Submitted to Engineering Computations, April 2016

A new sequential sampling method for constructing the high-order polynomial surrogate models

by

Jinglai Wu[1,2], Zhen Luo[1,*], Yunqing Zhang[2], Nong Zhang[1]

[1] School of Electrical, Mechanical and Mechatronic Systems
The University of Technology, Sydney, NSW 2007, Australia

[2] National Engineering Research Center for CAD
Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

Original submission: April, 2016

*Correspondence author in manuscript submission
(Dr. Z. Luo, Email: zhen.luo@uts.edu.au, Tel: +61-2-9514-2994; Fax +61-2-9514-2655)

This paper is submitted for possible publication in ***Engineering Computations***. It has not been previously published, is not currently submitted for review to any other journals, and will not be submitted elsewhere during the peer review.

Abstract

Purpose - The sampling methods (or design of experiments) which have large influence on the performance of the surrogate model will be mainly studied. To improve the adaptability of modelling, a new sequential sampling method termed as sequential Chebyshev sampling method (SCSM) is proposed in this study.

Design/methodology/approach - The high-order polynomials are used to construct the global surrogated model, which retains the advantages of the traditional low-order polynomial models while overcoming their disadvantage in accuracy. Firstly, the zeros of Chebyshev polynomials with the highest allowable order will be used as sampling candidates to improve the stability and accuracy of the high-order polynomial model. In the second step, some initial sampling points will be selected from the candidates by using a coordinate alternation algorithm, which keeps the initial sampling set be uniformly distributed. Thirdly, a fast sequential sampling scheme based on the space-filling principle is developed to collect more samples from the candidates, and the order of polynomial model is also updated in this procedure. The final surrogate model will be determined as the polynomial that has the largest adjusted R-square after the sequential sampling is terminated.

Findings - The SCSM has better performance in efficiency, accuracy, and stability compared with several popular sequential sampling methods, e.g. LOLA-Voronoi algorithm and global Monte Carlo method from the SED toolbox, and the Halton sequence.

Originality/value – The SCSM has good performance in building the high-order surrogate model, including the high stability and accuracy, which may save a large amount of cost in solving complicated engineering design or optimization problems.

Keywords: Global surrogate models; High-order polynomials; Sequential space-filling sampling; Chebyshev polynomials.

1. Introduction

The surrogate model, also termed as meta-model, response surface or emulator, refers to any relatively simple relationship between parameters and response often based on limited data (Rutherford et al., 2005). In order to save the computational cost of some complicated simulations, e.g. the finite element analysis of vehicle crashworthiness usually takes tens of hours to run one simulation, the surrogate models have been widely used in engineering and many other disciplines that have a relationship with statistics and computer simulations. The surrogate model can be divided into two types, which are the local surrogate model and global surrogate model. The local surrogate model is used as an intermediate medium to find the global optimum and is discarded afterwards while the global surrogate model is used as a full replacement of original complex system (Crombecq et al., 2011b). This study is focused on the global surrogate model.

Many kinds of surrogate models have been proposed, including the traditional response surface (low-order polynomials) (Myers and Montgomery, 1995), radial basis function (RBF) (Shan and Wang, 2010), Kriging model (Simpson et al., 2001b), multivariate adaptive regression splines (MARS) (Friedman, 1991), support vector regression (SVR) (Clarke et al., 2005), high dimensional model representation (HDMR) (Hajikolaei and Wang, 2012), and the combination of these models (Goel et al., 2006, Acar, 2010, Marrel et al., 2011). Jin et al. (Jin et al., 2001) show that the polynomial model have advantages in efficiency, transparency, and conceptual simplicity. They also noted that the performance of a surrogate model was influenced by the sampling points. Simpson et al. (Simpson et al., 2001a) demonstrate that the quadratic polynomials are better for low-order nonlinear functions but have a low accuracy for the high-order nonlinear problems. Further studies about the comparison of different surrogate models can be found in references (Barton, 1994, 1998, Wang and Shan, 2007, Gallina, 2009).

Although the low-order polynomials do not have a high accuracy for some strong nonlinear problems, they are still widely used in practical engineering problems due to its advantages in efficiency and transparency. To retain the advantages of low-order polynomials while overcome its weakness, the high-order polynomials (Lin, 2006) will be used to establish the surrogate model. However, the high-order polynomials surrogate model is rarely studied because of the numerical instabilities, e.g. the Runge phenomena. To solve this problem, we can select some particular points as the samples to avoid the numerical instability, e.g. the zeros of Chebyshev polynomials (Fox and Parker, 1968, Gil et al., 2007, Wu et al., 2013a, Wu et al., 2013b, Wu et al., 2014a, Wu et al., 2014b). On the other hand, a large number of coefficients to be estimated in the high-order polynomials make the sampling process expensive, particularly for high-dimensional problems (Barton, 1992). Therefore, how to choose the sampling points has large influence on both the accuracy and efficiency of high-order surrogate model.

The accuracy of a surrogate model will be improved when more data are sampled, but it is impossible to choose too many sampling points due to the computational cost. A good sampling scheme (or DOE) should have a good trade-off between the computational cost and accuracy. In this study, the original model is realized by the deterministic computer simulator, so the output will be the same if the simulator runs twice at the same input values. As a result, the classical DOE (Rutherford et al., 2005), e.g. the Central Composite Design (CCD) (Montgomery, 2007) which is mainly used for laboratory experiments (where the random errors are assumed to exist) will not be considered in this paper. Since no additional information about the inner mechanism of the original model is available, the sampling points are hoped to scatter in the entire design domain uniformly so that the space-filling sampling methods are the most suitable technique to handle this problem.

There are various types of space-filling sampling methods, such as the Factor Design (FD) (Montgomery, 2007), Pseudo-Monte Carlo Sampling (Giunta et al., 2003) (PMCS, e.g. the Latin Hypercube sampling (Mckay et al., 1979, Chen et al., 2012) and Orthogonal Sampling (Owen, 1992)), and Quasi-Monte Carlo Sampling (Fang et al., 2000, Giunta et al., 2003, Rutherford et al., 2005, Pronzato and Müller, 2011) (QMCS, e.g. the Halton sequence (Wong et al., 1997)). Most of the sampling methods mentioned above belong to the "one-shot" sampling schemes, as the samples are chosen once and fixed in the model fitting process (Crombecq et al., 2009). These methods can be easily implemented and provide a good coverage of the design space without incorporating prior knowledge of the system. However, the "one-shot" DOE may suffer from its inflexibility to learn the special characteristics of the shape of the response surface (Ajdari and Mahlooji, 2014), and the number of sampling points is easily over or under estimated. To improve the flexibility and efficiency of sampling, the sequential sampling strategies (also termed as the adaptive sampling (Busby et al., 2007, Gorissen et al., 2010), and incremental sampling (Romero et al., 2004)) have been developed. In the sequential sampling, a surrogate model is firstly built using an initial set of samples and then sequentially updated by adding new samples that are selected by analysing the previous data.

The global exploration and local exploitation are two schemes used in the sequential sampling strategy. The global exploration scatters samples in the regions containing no sampling points while the local exploitation adds more samples to regions identified to be interesting (Crombecq et al., 2011a). Crombecq et al. (Crombecq et al., 2011b) illustrate the trade-off between the exploration and exploitation and indicate that any sequential design strategy will miss important regions in the response surface without proper design space exploration. Therefore, the sequential design must be space-filling to a certain degree firstly, and then the exploitation-based techniques can be implemented in some interesting

regions to improve the sampling quality. Furthermore, the polynomials surrogate model is more suitable for grasping the global characteristics, so only the exploration-based sequential sampling is studied in this work. More studies combining exploration and exploitation to build surrogate model can be referred to (Crombecq et al., 2009, Crombecq et al., 2011a, Ajdari and Mahlooji, 2014).

Granularity (Crombecq et al., 2011b) is an important evaluation index of the sequential sampling. A sequential sampling technique with small granularity adds a small number of points in each sampling while the large granularity makes the sampling increment too large, which reduces the flexibility of sequential sampling. Romero et al. (Romero et al., 2004) used the Progressive Lattice Sampling (PLS) designs to construct the progressive response surface. The PLS allows only a quantized increment M of samples to be added to an existing PLS level (point set) to achieve to a new level. This quantized incremental cost M accelerates quickly with the increase of the PLS level and dimension of the parameter space. To make the sequential sampling more flexible, Romero et al. (Romero et al., 2005) suggested to use the Halton points to build the progressive response surface, because Halton sampling does not suffer from the cost-scaling problems that the PLS does. Halton (Wong et al., 1997) is a lower discrepancy sequence method and has a hierarchical structure. A comparison of different types of sequential sampling strategies was given by (Crombecq et al., 2011b), in which the criterions including the granularity, space-filling, and projective properties were used to evaluate the performance of sequential sampling methods. However, the comparison was only based on the sampling points without considering the final surrogate model. Since the performance of a surrogate model is determined by both the sampling points and fitting model, even the sampling methods having the best performance in the three criterions of (Crombecq et al., 2011b), there is possible that they produce bad performance in the final surrogate model. Therefore, there is necessary to study the sequential sampling techniques combining with fitting model.

This study will focus on the proposal of a new sequential sampling method to build the high-order polynomials surrogate model (HOPSM), which has higher approximation accuracy than traditional low-order polynomial models. To make the HOPSM more stable and accurate, we will propose a new sequential sampling method, in which the sampling points will be selected from a candidate set that is comprised of the zeros of Chebyshev polynomials, so it is termed as sequential Chebyshev sampling method (SCSM). A uniformly distributed initial sampling set will be selected from the candidate set by a coordinate alternation algorithm, and then a fast sequential sampling algorithm will be used to pick more points iteratively based on the space-filling principle (Jonson et al., 1990). At the same time, the order of HOPSM will be updated in the process of sequential sampling. Another three sequential sampling methods, which are the Halton sequence, LOLA-Voronoi algorithm, and global Monte Carlo method (Crombecq et al., 2011b), will be employed as the referees of the proposed SCSM. Several mathematical

test examples and two engineering applications are used to show the efficiency, accuracy and robustness of the proposed method.

2. High-order polynomial surrogate model

2.1 The 1-dimensional HOPSM

In this paper the simulation model is assumed to be continuous, so it can be approximated by using polynomials. To simplify the problem, we consider a continuous 1-dimensional function $f(x)$ on $x\in[-1,1]$ firstly (any interval $[a, b]$ can be normalized to [-1,1] via a linear transformation). $f(x)$ can be approximated by the truncated Chebyshev series $p_n(x)$ as follows:

$$f(x) \approx p_n(x) = \frac{1}{2} f_0 + \sum_{i=1}^{n} f_i C_i(x) \tag{1}$$

where f_i are the Chebyshev coefficients that are constant, and $C_i(x)$ denotes the Chebyshev polynomials

$$C_i(x) = \cos i\theta,\ \ \theta = \arccos\left(x\right) \in [0, \pi] \tag{2}$$

The Eq. (1) is the Chebyshev approximation polynomials which has high approximation accuracy and is close to the best uniform approximation polynomial (Fox and Parker, 1968). Based on the orthogonality of the Chebyshev polynomials, the coefficients can be obtained via the following integral formula

$$f_i = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)C_i(x)}{\sqrt{1-x^2}}dx \approx \frac{2}{\pi}\frac{\pi}{m}\sum_{j=1}^{m} f(x^{(j)})C_i(x^{(j)}) = \frac{2}{m}\sum_{j=1}^{m} f(\cos\theta^{(j)})\cos i\theta^{(j)} \tag{3}$$

Here the Gauss-Chebyshev quadrature is used to produce the Eq. (3). m is the order of the numerical integral formula, $x^{(j)}$ and $\theta^{(j)}$ are the interpolation points of the integral formula in the $\mathbf{x}$ space and $\boldsymbol{\theta}$ space, respectively. The interpolation points are the zeros of the Chebyshev polynomials of degree m, given by

$$x^{(j)} = \cos\theta^{(j)},\ \text{where}\ \theta^{(j)} = \frac{2j-1}{m}\frac{\pi}{2}, j = 1, 2, ..., m \tag{4}$$

To guarantee the accuracy of the Gauss-Chebyshev quadrature in computing the highest order coefficient f_n, the order m in Eq. (3) should be larger than n, i.e. $m \geq n+1$. More details about the Chebyshev polynomial can be found in references (Wu et al., 2013a, Wu et al., 2013b, Wu et al., 2014b).

Noting $\boldsymbol{\beta} = \left[\beta_0 \quad \beta_1 \quad \cdots \quad \beta_n\right]^{\mathrm{T}} = \left[f_0/2 \quad f_1 \quad \cdots \quad f_n\right]^{\mathrm{T}}$, the Eq. (1) can be transformed to the following expression

$$f(x) \approx \frac{1}{2} f_0 + \sum_{i=1}^{n} f_i C_i(x) = \sum_{i=0}^{n} \beta_i C_i\left(x\right) \tag{5}$$

In this case, the coefficients β_i can be calculated by using the least square method directly

$$\boldsymbol{\beta} = \left[\beta_0 \quad \beta_1 \quad \cdots \quad \beta_n\right]^{\mathrm{T}} = \left(\mathbf{C}^{\mathrm{T}}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{f} \tag{6}$$

where **C** is the transform matrix with size $m\times(n+1)$ at the interpolation points, and **f** is the column vector including function values over all the interpolation points $\mathbf{x}=\left[\begin{array}{ccc} x^{(1)} & \cdots & x^{(m)} \end{array}\right]^{\mathrm{T}}$

$$\mathbf{C}=\left[\begin{array}{cccc} C_0(\mathbf{x}) & C_1(\mathbf{x}) & \cdots & C_n(\mathbf{x}) \end{array}\right], \text{ and } \mathbf{f}=f(\mathbf{x}) \tag{7}$$

If the interpolation points used in Eq. (6) are the same as the interpolation points used in Eq. (3), the two formulas will produce the same coefficients. When the $m=n+1$, the Eq. (5) will be an n-th order interpolation polynomial. Generally, when the order of interpolation points is too high, the numerical instability (i.e. the Runge phenomenon) may happen. However, this numerical problem is avoided by using the zeros of Chebyshev polynomials as the interpolation points (Fox and Parker, 1968). This is the reason why we use the zeros of Chebyshev polynomials as sampling candidates to construct the surrogate model.

