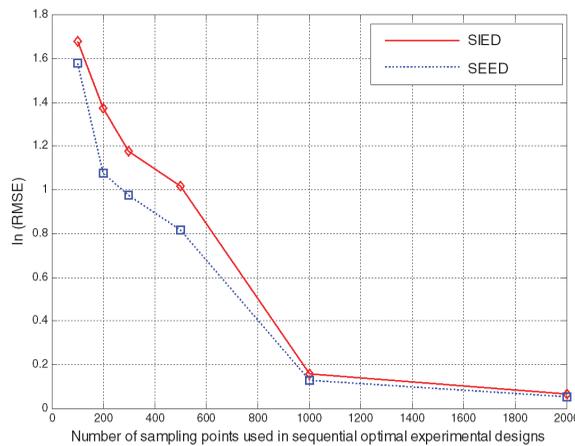


Figure 28: Results on model accuracy of the SIED algorithm and the SEED method for the inner flow field model of the cycloid gear pump.

6 Conclusions

This paper investigates and verifies the SIED algorithm for global Kriging approximation with. It is worth noting that the SIED method based on IKM is employed to ensure efficiency, stability and accuracy of global Kriging model. In addition, the feasibility of the presented methods is validated by seven test functions and three engineering examples. From the results, we can draw the following conclusions.

Table 1: Modeling time of IKM and the DACE method by using different numbers of sampled points for the five-bar planar truss problem

Add a sample	Type Time	IKM	DACE
From 99 to 100	Modeling Time (s)	0.004549	0.1286
From 299 to 300	Modeling Time (s)	0.008761	0.7437
From 499 to 500	Modeling Time (s)	0.02058	2.5281
From 999 to 1000	Modeling Time (s)	0.06038	10.1873
From 1999 to 2000	Modeling Time (s)	0.2046	46.0460

1. Time cost for construction of Kriging model in SIED is mainly affected by Cholesky factorization of \mathbf{R} . However, IKM can efficiently reduce the time consuming because it is unnecessary to solve Cholesky factor again.
2. Compared with the DACE method, modeling time of IKM of SIED is almost ignored without considering the sequent sampling process. Such as the five-bar planar truss problem, modeling time (only adding a new sampling point) of IKM and the DACE method is compared in Table 1. When sample number increases from 99 to 100, it costs 0.0148 s to build Kriging model by DACE, but it takes only 0.0045 s to do it by IKM. The advantage may be not obvious. However, DACE modeling time increase to 46.0406 s, and IKM only increase to 0.2046 s when sample number increases from 1999 to 2000. This is why IKM is used to enhance modeling efficiency in SIED.
3. Maximum MSE sampling criterion is more uniform and effective than LHD in the whole domain. Seen from RMSE, convergence effect for maximum MSE sampling criterion is better than LHD, and model accuracy is also higher than it.
4. Update criterion may be effective to weigh the update efficiency and model precision in the early days of the SIED algorithm. However, parameter θ

usually maintains a constant value when number of sample reaches a certain level, in this case, the effect of update criterion for model precision is greatly diminished, which is why accuracy of the SIED algorithm is almost equivalent to that of the SEED method in later period of sequential sampling.

As an additional discussion, the more effective algorithm based on IKM may be researched by adding multiple sampling points each time in the SIED algorithm. The computational cost may be further reduced, but it is inevitable to face the difficulties and complexity of these problems.

Acknowledgement: This work is supported by the National Natural Science Foundation of China (No. 51175198 and No. 61173115).

References

Balling, R. J.; Clark, D. T. (1992): A flexible approximation model for use with optimization. *AIAA journal*, 2, pp. 886-894.

Beers, W van.; Kleijnen, J. P. C. (2005): Customized Sequential Designs for Random Simulation Experiments: Kriging Metamodelling and Bootstrapping.

Boukouvala, F.; Ierapetritou, M. G. (2012): Feasibility analysis of black-box processes using an adaptive sampling Kriging-based method. *Computers & Chemical Engineering*, vol. 36, pp. 358-368.

Cressie, N. (1992): Statistics for spatial data. *Terra Nova*, vol. 4, no. 5, pp. 613-617.

Currin, C.; Mitchell, T.; Morris, M.; Ylvisaker, D. (1991): Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, vol. 86, no. 416, pp. 953-963.

De Munck, M.; Moens, D.; Desmet, W.; Vandepitte, D. (2009): An efficient response surface based optimisation method for non-deterministic harmonic and transient dynamic analysis. *Computer Modeling in Engineering and Sciences (CMES)*, vol. 47, no. 2, pp. 119.