Considering Eqs. (5)-(7), it can be found that the procedure of constructing the Chebyshev approximation polynomial is just the same process of constructing a response surface. Therefore, when the sampling points are selected as the zeros (**x** vector) of the Chebyshev polynomials $C_i(x)$, we can get the approximated Chebyshev series $p_n(x)$, which is a polynomial with order n. As aforementioned, it is termed as high-order polynomial surrogate model (HOPSM). The Eq. (5) is the expression of 1-dimensional HOPSM, the multi-dimensional HOPSM will be developed in the following subsection.

2.2 The multi-dimensional HOPSM

Extending the 1-dimensional space to k-dimensional space, the input variables are expressed as $\mathbf{x}\in[-1,1]^k$. The HOPSM of the k-dimensional continuous function $f(\mathbf{x})$ with order n can be expressed as the tensor product of the HOPSM, with respect to each dimensional variable x_i (Wu et al., 2014b)

$$\hat{f}_n(\mathbf{x})=\sum_{0\le i_1,\ldots,i_k\le n}\beta_{i_1\ldots i_k}C_{i_1}(x_1)\ldots C_{i_k}(x_k), \text{ and } i_1,\ldots,i_k=0,1,\ldots,n \tag{8}$$

where $\beta_{i_1\ldots i_k}$ denotes the coefficients. The subscript of coefficients $\beta_{i_1\ldots i_k}$ forms a "hypercube" (from 0 to n), so Eq. (8) is called the hypercube format of HOPSM. It should be noted that the number of the coefficients need to be determined in Eq. (8) is $(n+1)^k$, which is also equal to the least required sampling points. It can be found that the sampling size will be extremely large when k and n are large, so it is not suitable for high dimensional problems.

Besides the hypercube format, the HOPSM can also be expressed as a "simplex" format, which only contains the terms when the order without exceeding n

$$\hat{f}_n(\mathbf{x})=\sum_{0\le i_1+\ldots+i_k\le n}\beta_{i_1\ldots i_k}C_{i_1}(x_1)\ldots C_{i_k}(x_k), \text{ and } i_1,\ldots,i_k=0,1,\ldots,n \tag{9}$$

It should be noted that Eq. (9) is different from Eq. (8), where the subscript in the former forms a simplex while the Eq. (8) forms a hypercube. As a result, the number of coefficients to be estimated in Eq. (9) is much less than that in Eq. (8) for high dimensional problems, which is given by

$$N_C\left(n,k\right)=\frac{\left(k+n\right)!}{k!n!} \tag{10}$$

To reduce the number of coefficients to be estimated, we use the simplex format to build the HOPSM. To simplify the notation, noting $\psi_i(\mathbf{x})$ as the multi-dimensional Chebyshev polynomials, the relationship between $\psi_i(\mathbf{x})$ and $C_i(x)$ is given by

$$\begin{aligned}
&\psi_1(\mathbf{x})=1,\\
&\psi_2(\mathbf{x})=C_1\left(x_1\right),...,\psi_{N_C(1,k)}(\mathbf{x})=C_1\left(x_k\right),\\
&\psi_{N_C(1,k)+1}(\mathbf{x})=C_2\left(x_1\right),\psi_{N_C(1,k)+2}(\mathbf{x})=C_1\left(x_1\right)C_1\left(x_2\right),...,\psi_{N_C(2,k)}(\mathbf{x})=C_2\left(x_k\right),\\
&...
\end{aligned} \tag{11}$$

Substituting Eq. (11) into Eq. (9), the HOPSM of k-dimensional function with order n can be expressed by the following equation

$$\hat{f}_n\left(\mathbf{x}\right)=\sum_{i=1}^{N_C(n,k)}\alpha_i\psi_i\left(\mathbf{x}\right) \tag{12}$$

The coefficient α_i is corresponding to the coefficient $\beta_{i_1...i_k}$, and it can still be produced by using the LSM, similar to the Eq. (6)

$$\boldsymbol{\alpha}=\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_{N_C(n,k)}\end{bmatrix}^{\mathrm{T}}=\left(\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi}\right)^{-1}\boldsymbol{\psi}^{\mathrm{T}}\mathbf{f} \tag{13}$$

where the $\boldsymbol{\psi}$ is the transform matrix at the sampling points, and $\mathbf{f}$ is the column vector of original function value at the sampling points, given be

$$\boldsymbol{\psi}=\begin{bmatrix}\psi_1\left(\mathbf{x}^{(1)}\right) & \cdots & \psi_{N_C(n,k)}\left(\mathbf{x}^{(1)}\right)\\ \vdots & \ddots & \vdots\\ \psi_1\left(\mathbf{x}^{(s)}\right) & \cdots & \psi_{N_C(n,k)}\left(\mathbf{x}^{(s)}\right)\end{bmatrix},\mathbf{f}=\begin{bmatrix}f\left(\mathbf{x}^{(1)}\right)\\ \vdots\\ f\left(\mathbf{x}^{(s)}\right)\end{bmatrix} \tag{14}$$

Here, $\mathbf{x}^{(i)}$ is the sampling point, and s is the number of sampling points. Based on the 1-dimensional HOPSM, the sampling points of k-dimensional HOPSM can still be selected as the zeros of Chebyshev polynomials. Using the tensor product operation, the zeros of multi-dimensional Chebyshev polynomials are expressed by

$$\boldsymbol{X}=\mathbf{x}_1\otimes...\otimes\mathbf{x}_k,\quad \boldsymbol{\theta}=\boldsymbol{\theta}_1\otimes...\otimes\boldsymbol{\theta}_k \tag{15}$$

where $\mathbf{x}_i$ and $\boldsymbol{\theta}_i$ denote the zeros of the Chebyshev polynomial of the ith variable in $\mathbf{x}$ space and $\boldsymbol{\theta}$ space, respectively. Based on Eq. (4), the $\mathbf{x}_i$ is expressed by

$$\mathbf{x}_i = \cos\boldsymbol{\theta}_i, \boldsymbol{\theta}_i = \left\{ \theta^{(j)} = \frac{2j-1}{m}\frac{\pi}{2}, j = 1, 2, ..., m \right\} \tag{16}$$

Considering the Eq. (15), the number of zeros of *k*-dimensional Chebyshev polynomials with order *m* is equal to

$$N_S\left(m,k\right) = m^k \tag{17}$$

As aforementioned in Section 2.1, the *m* should satisfy the condition $m \geq n+1$. It can be found that the $N_S\left(m,k\right) >> N_C\left(n,k\right)$ when the *k* and *n* are relatively large, so we cannot use all the zeros of multi-dimensional Chebyshev polynomials as the sampling points, otherwise the computational cost of sampling will be extremely expensive. In fact, as long as the sampling size is not smaller than the number of coefficients to be estimated, the LSM can be implemented. Hence, we can use the zeros of *k*-dimensional Chebyshev polynomials as a sampling candidate set, and then choose the most fitful candidates as the final sampling points. How to select the sampling points will be mainly studied in Section 3.

2.3 The evaluation index of surrogate models

Since the LSM is employed to calculate the coefficients of HOPSM, the numerical stability is an important evaluation index of this surrogate model. Considering the Eq. (13), there is an inverse operation of matrix $\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi}$. If the matrix is ill-conditioned, there will be a large numerical error for the inverse operation. The condition number is used to describe the ill-condition of a matrix, i.e. a low condition number is said to be well-conditioned, while a matrix with a high condition number is said to be ill-conditioned. Denoting $\mathbf{A} = \boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi}$, the condition number of the **A** matrix is defined as

$$\mathrm{COND}\left(\mathbf{A}\right) = \left\|\mathbf{A}^{-1}\right\| \cdot \left\|\mathbf{A}\right\| \tag{18}$$

Since the **A** is a normal matrix, the condition number can be further expressed as

$$\mathrm{COND}\left(\mathbf{A}\right) = \left|\frac{\lambda_{\max}\left(\mathbf{A}\right)}{\lambda_{\min}\left(\mathbf{A}\right)}\right| \tag{19}$$

where $\lambda_{\max}\left(\mathbf{A}\right)$ and $\lambda_{\min}\left(\mathbf{A}\right)$ are the maximal and minimal eigenvalues of the **A** matrix, respectively.

To evaluate the accuracy denoting the fitting goodness of the surrogate model, some other evaluation indexes should be introduced. The root-mean-square error (RMSE) and average absolute error (AAE) are widely employed to describe global accuracy, while the maximum absolute error (MAE) is usually used to indicate the local accuracy. In order to avoid ambiguity or to enable comparisons of surrogate models across disciplines, then relative error averages are sometimes used in the literature, such as AAE relative to the standard deviation, R-square (Jin et al., 2001), or AAE and RMSE measures relative to the average response *f* (Qu et al., 2004). This paper will use R-square, AAE, and MAE to describe the accuracy, but some normalized operation will be used to make these indexes more clearly.

The R-square, denoted as R^2, is a statistical characteristic, which is expressed as follows:

$$R^2 = 1 - \frac{SSE}{SST} \tag{20}$$

where SSE and SST denote the residual sum of square and the total sum of square, respectively

$$SSE = \sum_{i=1}^{N} \left(\hat{f}_n(\mathbf{x}_i) - f(\mathbf{x}_i) \right)^2 \text{ and } SST = \sum_{i=1}^{N} \left(f(\mathbf{x}_i) - \bar{f} \right)^2 \tag{21}$$

where $\bar{f}$ is the mean value of observed data (real value at the test points), $\mathbf{x}_i$ is the test point, and N is the size of test points. The R^2 can also be used to pre-estimate the accuracy by using the sampling points as the test points in the construction of HOPSM. Generally, the larger value of R^2 means higher fitness of the regression model. However, if the sampling points are used as the test points, the R^2 will improve when the number of estimated coefficients increases, so we need to adjust the R-square to pre-estimate the fitting degree of regression model. The adjusted R-square $\overline{R}^2$ has been widely used, defined as

$$\overline{R}^2 = 1 - \frac{SSE / \left(s_0 - N_C(n,k) - 1 \right)}{SST / \left(s_0 - 1 \right)} \tag{22}$$

where s_0 denotes the number of samples, $N_C(n,k)$ is the number of estimated coefficients given by Eq. (10). In this case, even when the number of coefficients increases, the adjusted R-square $\overline{R}^2$ may still decrease. $\overline{R}^2$ will be used to pre-estimate the fitness of HOPSM in the model constructing period, while R^2 will be used to indicate its actual fitness in the validation period.

The definition of AAE and MAE are given as

$$\text{AAE} = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{f}_n(\mathbf{x}_i) - f(\mathbf{x}_i) \right| \text{ and } \text{MAE} = \max_i \left| \hat{f}_n(\mathbf{x}_i) - f(\mathbf{x}_i) \right| \tag{23}$$

In solving different problems, the AAE and MAE may change in large range. To make the two indexes more clearly, we transform them into the normalized average absolute error and the normalized maximum absolute error, respectively, noted as NAAE and NMAE.

$$\text{NAAE} = \text{AAE} \Big/ \left(\frac{1}{N} \sum_{i=1}^{N} \left| f(\mathbf{x}_i) \right| \right) \text{ and } \text{NMAE} = \text{MAE} \Big/ \left(\max_i f(\mathbf{x}_i) - \min_i f(\mathbf{x}_i) \right) \tag{24}$$

The smaller values of NAAE and NMAE are, the better the surrogate model is. Both NAAE and NMAE will be used in constructing the HOPSM, to pre-estimate the fitness by using the sampling points as the test points, which will be illustrated in the following section.

3. Sequential Chebyshev sampling method for HOPSM

3.1 The sampling candidate set

As described in Section 2, the sampling points of k-dimensional HOPSM are chosen from a candidate set, which is comprised of the tensor product of the zeros of Chebyshev polynomials with order m, and the m

should be larger than n (the order of HOPSM). How to choose the order of sampling candidate set m is important, because a very small m may reduce the accuracy of the surrogate model, while a very large m will usually waste the sampling information.

We choose m based on the highest allowable order of HOPSM, noted by $n_{\max}$, which can be determined by the most allowable sampling size $N_{\max}$. In practical application, the $N_{\max}$ can be given through the allowable computation cost. For example, if the allowable computation time is 10 hours, and each running of model simulation is 1 min, then the $N_{\max}$ should be 10 hours divided by 1 min, i.e. $N_{\max}$ =360. It should be noted that $N_{\max}$ just restricts the maximal allowable sampling size, but the actual number of sampling points may be much smaller than it due to the sequential sampling strategy. The number of coefficients must be smaller than the sampling size, so $n_{\max}$ will be defined as

$$n_{\max} = \left\{ \max\ n \;\middle|\; N_C(n,k) = \frac{(k+n)!}{k!n!} < N_{\max} \right\} \tag{25}$$

The order m of the candidate set should be larger than $n_{\max}$, so it will be set as $m=n_{\max}+1$ to minimize the size of the candidate set. We should note that the number of sampling candidates N_S shown in Eq. (26) is much larger than the number of the coefficients N_C, especially for the large k.

$$N_S(m,k) = (n_{\max}+1)^k \tag{26}$$

For example, if the $N_{\max}$=100, and k=2 or 3, then $n_{\max}$=12 or 6, so m=13 or 7, respectively. The plot of candidate sets in **x** space and **θ** spaces for the two cases are shown in Figs. 1 and 2. It is noted that the candidate set is a symmetrical grid in **θ** space but an asymmetrical grid in **x** space. The samples distribute denser in the regions closing the boundaries of the **x** design space, which is different from the traditional equidistant full factor design.

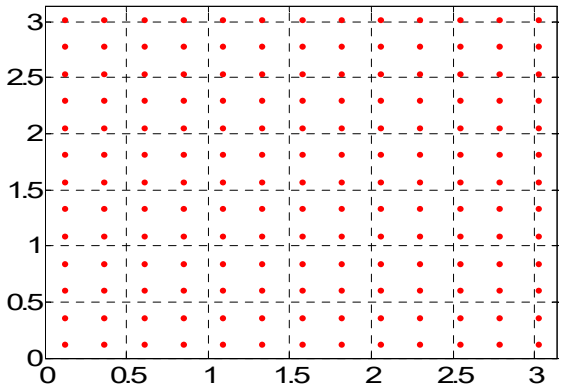


Figure.1 (a) Candidates for k=2 in **θ** space

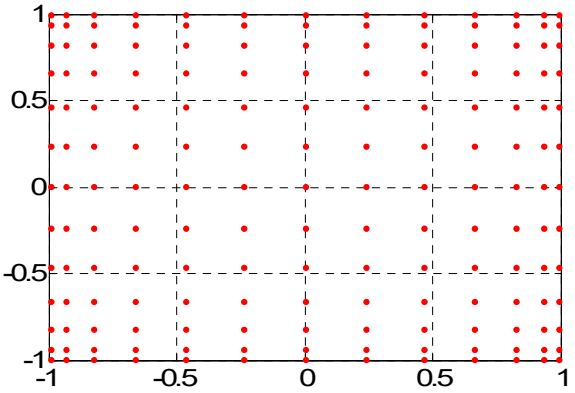


Figure.1 (b) Candidates for k=2 in **x** space

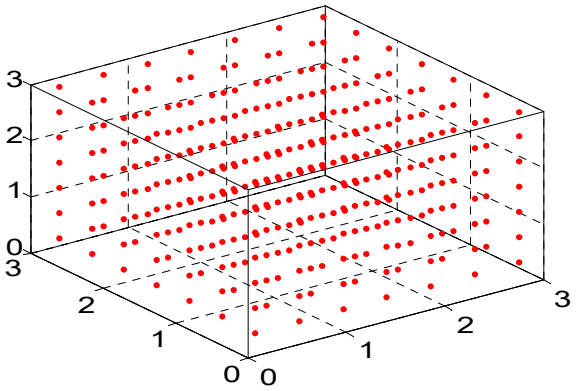


Figure.2 (a) Candidates for k=3 in $\boldsymbol{\theta}$ space

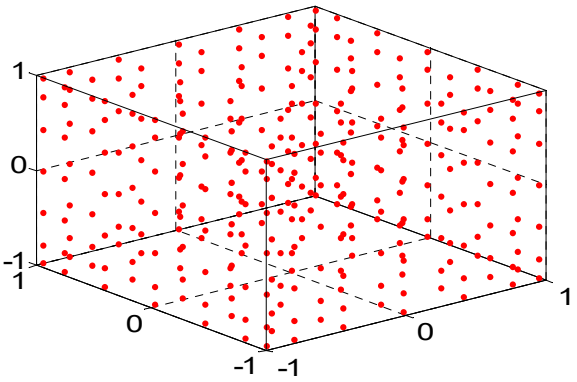


Figure.2 (b) Candidates for k=3 in $\mathbf{x}$ space

3.2 The coordinate alternative algorithm for initial sampling

Generally, the initial samples are expected to distribute uniformly in the entire space. One of the most widely used measures to evaluate the uniformity of a sampling set is the *maximin* metric introduced by Johnson et al (Jonson et al., 1990). The following scalar-valued criterion function (Morris and Mitchell, 1995, Viana et al., 2009) is mainly used to rank competing sampling set as

$$\Phi_p\left(\boldsymbol{\Theta}\right)=\left(\sum_{i=1}^{s}\sum_{j=i+1}^{s} d\left(\boldsymbol{\theta}^{(i)},\boldsymbol{\theta}^{(j)}\right)^{-p}\right)^{1/p} \tag{27}$$

where the p is a large positive integer and set as 100 in this paper, $\boldsymbol{\Theta}$ is the sampling set, and the distance $d\left(\boldsymbol{\theta}^{(j_1)},\boldsymbol{\theta}^{(j_2)}\right)$ is measured by the Euclidean norm in the $\boldsymbol{\theta}$ space:

$$d\left(\boldsymbol{\theta}^{(j_1)},\boldsymbol{\theta}^{(j_2)}\right)=\sqrt{\sum_{i=1}^{k}\left|\theta_i^{(j_1)}-\theta_i^{(j_2)}\right|^2} \tag{28}$$

A smaller Φ_p indicates more uniformity of the sampling set. However, minimizing Φ_p is an NP-complete problem. In this study, a new sequential algorithm will be proposed to produce the uniformly distributed initial samples. To seek the uniformity, the sampling points should be located in all levels in each dimension, and they are expected to locate in each level equivalently.

We propose a coordinate alternation algorithm to fast produce the initial sampling points which may be located in the design space uniformly. The number of initial samples is set to N_0=mk. To simplify the notation, we use the levels (from 1 to m) to denote the location of samples. The samples can be expressed by a matrix with a size of k×mk, where the column denotes the sequence of variables and row denotes the sequence of sampling points. The first m elements of the first row (x_1) are set as 1, 2 … m sequentially, which makes the samples locate in each level of x_1. The level of the first column for the ith variables is set as i, which gives the first sampling point. If the order m is smaller than i, the level of the ith variable will be set as the remainder of i/m. The first row is called as the reference row, marked by "*". For example,

considering the case m=4 and k=3, the matrix is given in Table 1, while numbers in the reference row are marked with blue colour.

Table 1. The initial initialization design matrix for k=3, m=4 (the first step)

No. of samples / Design variables	1	2	3	4	5	6	…	12
x_1*	**1**	**2**	**3**	**4**				
x_2	**2**							
x_3	**3**							

Table 1 has given the first column and the first 4 elements of the first row, so we need to determine the remaining elements to uniformly distribute the samples. Choose new sampling points in sequentially, through minimizing the *maximin* metric Φ_p. Denote the sampled set as $\mathbf{\Theta}$ and the remaining candidate set as $\mathbf{\Psi}$. When a new point from $\mathbf{\Psi}$ is added into the sampled set $\mathbf{\Theta}$, this point should minimize Φ_p of the new sampled set comprised by $\mathbf{\Theta}$ and the new sampling point $\boldsymbol{\theta}_1^{(j)}$, expressed as

$$\Phi_p\left(\mathbf{\Theta},\boldsymbol{\theta}_1^{(j)}\right)=\left(\sum_{i_1=1}^{s_0}\sum_{i_2=i_1+1}^{s_0} d\left(\boldsymbol{\theta}_0^{(i_1)},\boldsymbol{\theta}_0^{(i_2)}\right)^{-p}+\sum_{i_1=1}^{s_0} d\left(\boldsymbol{\theta}_0^{(i_1)},\boldsymbol{\theta}_1^{(j)}\right)^{-p}\right)^{1/p}=\left(\Phi_p\left(\mathbf{\Theta}\right)^p+\sum_{i_1=1}^{s_0} d\left(\boldsymbol{\theta}_0^{(i_1)},\boldsymbol{\theta}_1^{(j)}\right)^{-p}\right)^{1/p} \quad (29)$$

where $\boldsymbol{\theta}_0^{(i)}\in\mathbf{\Theta}$ are the sampled points, $\boldsymbol{\theta}_1^{(j)}\in\mathbf{\Psi}$ is the new sampling point, and s_0 is the number of points in the sampled set. Since the first term in the right side keeps unchanged, the minimizing operation will be applied to the second term, noted as

$$\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0\right)=\left(\sum_{i=1}^{s_0} d\left(\boldsymbol{\theta}_0^{(i)},\boldsymbol{\theta}_1^{(j)}\right)^{-p}\right)^{1/p} \quad (30)$$

Therefore, the new sampling points can be chosen by minimizing ϕ_p. If all the candidates are directly used to calculate ϕ_p, the computational cost will still be expensive as O(m^k). Producing the level of each dimensional variable sequentially, the number of calculation will be reduced to O(m(k-1)).

To describe the procedure more clearly, we still use the data shown in Table 1. The 3-dimensional design space with 4 levels is shown as the Fig. 3, where the first point has been shown as the black point (the first column in Table1). We firstly calculate the second element of the second column (the level of the 2nd sampling point in the 2nd variable). Since the level of x_1 for the second sampling point has been given as 2 first element, the second sampling point will be chosen from those points located in the plan x_1=2, shown as the yellow frame in Fig. 3(a). In this period, only the space comprised by x_1 and x_2 is considered, so there are only four candidates, shown as the blue points, are considered. After the

evaluation of Eq. (30), the ϕ_p get the minimum value when the level of the second variable is 4, i.e. x_2=4. Go to the third row of the same column, and determine the level of the third variable by minimizing ϕ_p.

The first two elements of the second sample have been determined, so the sampling points will be selected from those candidates located on the line x_1=2 and x_2=4, which is the red line shown in the Fig. 3(b). However, the distance in Eq. (30) will be calculated in the 3-dimensional space in this case. After the evaluation, we get x_3=1 that minimizes ϕ_p. Therefore, we have obtained all the elements of the second column, and fill it in the Table 2. We can find that, only 8 candidates are evaluated to produce the second sampling point, the computational cost has been reduced much related to the number of candidates. After the second column is determined, the sampled set **Θ** should be updated, and then we can calculate the elements in the third column based on the same approach. Repeat the process until the *m*-th column is determined, and the results of the first 4 sampling points are shown in Table 2.

Table 2. The initial design matrix for k=3 and m=4(the second step)

Design variables \ No. of samples	1	2	3	4	5	6	…	12
x_1*	**1**	**2**	**3**	**4**				
x_2	**2**	**4**	**1**	**3**				
x_3	**3**	**1**	**1**	**4**				

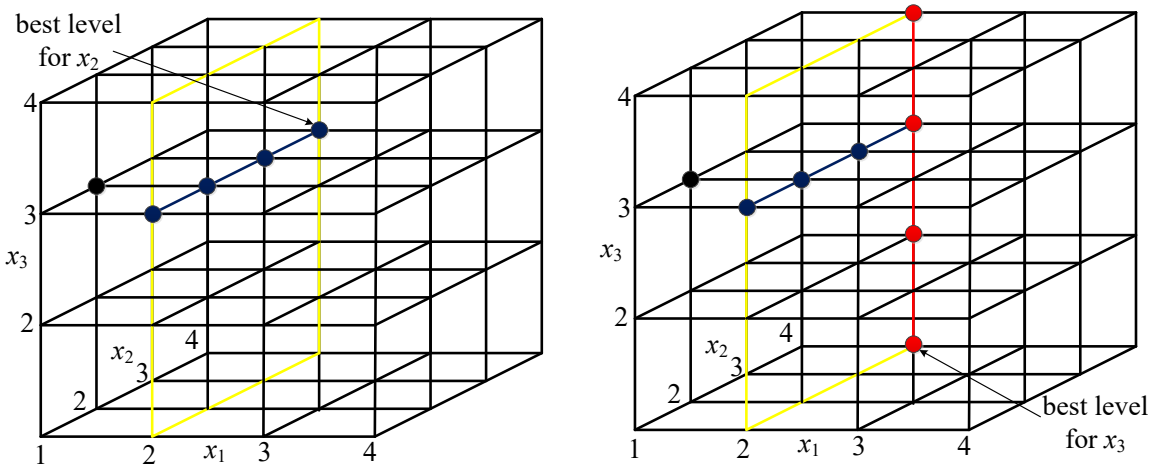


Figure. 3 The procedure of the coordinate alternative algorithm

The remaining columns need to be determined, which can be implemented by using the same approach. However, we can find that the levels of the third variable in Table 2 are not distributed uniformly. To make the levels of every design variable be distributed uniformly, we can pre-fix the levels of each design

variable by moving the first row to the last row and moving other rows forward in sequence, so this method is called as coordinate alternative algorithm. After that, the reference row will be the second variable x_2. Pre-fix the elements from the column 5 to 8 of the reference row as 1 to 4, shown in Table 3, and then determine the levels of other rows of columns 5 to 8. The rows-moving operation will be implemented per m new samples are determined. The rows-moving operation will guarantee every level occurring in each variable at least once.

Table 3. The initial design matrix for k=3 and m=4 (the third step)

No. of samples / Design variables	1	2	3	4	5	6	7	8	9	10	11	12
x_2*	**2**	**4**	**1**	**3**	**1**	**2**	**3**	**4**				
x_3	**3**	**1**	**1**	**4**								
x_1	**1**	**2**	**3**	**4**								

Table 4. The initial design matrix for k=3 and m=4

No. of samples / Design variables	1	2	3	4	5	6	7	8	9	10	11	12
x_1	**1**	**2**	**3**	**4**	**3**	**1**	**4**	**1**	**4**	**2**	**3**	**2**
x_2	**2**	**4**	**1**	**3**	**1**	**2**	**3**	**4**	**4**	**1**	**4**	**3**
x_3	**3**	**1**	**1**	**4**	**4**	**1**	**2**	**3**	**1**	**2**	**3**	**4**

After two rows-moving operations, the 12 initial sampling points will be produced, shown in Table 4. The plot of the initial sampling points is shown in Fig. 3, expressed through the level. The number of each level located in each dimensional variable, termed as occurrence number, is summarized in Table 5. It can be found that the occurrence number for each level of each dimensional variable is close to 3 (changing from 2 to 4), so the samples distribute uniformly.

Table 5. The occurrence number of each dimensional variable

No. of levels / Design variables	1	2	3	4
x_1	**3**	**3**	**3**	**3**
x_2	**3**	**2**	**3**	**4**
x_3	**4**	**2**	**3**	**3**

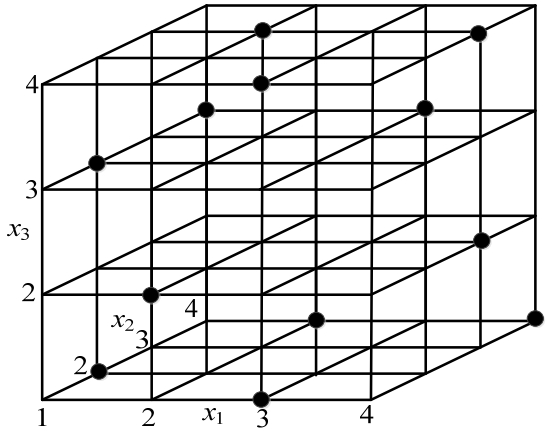


Figure.4 The distribution of initial sampling points

3.3 The fast sequential sampling scheme

The exploration strategy, to fill up the domain as evenly as possible, will be used to choose more sampling points. The new samples can still be chosen by minimizing ϕ_p in Eq. (30), to locate the new sample to the region without containing samples. This can be described by the following optimization

$$\begin{aligned} &\text{Find:} \quad \boldsymbol{\theta}_1^{(j)}, \quad j = 1, ..., s_1 \\ &\min: \quad \phi_p\left(\boldsymbol{\theta}_1^{(j)}, s_0\right) \end{aligned} \tag{31}$$

where s_1 denotes the number of points in the remaining candidate set $\boldsymbol{\Psi}$.