Fabiani, M.; Mancò, S.; Nervegna, N.; Rundo, M.; Armenio, G.; Pachetti, C.; Trichilo, R. (1999): Modeling and simulation of gerotor gearing in lubricating oil pumps. *SAE transactions*, vol. 108, no. 2, pp. 989-1003.

Golub, G. H.; Van Loan, C. F. (2012): *Matrix computations*, (vol. 3): JHU Press.

Han, Z.; Görtz, S.; Zimmermann, R. (2012): Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aerospace Science and Technology*.

Hartman, L.; Hössjer, O. (2008): Fast kriging of large data sets with Gaussian Markov random fields. *Computational Statistics & Data Analysis*, vol. 52, no. 5, pp. 2331-2349.

Jia, G.; Taflanidis, A. A. (2013): Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. *Computer Methods in Applied Mechanics and Engineering*.

Jin, R.; Chen, W.; Sudjianto, A. (2005): An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, vol. 134, no. 1, pp. 268-287.

Jones, D. R. (2001): Direct global optimization algorithm Direct Global Optimization Algorithm. *Encyclopedia of optimization*, pp. 431-440, Springer.

Jones, D. R.; Schonlau, M.; Welch, W. J. (1998): Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, vol. 13, no. 4, pp. 455-492.

Krige, D. G. (1951): A statistical approach to some basic mine valuation problems on the Witwatersrand. *Jnl. C'hem. Met. and Min. Soc. S. Afr.*

Li, F.; Luo, Z.; Rong, J.; Zhang, N. (2013): Interval multi-objective optimisation of structures using adaptive Kriging approximations. *Computers & Structures*, vol. 119, pp. 68-84.

Lin, Yao. (2004): An efficient robust concept exploration method and sequential exploratory experimental design.

Lophaven, S. N.; Nielsen, H. B.; Søndergaard, J. (2002): A Matlab Kriging Toolbox.

Lophaven, S. N.; Nielsen, H. B.; Søndergaard, J. (2002): Aspects of the matlab toolbox DACE: Informatics and Mathematical Modelling, Technical University of Denmark, DTU.

Mao, H.; Li, G.; Hu, Y.; Liu, H.; Wang, W. (2005): Design and calculation for oil chambers of cycloidal rotor pump. *Journal of Shandong University (Engineering Science)*, vol. 5, pp. 004.

Martin, J. D. (2009): Computational improvements to estimating kriging metamodel parameters. *Journal of Mechanical Design*, vol. 131, no. 8, pp. 084501.

Martin, J. D.; Simpson, T. W. (2005): Use of kriging models to approximate deterministic computer models. *AIAA journal*, vol. 43, no. 4, pp. 853-863.

Matheron, G. (1963): Principles of geostatistics. *Economic geology*, vol. 58, no. 8, pp. 1246-1266.

Panda, S. S.; Manohar, C. S. (2008): Applications of meta-models in finite ele-

ment based reliability analysis of engineering structures. *Comput Model Eng Sci*, vol. 28, pp. 161-184.

Sacks, J.; Welch, W. J.; Mitchell, T. J.; Wynn, H. P. (1989): Design and analysis of computer experiments. *Statistical science*, vol. 4, no. 4, pp. 409-423.

Sakata, S.; Ashida, F.; Zako, M. (2004): An efficient algorithm for Kriging approximation and optimization with large-scale sampling data. *Computer methods in applied mechanics and engineering*, vol. 193, no. 3, pp. 385-404.

Santner, T. J.; Williams, B. J.; Notz, W. (2003): *The design and analysis of computer experiments*: Springer.

Sasena, M. J. (2002): *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*. Citeseer.

Shahsavani, D.; Grimvall, A. (2009): An adaptive design and interpolation technique for extracting highly nonlinear response surfaces from deterministic models. *Reliability Engineering & System Safety*, vol. 94, no. 7, pp. 1173-1182.

Sobieszczanski-Sobieski, J.; Barthelemy, J. F.; Riley, K. M. (1982): Sensitivity of optimum solutions of problem parameters. *AIAA journal*, vol. 20, no. 9, pp. 1291-1299.

Stein, M. L. (1999): *Interpolation of spatial data: some theory for kriging*: Springer.

Van, B.; Wim, C. M.; Kleijnen, J. P. C. (2004): *Kriging interpolation in simulation: a survey*. Paper presented at the Simulation Conference, 2004. Proceedings of the 2004 Winter.

Wang, G. G.; Shan, S. (2007): Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design*, vol. 129, pp. 370.

Zhao, W.; Liu, J. K.; Li, X. Y.; Yang, Q. W.; Chen, Y. Y. (2013): A Moving Kriging Interpolation Response Surface Method for Structural Reliability Analysis. *CMES: Computer Modeling in Engineering & Sciences*, vol. 93, no. 6, pp. 469-488.