The sampling point can be chosen by performing the above optimization one by one, until the termination condition is satisfied. However, in each exploration, the optimization model requires to compute the distance d in O($s_0 \times s_1$) time. This process may be expensive computational cost, especially when k and m are relatively large, since s_1 is close to m^k. However, we do not need to compute the *maximin* criterion for all candidates except in the first exploration, as ϕ_p always keeps increasing after each exploration, which can be used to reduce the computational cost.

For the first exploration, the *maximin* criterion ϕ_p of each candidate point will be calculated firstly. Then we can sort the candidates in ascending order of ϕ_p, such that $\phi_p(\boldsymbol{\theta}_1^{(1)}, s_0) \le \cdots \le \phi_p(\boldsymbol{\theta}_1^{(s_1)}, s_0)$. The candidate that has the minimum $\phi_p(\boldsymbol{\theta}_1^{(j)}, s_0)$ will be chosen as the new sampling point, and so the first candidate $\boldsymbol{\theta}_1^{(1)}$ will be the new sampling point. If we consider the case of k=2 and m=7, the Fig. 5(a) is used to illustrate this process directly. There are 49 candidates on the 7 by 7 lattice, where the numbers on the coordinates are the levels of each design variables. The sampled set $\boldsymbol{\Theta}$ is composed by the14 black points (s_0=14) which are selected by using the previous coordinate alternation algorithm, and the candidate set $\boldsymbol{\Psi}$ is composed by the remaining red and blue points (s_1=35). After the evaluation of ϕ_p and sort of all the remaining candidates, the sequential number is noted to all the candidates in italic number.

The red point in the coordinate x_1=2, x_2=5 is the No. 1 point which minimizes the ϕ_p. Therefore, the red point will be chosen as the new sampling point.

After the new sampling point is picked, we can update the amount of samples to $s_0+1=15$. In this way, a new sample ($\boldsymbol{\theta}_0^{(s_0+1)}=\boldsymbol{\theta}_1^{(1)}$) will be added into the sampled set $\boldsymbol{\Theta}$, while the candidate $\boldsymbol{\theta}_1^{(1)}$ is also removed from the candidate set $\boldsymbol{\Psi}$. Subtracting 1 for the sequential number of each candidate, the remaining candidates still keep the ascending sequence as $\phi_p(\boldsymbol{\theta}_1^{(1)},s_0)\le\cdots\le\phi_p(\boldsymbol{\theta}_1^{(s_1-1)},s_0)$. After the first exploration, the distribution of sampling points (black points) and the candidates (green points and blue points) are shown as Fig. 5(b).

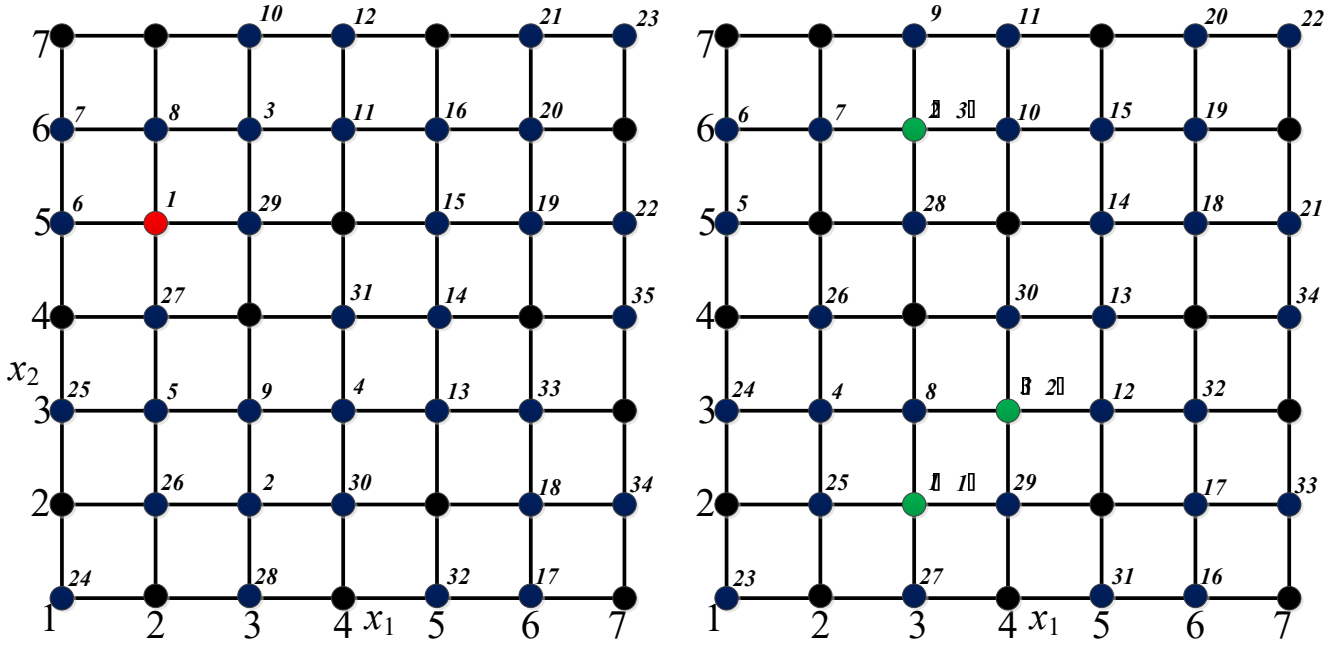


Figure. 5 The process of fast sequential sampling scheme

For the second exploration, based on the Eq. (30), ϕ_p can be calculated as follows:

$$\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0+1\right)=\left(\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0\right)^p+d\left(\boldsymbol{\theta}_0^{(s_0+1)},\boldsymbol{\theta}_1^{(j)}\right)^{-p}\right)^{1/p} \tag{32}$$

Eq. (32) will save computational cost related to Eq. (30), as we have known $\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0\right)$. Calculating $\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0+1\right)$ in ascending order until a candidate is found to satisfy $\phi_p\left(\boldsymbol{\theta}_1^{(J)},s_0+1\right)\le\phi_p\left(\boldsymbol{\theta}_1^{(J+1)},s_0\right)$, which means the *J*-th candidate is better than the candidates with larger sequential number than *J*, as the *maximin* criterion always increases after each exploration ($\phi_p\left(\boldsymbol{\theta}_1^{(j)},s_0+1\right)>\phi_p(\boldsymbol{\theta}_1^{(j)},s_0)$). The *J* is termed as criterion sequential number. This operation would also save more computational cost, because most of

the candidates which are close to the sampled points can never be selected in the exploration. The ϕ_p of the first J candidates have been updated, denoted as $\phi_p\left(\boldsymbol{\theta}_1^{(j)}, s_0+1\right)$ for $j \le J$, while other candidates have not updated yet, expressed by $\phi_p(\boldsymbol{\theta}_1^{(j)}, s_0)$ for $j > J$. To distinguish them, we introduce another notation M_j, which denotes the number of points in the sampled set used to calculate the ϕ_p of $\boldsymbol{\theta}_1^{(j)}$. Therefore, ϕ_p of $\boldsymbol{\theta}_1^{(j)}$ is re-noted as $\phi_p(\boldsymbol{\theta}_1^{(j)}, M_j)$. After the second evaluation of ϕ_p, $M_j = s_0 + 1\left(j \le J\right)$, $M_j = s_0\left(j > J\right)$. Sorting the remaining candidates in ascending order once again, the first candidate $\boldsymbol{\theta}_1^{(1)}$ will be added into the sampled set $\boldsymbol{\Theta}$ and removed from the candidate set $\boldsymbol{\Psi}$. This process can be described by Fig. 5(b) more directly. After the evaluation of the first three candidates (green points), the computation is terminated, i.e. the criterion sequential number is J=3. Sort all the remaining candidates in ascending order, where the updated sequential number is shown in the bracket. The No. 1 candidate with coordinate x_1=3, x_2=2 will be chosen as the new sampling point of the second exploration. It can be found that only the ϕ_p of the first three candidates are evaluated, which saves much computational cost.

After the second exploration, in a more general case, the ϕ_p can be calculated by

$$\phi_p\left(\boldsymbol{\theta}_1^{(j)}, s\right) = \left(\left(\phi_p\left(\boldsymbol{\theta}_1^{(j)}\right), M_j\right)^p + \sum_{i=1}^{s-M_j} d\left(\boldsymbol{\theta}_0^{(s-i+1)}, \boldsymbol{\theta}_1^{(j)}\right)^{-p}\right)^{1/p} \tag{33}$$

where s is the size of the sampled set before this exploration.

Repeat the same process for the second exploration, another new sample can be selected. After the evaluation of ϕ_p, the M_j should be updated, i.e. $M_j = s\left(j < J\right)$. The procedure of the proposed exploration can be described by the following **Algorithm 1**. When the number of sampling points exceeds the required samples $N_C(n,k)$, the least square method can be used to construct the HOPSM.

Algorithm 1

Input: k, p, $\boldsymbol{\Theta}$, $\boldsymbol{\Psi}$, s_0, s_1

Initialize: $s = s_0$, $M_j = 0$, $\phi_p\left(\boldsymbol{\theta}_1^{(j)}, M_j\right) = 0$, $\left(j = 1, ..., s_1\right)$

Compute: Do

for j=1:s_1

$\phi_p\left(\boldsymbol{\theta}_1^{(j)}, s\right)$; $M_j = s$;

if $\phi_p\left(\boldsymbol{\theta}_1^{(j)}, M_j\right) \le \phi_p\left(\boldsymbol{\theta}_1^{(j+1)}, M_{j+1}\right)$

Break

end if

end for

Sort: $\phi_p\left(\boldsymbol{\theta}_1^{(j)}, M_j\right),\ \left(j=1,...,s_1\right)$;

$s=s+1; s_1=s_1-1;$

$\boldsymbol{\theta}_0^{(s)}=\boldsymbol{\theta}_1^{(1)}$ (Add $\boldsymbol{\theta}_1^{(1)}$ into $\boldsymbol{\Theta}$);

$\phi_p\left(\boldsymbol{\theta}_1^{(j)}, M_j\right)=\phi_p\left(\boldsymbol{\theta}_1^{(j+1)}, M_{j+1}\right),\ \boldsymbol{\theta}_1^{(j)}=\boldsymbol{\theta}_1^{(j+1)}, M_j=M_{j+1}, j=1,...,s_1$ (Remove $\boldsymbol{\theta}_1^{(1)}$ from $\boldsymbol{\Psi}$);

End Do

Output: $\boldsymbol{\Theta}, \boldsymbol{\Psi}, s, s_1$

3.4 The incremental modelling of HOPSM

The order of HOPSM will largely influence the approximation accuracy. A problem to be solved is that which order is mostly suitable for balancing accuracy and efficiency, since the higher order polynomial may have higher approximation accuracy but requires more samples. This problem can be solved by the order increment strategy, which changes the polynomials from low-order to high-order gradually. The criterions of order increase and termination are based on the evaluation indexes given in Section 2.

The order of HOPSM starts from 2, and then it will increase successively until the termination conditions are satisfied. When the order of HOPSM keeps fixed, the approximation accuracy of the surrogate model will improve gradually if the data points are added using the sequential sampling scheme. However, the improvement of the approximation accuracy will become very small when the sampling size has got a certain large number, which is named as the sampling saturation. When the sampling has saturated, the increase the order of HOPSM to 3 may further improve the approximation accuracy. After enhancing the order, the sampling saturation may become unsaturation, as the number of coefficients to be estimated has increased. Therefore, the sequential sampling procedure will be implemented again until the sampling saturation is reached. Then the implementation of the order increment operation was followed by the sequential sampling process repeatedly. Three evaluation indexes are used to judge when the order incremental operation should be implemented.

In the sampling saturation period, the three evaluation indexes have little variation, which can be used as the variation of these evaluation indexes as the criterion of incremental operation. The variation of the three evaluation indexes can be defined as follows:

$$\Delta\overline{R}_s^2=\left|\overline{R}_s^2-\overline{R}_{s-k}^2\right|,\ \Delta\mathrm{NAAE}_s=\frac{\left|\mathrm{NAAE}_s-\mathrm{NAAE}_{s-k}\right|}{\mathrm{NAAE}_s}\ \text{and}\ \Delta\mathrm{NMAE}_s=\frac{\left|\mathrm{NMAE}_s-\mathrm{NMAE}_{s-k}\right|}{\mathrm{NMAE}_s} \quad (34)$$

where the subscript denotes the number of sampling points, and k is the dimensional size of design variables. The incremental criterions are given by

$$\Delta\overline{R}_s^2<\varepsilon_1, \Delta\mathrm{NAAE}_s<\varepsilon_2,\ \text{and}\ \Delta\mathrm{NMAE}_s<\varepsilon_3 \quad (35)$$

where $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are some small positive percentage, e.g. 1%.

The order will increase when all the three inequalities are satisfied. To guarantee the robustness of surrogate modelling, the sampling size should not be smaller than the twice of coefficients size, so the following condition should be satisfied

$$s \geq 2N_C(n,k) \tag{36}$$

When one of the criterions in (35) and (36) is satisfied, the order increment operation will be performed. The following termination criterions are also based on the three indexes. The first one is the same as incremental criterion given in inequality (28), while the second is expressed as follows

$$\overline{R}_s^2 \geq \mu_1, \text{NAAE}_s \leq \mu_2, \text{ and } \text{NMAE}_s \leq \mu_3 \tag{37}$$

where μ_1 is a number close to but smaller than 1, e.g. 0.99, and μ_2 and μ_3 are two small positive percentages (e.g. 2% and 10%), respectively. The third termination criterion is the number of sampling points, which should not exceed the allowed maximum sampling size.

$$s > N_{\max} \tag{38}$$

The procedure of the sequential sampling will be stopped when both the inequalities in (35) and (37) are satisfied simultaneously, or the inequality (38) is satisfied.

When the sampling process is finished, the highest order of the HOPSM may not be the best order in demand, and the best order will be chosen by using the adjusted R-square. Therefore, after the sequential sampling is terminated, we will fit the HOPSM with respect to all the allowable orders, the coefficients to be estimated are less than the sampling points. The order of the HOPSM which has the largest value of adjusted R-square will be selected as the final order. Use $\overline{R}_{n,s}^2$ to denote the adjusted R-square of HOPSM with the order *n* and sampling size *s*. The best order of HOPSM will be determined by

$$n_{opt} = \left\{ n \mid \max_n \overline{R}_{n,s}^2, \kappa N_C(n,k) < s \right\} \tag{39}$$

where κ is a parameter between 1 and 2, e.g. 1.5 used in this paper. The final order of HOPSM will be n_{opt}, and the best HOPSM is given by $\hat{f}_{n_{opt}}(\mathbf{x})$ produced by *s* samples.

The flow chart of constructing HOPSM is shown in Fig. 6, which mainly includes five steps: candidates generation, initial sampling, exploration sampling, order increment, and order determination.

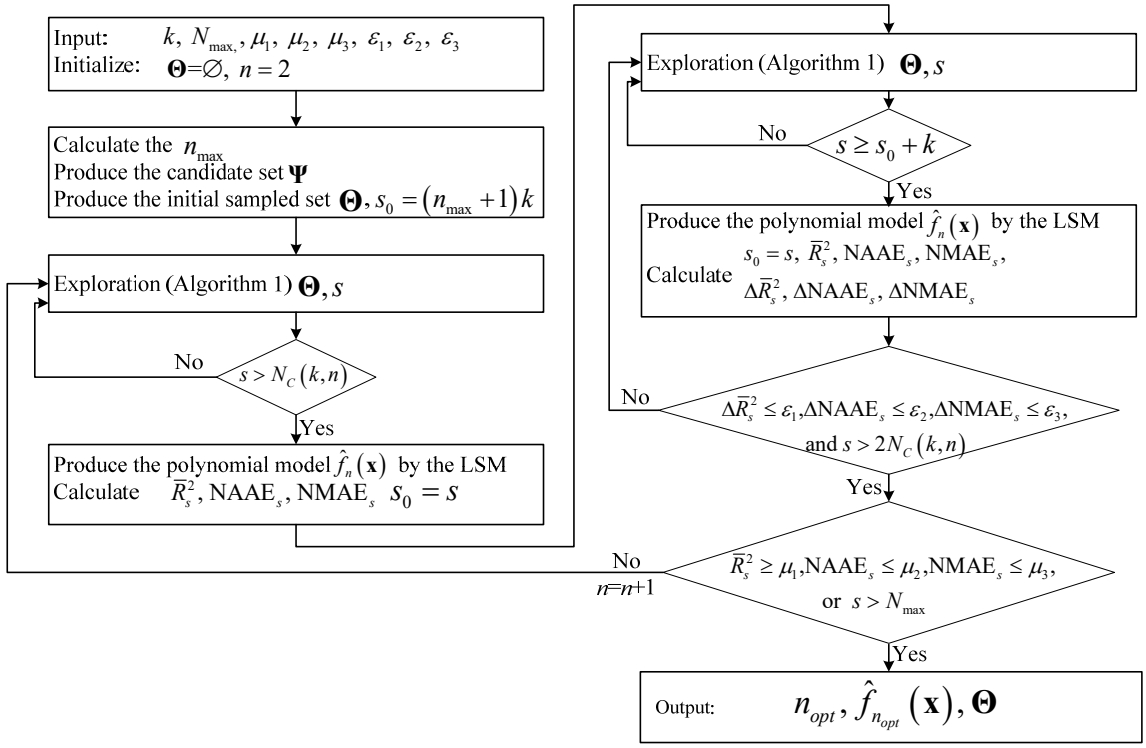


Figure.6 The flow chart of constructing HOPSM

The candidates $\mathbf{\Psi}$ are the tensor product of the zeros of Chebyshev polynomials, shown in Section 3.1. The initial sampling method has been given in Section 3.2, and it will produce a small uniform sampling set $\mathbf{\Theta}$. The algorithm of the exploration sampling is given in Section 3.3, which selects more samples from the candidates and keeps the data points distribute uniformly. The order incremental plan is used to update the order of HOPSM, and improve the approximation accuracy gradually. The exploration sampling and order incremental operation will be repeated until the termination criterions are satisfied. Lastly, the order determined by Eq. (39) may make HOPSM best.

4. Numerical examples

4.1 Mathematical test examples

Here we consider several benchmark mathematical testing examples, the expressions of which are given in Table 6. These mathematical examples contain strong nonlinear characteristics and different dimensional size, which are suitable for demonstrate the accuracy and robustness of the proposed surrogate model. The plot of the 2-dimensional functions, Michalewicz function and Deceptive function are given in Figs. 7 and 8, which show their nonlinear characteristics.

Table 6 The mathematic test functions

Functions	Expression	Domain	Dimension (k)
Michalewicz	$f(\mathbf{x})=\sum_{i=1}^{k}\sin(x_i)\left[\sin\left(\frac{ix_i^2}{\pi}\right)\right]^{20}$	$0 \leq x_i \leq \pi$	2

Deceptive	$f(\mathbf{x})=-\left(\frac{1}{k}\sum_{i=1}^{k}g_i(x_i)\right)^{0.5}$, $g_i(x_i)=\begin{cases}-x_i/\alpha_i+0.8 & 0\le x_i\le 0.8\alpha_i\\ 5x_i/\alpha_i-4 & 0.8\alpha_i\le x_i\le \alpha_i\\ 5(x_i-\alpha_i)/(\alpha_i-1)+1 & \alpha_i< x_i\le 0.2+0.8\alpha_i\\ (1-x_i)/(\alpha_i-1)+0.8 & 0.2+0.8\alpha_i< x_i\le 1\end{cases}$	$0\le x_i\le 1$	2
Paviani	$f(\mathbf{x})=\sum_{i=1}^{k}\left(\ln^2(x_i-2)+\ln^2(10-x_i)\right)-\left(\prod_{i=1}^{k}x_i\right)^2$	$2.001\le x_i\le 9.999$	3
Ackley	$f(\mathbf{x})=-20\exp\left(-0.2\sqrt{\frac{1}{k}\sum_{i=1}^{k}x_i^2}\right)-\exp\left(\frac{1}{k}\sum_{i=1}^{k}\cos 2\pi x_i\right)+20+\exp(1)$	$-2\le x_i\le 2$	4
Schwefel	$f(\mathbf{x})=\sum_{i=1}^{k}\left(-x_i\sin\left(\sqrt{\lvert x_i\rvert}\right)\right)$	$-100\le x_i\le 100$	5
Rastrigin	$f(\mathbf{x})=10k+\sum_{i=1}^{k}\left(x_i^2-10\cos(2\pi x_i)\right)$	$-1\le x_i\le 1$	6

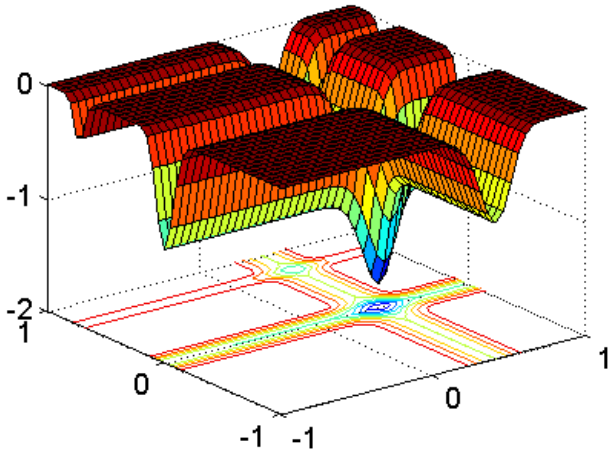


Figure. 7 The plot of Michalewicz function

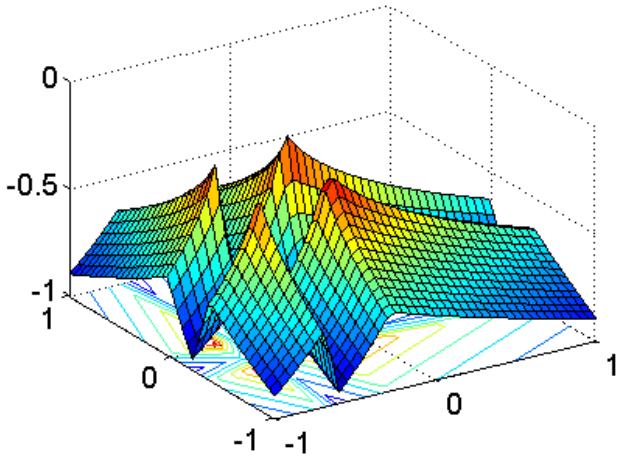


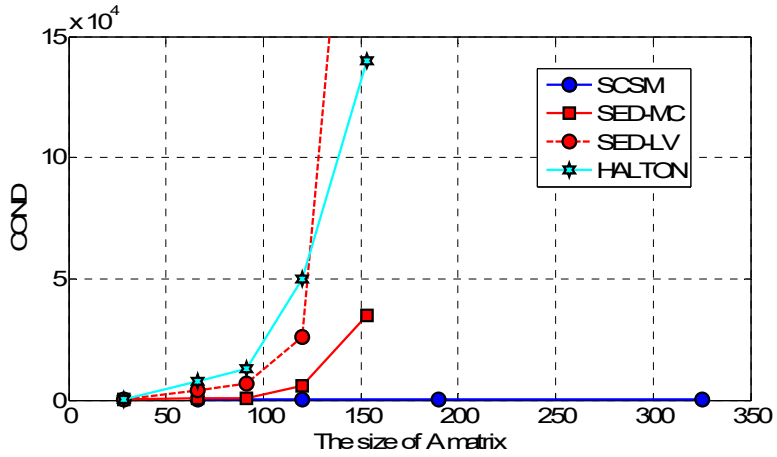
Figure. 8 The plot of Deceptive function

The most allowable sampling size Nmax is set as 500 for 2-dimensional functions, and 1000 for higher dimensional functions. The condition number is used to illustrate the stability of the SCSM, while the R2 and NMAE shown in Section 2.3 are used to validate the approximation accuracy of the HOPSM. Besides the proposed SCSM, the Halton sequential method, LOLA-Voronoi algorithm, and the global Monte Carlo method (Crombecq et al., 2011b) will be employed as the referees. The LOLA-Voronoi algorithm and global Monte Carlo method are obtained from the SED toolbox in Matlab, so they are noted as SED-LV and SED-MC, respectively. To keep the uniformity of test points, we use the Hammersely (Kalagnanm and Diwekar, 1997) sequence to produce 10000 test points.

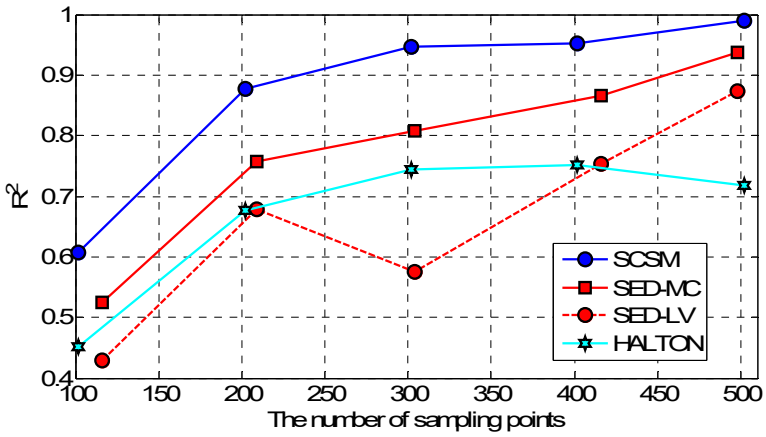
All the sampling methods are implemented on the workstation with Intel core running at 3.4 GHz. When 1000 sampling points are selected, the Halton sequential sampling only takes less than 1 min, and most of the cost is spent on the surrogate modelling (LSM). The computation time of SCSM is about 2 mins, and the third is the SED-LV, which takes about 20 mins to produce 1000 samples. The last one is the SED-MC, which takes about 2 hours to produce 1000 sampling points. Therefore, the proposed SCSM is very high efficient, just slightly lower than the Halton sequential sampling method.

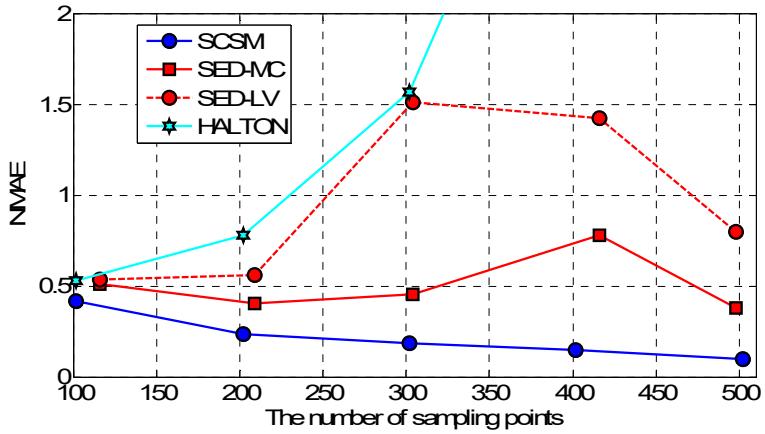
For the Michalewicz function shown in Fig. 9(a), the condition number of SCSM is much smaller than that of other methods. When the size of **A** matrix exceeds 100, the COND of the Halton, SED-LV, and SED-MC increases rapidly, but the COND of SCSM always keeps small even when the size of the **A** matrix exceeds 300, which means there are more possible to produce large numerical errors when using the LSM for the three reference methods. Therefore, the SCSM is much more stable than other methods for the Michalewicz function. For the accuracy, it can be found that the R-square (Fig. 9(b)) of SCSM is larger than other sampling method while the NMAE (Fig. 9(b)) of SCSM is the smallest, so the SCSM has the highest accuracy in both global error and local error.

(a)



(b)

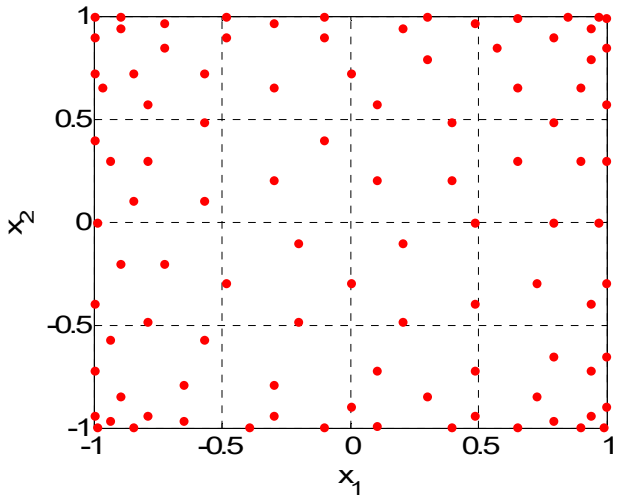




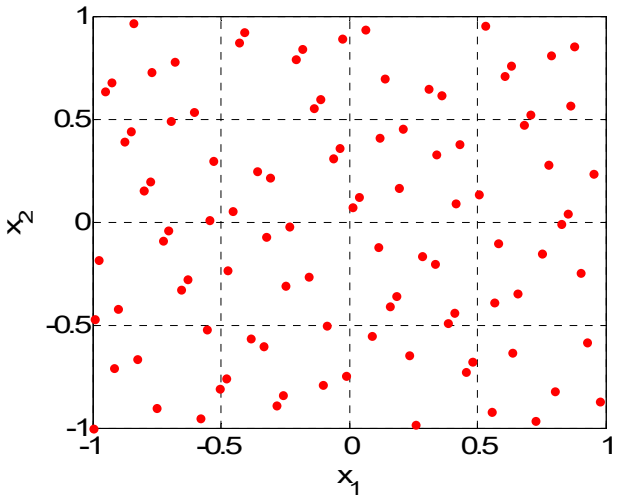
(c)

Figure.9 The COND, R2, and NAME for the 2d Michalewicz function

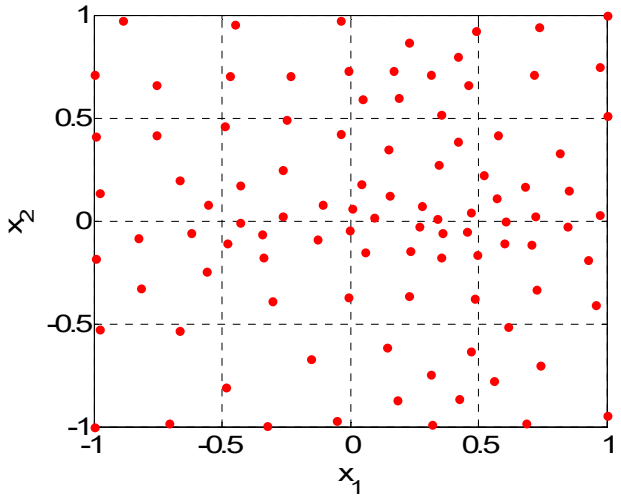
To show the distribution of the four types of samples directly, Fig.10 plots the different sampling points when the number of samples is set as 100.

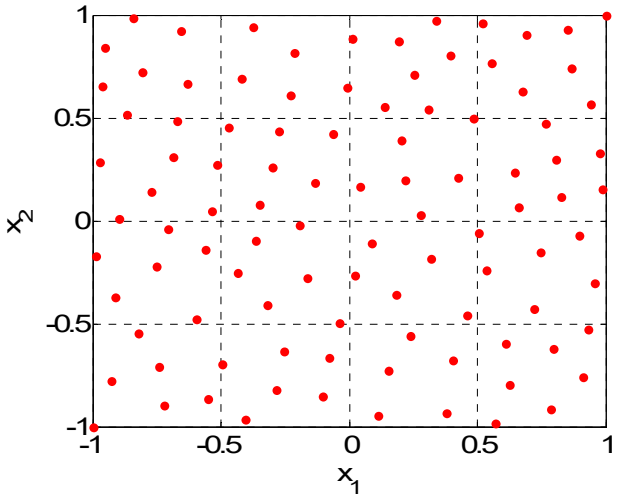


(a) SCSM



(b) HALTON





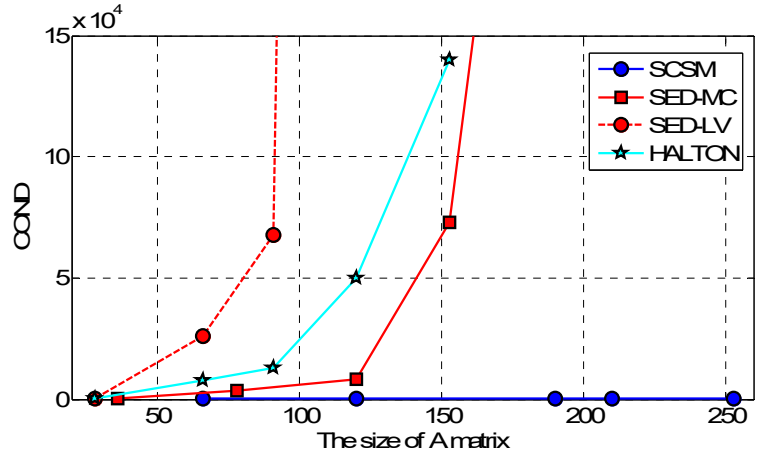
(c) SED-LV (d) SED-MC

Figure. 10 The distribution of 100 samples

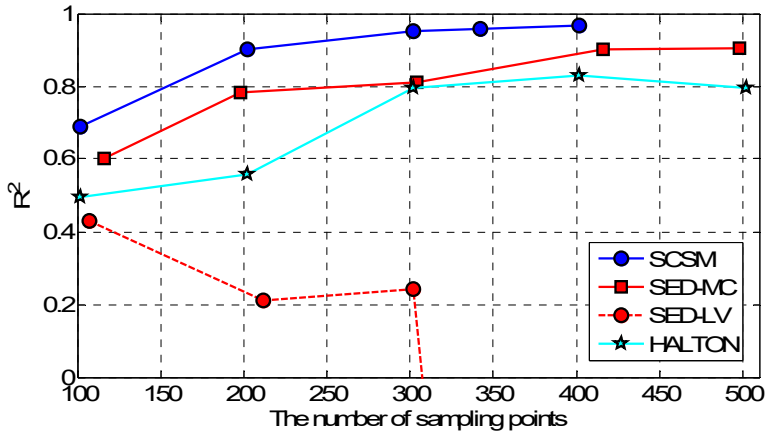
From Fig. 10, it can be found that the samples obtained by the SCSM and SED-MC are distributed more evenly than the other two sampling methods. The SED-MC seems to produce the most uniform distributed samples in the entire design space, while the SCSM makes more sampling points locate in region close to the edges of the design space. Combining the results shown in Fig. 9, we conclude that the more uniformly distributed samples does not have to produce higher approximation accuracy.

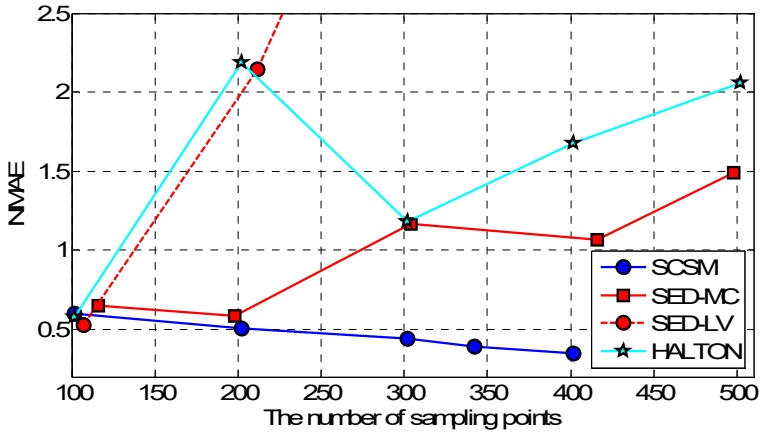
Figure. 11 shows the condition number and approximation errors of the Deceptive function. The results are consistent with the results of Michalewicz function. Thus, the proposed SCSM has advantages in both stability and approximation accuracy for the 2-dimensional functions.

(a)



(b)

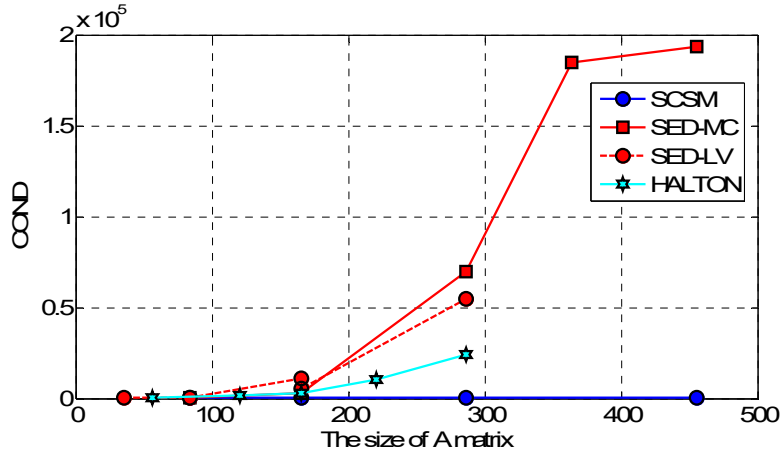




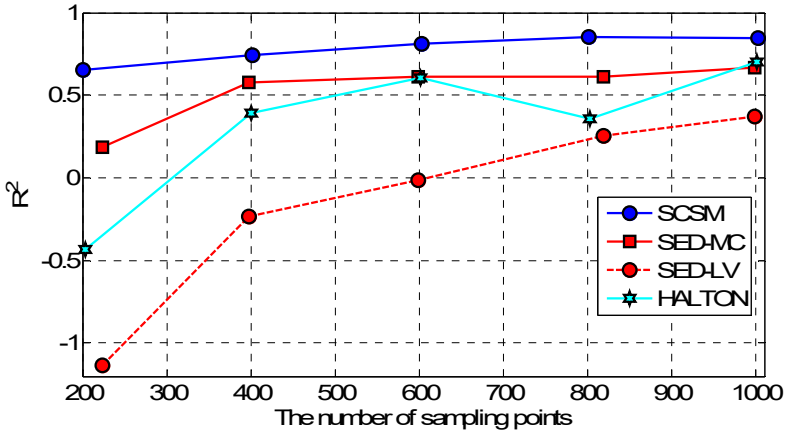
(c)

Figure.11 The COND, R2, and NAME for the 2d Deceptive function

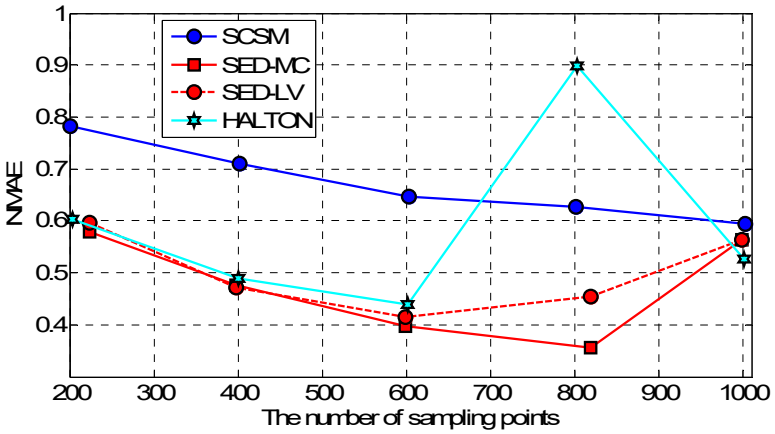
The COND, R^2, and NAME of the 3-dimensional Paviani function are shown in Fig. 12. The condition number of SCSM is much smaller than other three methods, no matter what the size of the **A** matrix is. The R^2 of SCSM is the largest, so the SCSM has higher global approximate accuracy for this test function. The NMAE of SCSM is worse than the SED-LV and SED-MC, the main reason of which is the extremum of the Paviani function occur at the edges of design space. The SCSM cannot sample the points on the edges, but the SED-LV and SED-MC will selected the points on the edges.



(a)



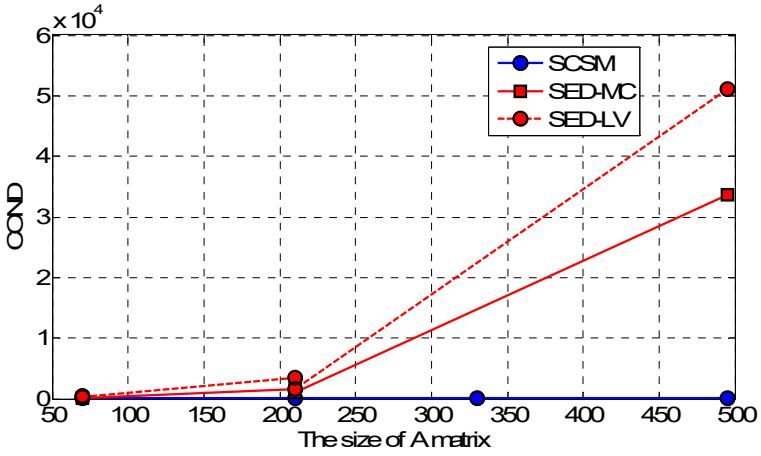
(b)



(c)

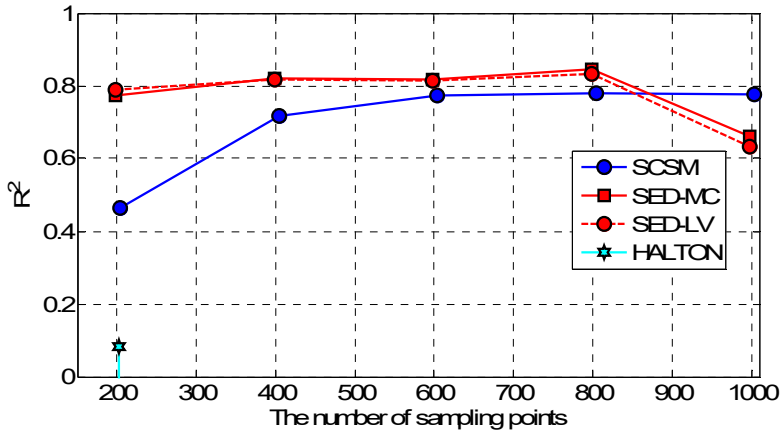
Figure.12 The COND, R^2, and NAME for the 3d Paviani function

Figure. 13 gives the results of the 4-dimensional Ackley function. The COND of the Halton method is not shown in the figure, because its value is too large. Similar to the previous three test functions, the SCSM has the best stability for the Ackley function, since its condition number is much smaller than the SED-LV and SED-MC. When the sampling size is less than 400, the SED-LV and SED-MC have better accuracy than the SCSM. With the sampling size increasing, the three methods provide similar approximation accuracy. However, the SED sampling methods become worse when the sampling size is 1000. From another point of view, it also indicates that the SED sampling methods may be unstable for constructing the HOPSM.



(a)

(b)



(c)

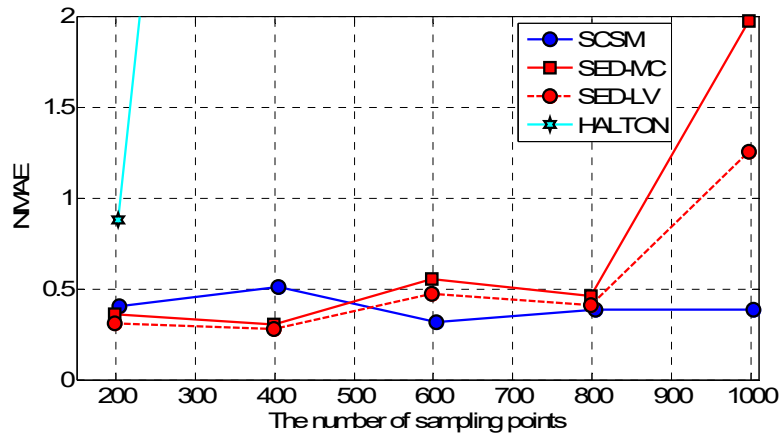
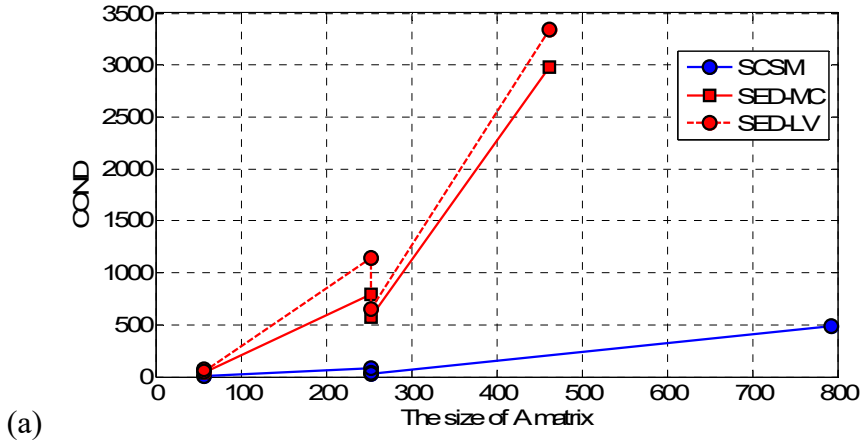
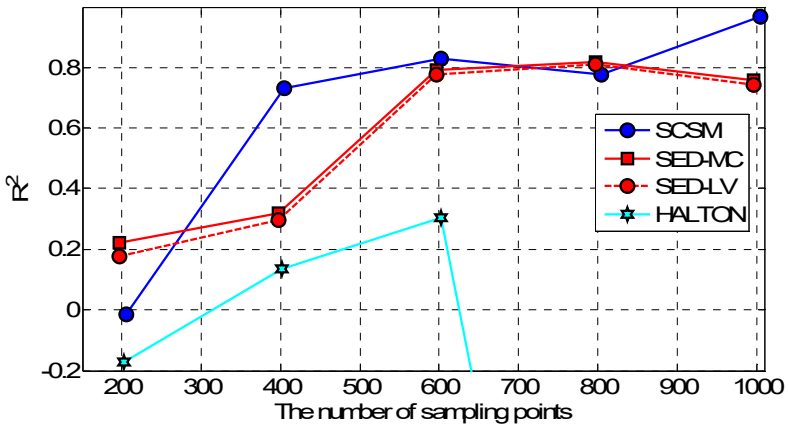


Figure.13 The COND, R2, and NAME for the 4d Ackley function

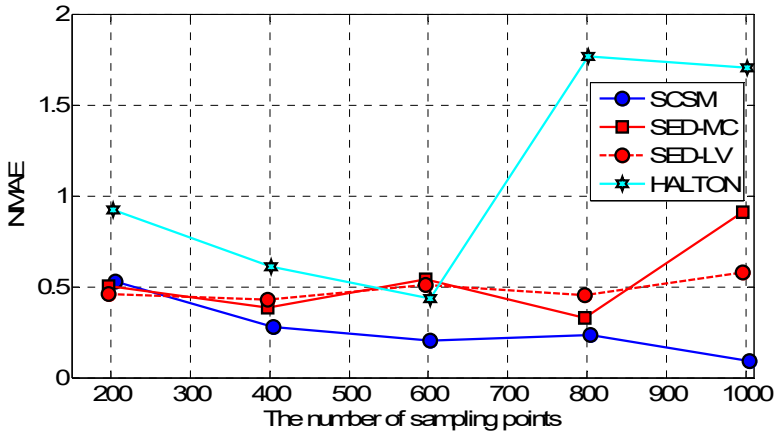
The results of 5-dimensional Schwefel function are shown in Fig. 14. It can be found that the condition number of SED-LV and SED-MC have been more than 500 when the size of the **A** matrix is around 250, while the COND of SCSM is only about 100 at the same size of the **A** matrix. The COND of SED-LV and SED-MC exceeds 3000 when the size of **A** is 450, but the condition number of SCCM is only 500 even when the size of **A** is 800. Hence, the SCSM is more stable than other sampling methods for the 5-dimensional Schwefel function. Figure. 11(b) and (c) indicate that the R^2 of SCSM is higher or close to the two SED methods under different sampling size, and the NMAE of SCSM is also smaller than other three sampling methods.



(a)



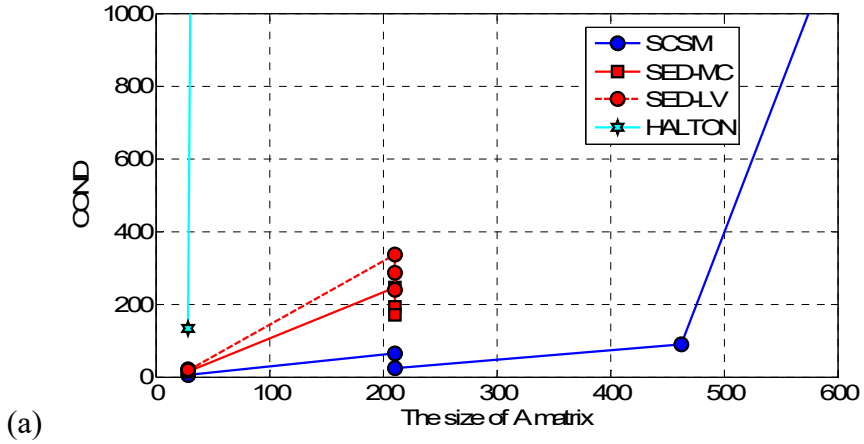
(b)



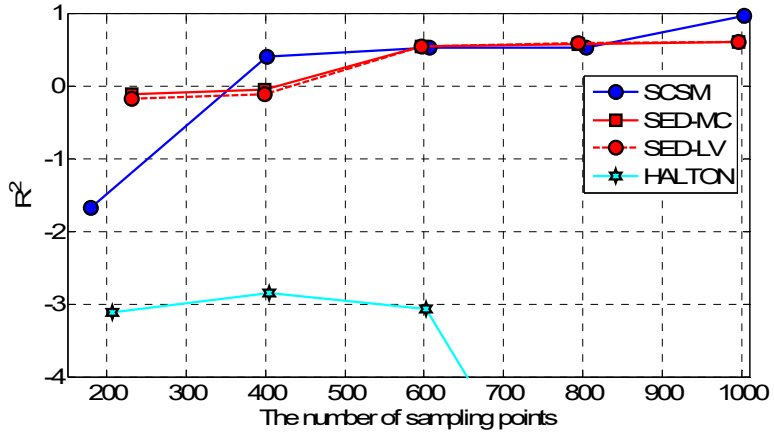
(c)

Figure.14 The COND, R2, and NAME for the 5d Schwefel function

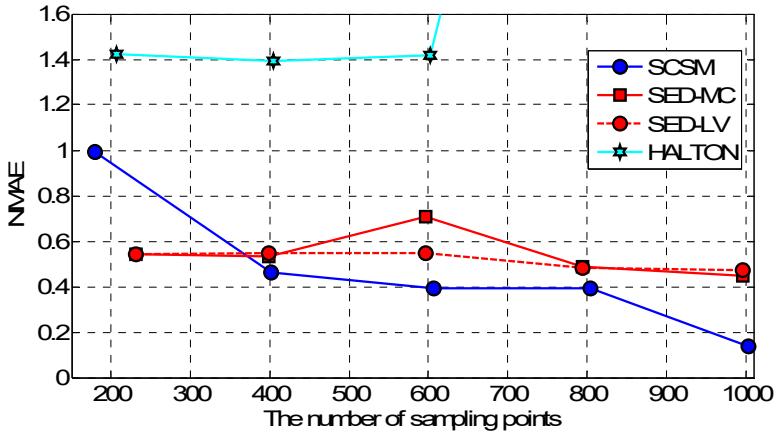
The results of the last test function, 6-dimensional Rastrigin function, are shown in Fig. 15. For the same size of the **A** matrix, the COND of SCSM is much smaller than that of other three methods, e.g. when the size of **A** is 200, the COND of SED-LV and SED-MC is about 300 and 200, respectively, while that of the SCSM is less than 100. For the approximation accuracy, except the case that the sampling size is small (e.g. 200), the SCSM has higher accuracy than the three references for all other cases.



(a)



(b)



(c)

Figure.15 The COND, R2, and NAME for the 6d Rastrigin function

From these mathematic test functions, we will have the following conclusions. The Halton sequential sampling method has the highest efficiency, while the proposed SCSM also has high efficiency which is quite close to the Halton sampling method, followed by the SED-LV, and the SCD-MC is the lowest efficient method. For the stability, the SCSM shows the outstanding performance than other three methods, followed with the two SED sampling methods, and the Halton method is extremely unstable. For the approximation accuracy, the SCSM has higher accuracy in most of the cases, while the two SED sampling methods only have better accuracy in some cases, e.g. the 5d Schwefel function and 6d

Rastrigin function in the case of fewer than 200 sampling points. The Halton sampling method almost gives the lowest approximation accuracy.

4.2 Engineering Applications

Application 1: slider crank mechanism

The slider crank is a very important mechanism in mechanical engineering, because it has the capability to transform the motion mode between the translation and rotation. One of the most important applications of slider crank is the engine in vehicle, which transforms the translation of slider to the rotation of crank and then output the rotation energy to the transmission system of vehicles. The schematic of slider crank mechanism is given in Fig. 16. Points A, B, and C are the gravity centers of the crank, connecting rod, and slider, respectively. θ_1 and θ_2 show the angles between the global coordinate and local coordinate of the crank and connecting rod, respectively. L1 and L2 are the lengths of the crank and connecting rod. m1, m2, and m3 represent the mass of the crank, connecting rod and slider, respectively. ω is the angular velocity of the crank, and τ is the external torque applied to the crank. F_{Xi} and F_{Yi} (i=1, 2, 3) are the reaction forces in X direction and Y direction acted on the three hinges.

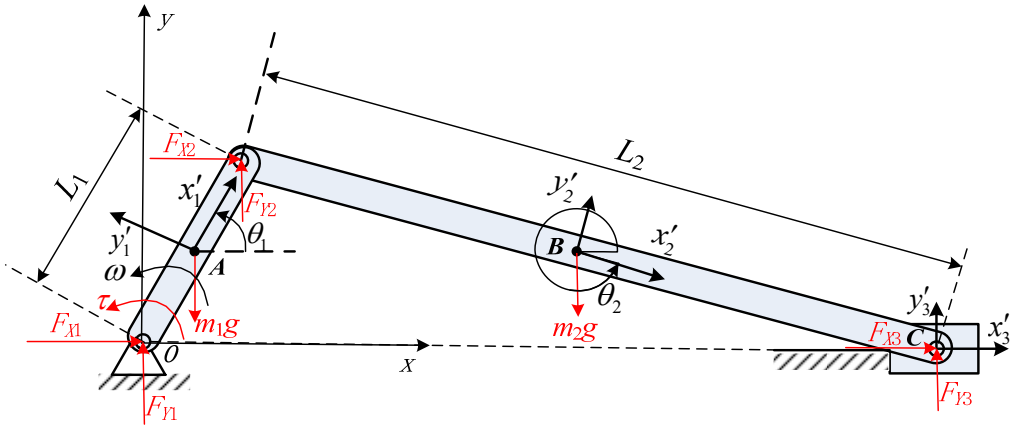


Figure. 16 The schematic of slider crank mechanism

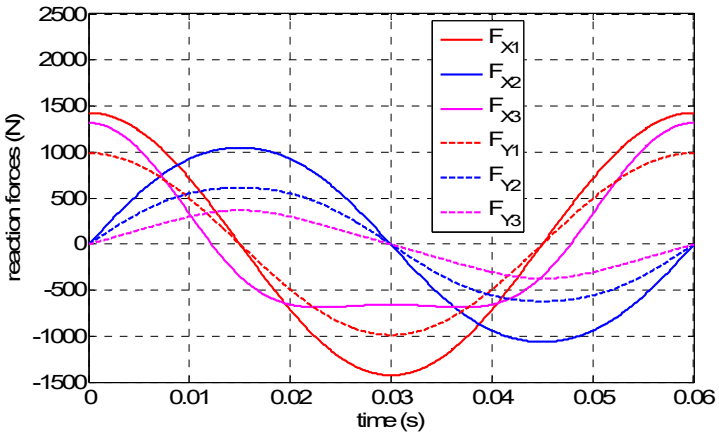


Figure. 17 The reaction forces of slider crank in one period

The crank and connecting rod are assumed to be uniform bar, and their density is 2 kg/m and 1.5 kg/m, respectively, while the mass of slider is set as 0.45 kg. When the angular velocity of crank (ω) is given,

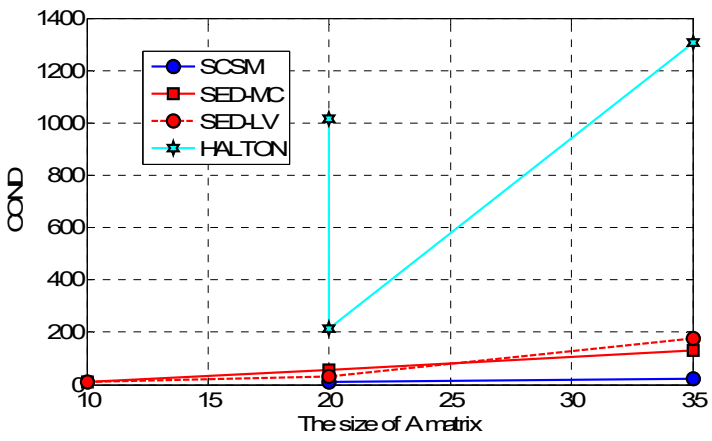
the external torque τ and other reaction forces on each hinge can be obtained by using the theory of rigid body dynamics (Wu et al., 2013a). If the ω is set as 1000 rpm, the reaction forces in one period are shown in Fig. 17.

It can be found that the F_{XI} gets the maximum value 1425 N at the time 0.03 s and 0.06 s. The maximum reaction force is an important factor to design a new mechanism, so we will use the proposed method to build the surrogate model of the maximum reaction force. In this example, three design variables are considered, i.e. the length of crank L1 and connecting rod L2, and the angular velocity ω. The range of the three design variables are defined as $0.2m \le L1 \le 0.4m$, $0.6m \le L2 \le 1.2m$, and $1000rpm \le \omega \le 2000rpm$.

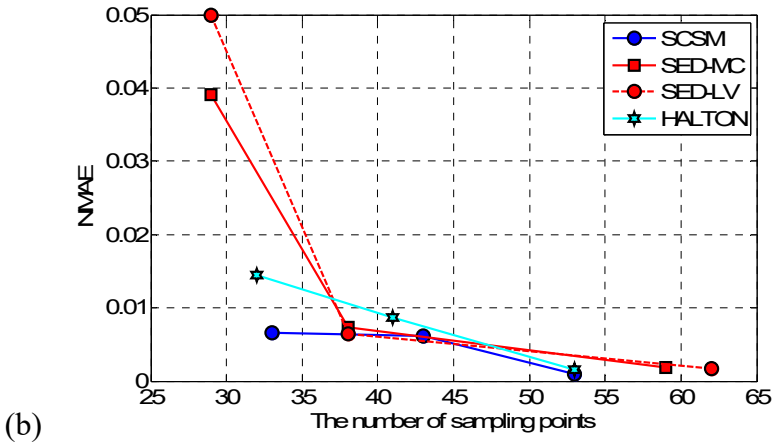
Using the four sampling methods shown in section 4.1, the results of these methods are given in Fig. 18. To distinguish the average error of these methods, the RMSE is used to replace the R2 which is very close for all these methods in this example. The condition number of the SCSM is the smaller than other three methods under the same size of the A matrix, while the SED-MC and SED-LV are better than the HALTON sampling. Therefore, the SCSM is the most stable sampling method for this example.

We use the RMSE to show the average error of the surrogate model. It may be found that when the sampling size is 29, the RMSE of SED-MC and SED-LV is much worse, because the order of the HOPSM is only 2 in this case. When the order of HOPSM exceeds 3, the RMSE of the four sampling methods is quite close. For the maximum error, the NMAE of SCSM has the highest convergence ratio, followed with the Halton sampling, while the SED-MC and SED-LV converge slowly.

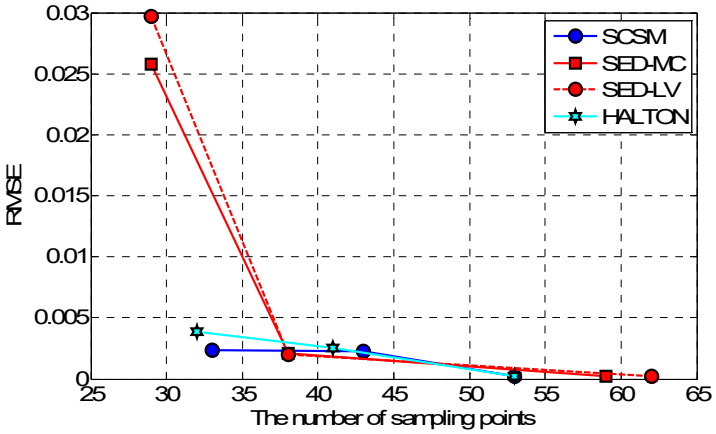
In summary, the SCSM has the best performance in stability and local accuracy for the slider crank mechanism. The global accuracy (average error) is very high and quite similar for all the four sampling methods



(a)



(b)



(c)

Figure. 18 The COND, RMSE, and NAME for the maximum reaction force

Application 2: Vehicle handling model

In this section we use the HOPSM to study the vehicle handling performance. The 27-DOF vehicle model in software CarSim is used to analysis the vehicle handling performance, shown as Fig. 19. The work condition double-loop control is used to analysis the yaw velocity of the vehicle, which is an important evaluation index of handling performance. Six design variables are considered to build the HOPSM, including the spring mass M_SU, the front unspring mass M_US1, the rear unspring mass M_US2, the wheelbase LX_AXLE, the distance from front axle to the gravity centre of spring mass LX_CG_SU, and the height of gravity centre of spring mass H_CG_SU. The variation range of the design variables are given in Table 7.

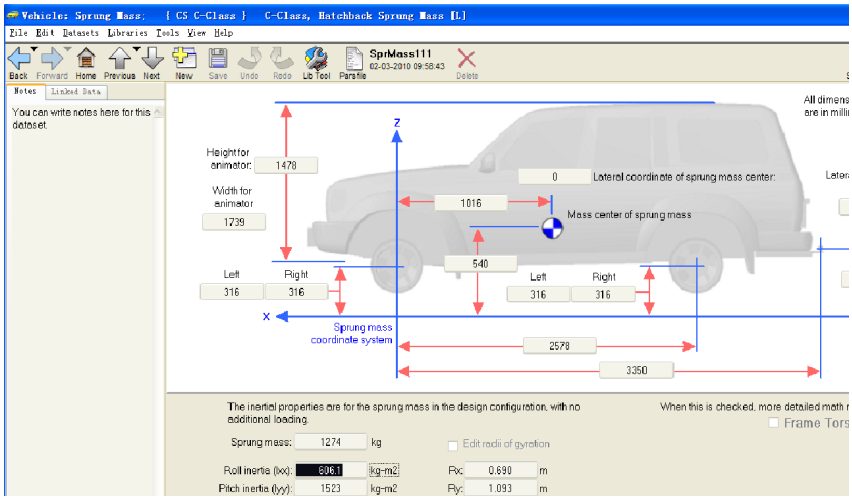


Figure. 19 The 27-DOF vehicle model in CarSim

Table 7 The range of design variables

Parameters	M_SU	M_US_1	M_US_2	LX_AXLE	LX_CG_SU	H_CG_SU
Range	[637,1901]*kg*	[30,101]*kg*	[30,101]*kg*	[2062,3094]*mm*	[813,1219]*mm*	[432,648]*mm*

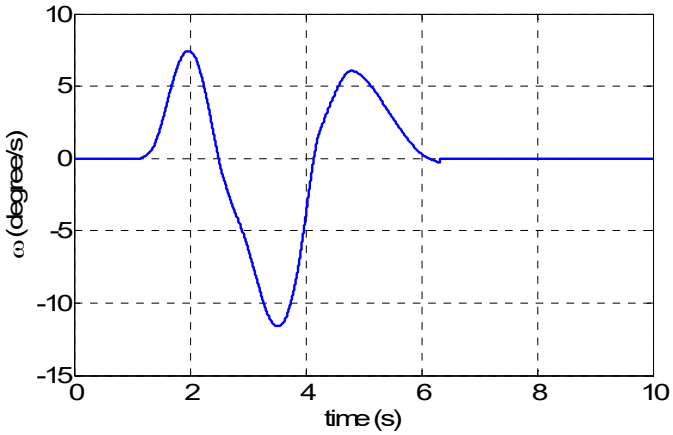


Figure. 20 The yaw velocity of vehicle under the double-loop control

Considering the case that the vehicle velocity is 120km/h and all design variables are set as the middle values of their variation range, the yaw velocity with respect to time is provided in Fig. 20. The smaller yaw velocity, the better handling performance of the vehicle is. The maximum yaw velocity (absolute value) $|\omega|_{\max}$ is used to indicate the vehicle handling performance, so we will build the HOPSM of the maximum yaw velocity based on the 6 design variables given in Table 7. The vehicle model is run in the software CarSim, and each running time is about 5 seconds. To limit the computational cost for building the surrogate model, the most allowable sampling size $N_{\max}$ is set as 1000.

Using the same four types of sampling methods shown in section 4.1 to build the HOPSM, the results about the stability and accuracy of these methods are summarised in Table 8, where the 10000 test data are still produced by the Hamerseley points.

Table 8 The stability and accuracy of HOPSM for the max yaw velocity

Sampling method	SCSM	SED-MC	SED-LV	HALTON
COND	46.54	1.13e3	1.54e3	1.12e6
R^2	0.86	0.83	0.82	-4.48
NMAE	0.24	0.45	0.44	4.10

For the condition number, it can be found that the proposed SCSM has the minimum value 46.54, while the condition number of other methods is much larger than it. Therefore, the SCSM is more stable than the SED sampling methods and Halton sampling method. For the global accuracy, the R2 of the SCSM is the highest 0.86, followed with the SED-MC (0.83) and SED-LV (0.82). The R2 of Halton sampling is quite small even be a negative number, which means that the approximation accuracy of this surrogate model is extremely bad. For the local accuracy, the NMAE gets the minimum value 0.24 for the SCSM, which is almost half of the SED-MC (0.45) and SED-LV (0.44), so the SCSM also has the best local accuracy. The NMAE of Halton sampling is quite large, which indicates that the Halton sampling is unsuitable for resolving this vehicle model.

5. Conclusions

A new sequential sampling method termed as sequential Chebyshev sampling method (SCSM) is proposed in this paper to build the high-order polynomials surrogate model (HOPSM). The SCSM uses the zeros of Chebyshev polynomials as the sampling candidates, and then sequentially choose the sampling points based on the space-filling principle. This paper contains three novelties: (1) an efficient uniform initial sampling technique, termed as coordinate alternation algorithm, is proposed; (2) a fast sequential space-filling sampling algorithm is proposed; (3) the high-order polynomials are used to build the surrogate model to improve the accuracy of traditional polynomials model. Two sampling methods from the sequential sampling toolbox SED in MATLAB and the Halton sampling method are used as the references of the proposed SCSM. Several mathematical examples and two engineering applications are used to demonstrate the performance of the proposed sequential sampling method. The results show that the SCSM has the best stability compared to other three sequential sampling methods. The accuracy of the SCSM is also higher than the SED-MC, SED-LV, and Halton sampling in most cases. The efficiency of SCSM is close to the Halton sampling and much higher than the SED sampling methods.

Reference

http://sumo.intec.ugent.be/SED. In (ed.), Vol. pp.

Acar E (2010), "Various approaches for constructing an ensemble of metamodels using local measures", *Structural and Multidisciplinary Optimization*, Vol. 42, pp. 879-96.

Ajdari A and Mahlooji H (2014), "An Adaptive Exploration-Exploitation Algorithm for Constructing Metamodels in Random Simulation Using a Novel Sequential Experimental Design", *Communications in Statistics - Simulation and Computation*, Vol. 43, pp. 947-68.

Barton RR (1994) Metamodeling: a state of the art review. In *the 1994 Winter Simulation Conference*. (ed.), Vol. pp. Lake Buena Vistak, Florida.

Barton RR (1992) Metamodels for simulation input-output relations. In *the 24th conferencee on Winter simulation*. (ed.), Vol. pp. 289-99, New York.

Barton RR (1998) Simulation metamodels. In *the 1998 Winter Simulation Conference*. (ed.), Vol. pp. Washington D. C.

Busby D, Farmer CL, and Iske A (2007), "Hierarchical Nonlinear Approximation for Experimental Design and Statistical Data Fitting", *SIAM Journal on Scientific Computing*, Vol. 29, pp. 49-69.

Chen R-B, Hsieh D-N, Hung Y, and Wang W (2012), "Optimizing Latin hypercube designs by particle swarm", *Statistics and Computing*, Vol. 23, pp. 663-76.

Clarke SM, Griebsch JH, and Simpson TW (2005), "Analysis of Support Vector Regression for Approximation of Complex Engineering Analyses", *Journal of Mechanical Design*, Vol. 127, pp. 1077.

Crombecq K, De Tommasi L, Gorissen D, and Dhaene T (2009) A novel sequential design strategy for global surrogate modeling. In *Proceeding in 2009 Winter Simulation Conference*. (ed.), Vol. pp. 731-42, Austin, TX, USA.

Crombecq K, Gorissen D, Deschrijver D, and Dhaene T (2011a), "A Novel Hybrid Sequential Design Strategy for Global Surrogate Modeling of Computer Experiments", *SIAM Journal on Scientific Computing*, Vol. 33, pp. 1948-74.

Crombecq K, Laermans E, and Dhaene T (2011b), "Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling", *European Journal of Operational Research*, Vol. 214, pp. 683-96.

Fang KT, Lin DKJ, Winker P, and Zhang Y (2000), "Uniform Design: Theory and Application", *Technometrics*, Vol. 39, pp. 237-48.

Fox L and Parker IB (1968), *Chebyshev Polynomials in Numerical Analysis*, Oxford University Press, London.

Friedman JH (1991), "Multivariate adaptive regression splines", *Annals of Statistics* Vol. 19, pp. 1-141.

Gallina A (2009) Response surface methodology as a tool for analysis of uncertainty in structural dynamics[Thesis]. Type, AGH - University of Science and Technology,

Gil A, Segura J, and Temme NM (2007), *Numerical Methods for Special Functions*, SIAM,

Giunta AA, Wojtkiewicz JSF, and Eldred MS (2003) Overview of Modern Design of Experiments Methods for Computational Simulations. In *41st AIAA Aerospace Sciences Meeting and Exhibit*. (ed.), Vol. pp. AIAA Paper 2003-0649, Reno, NV.

Goel T, Haftka RT, Shyy W, and Queipo NV (2006), "Ensemble of surrogates", *Structural and Multidisciplinary Optimization*, Vol. 33, pp. 199-216.

Gorissen D, Couckuyt I, Demeester P, Dhaene T, and Crombecq K (2010), "A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design", *Journal of Machine Learning Research*, Vol. 11, pp. 2051-55.

Hajikolaei KH and Wang GG (2012) Adaptive orthonormal basis functions for high dimensional metamodeling with existing sample points. In *ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (ed.), Vol. pp. 709-16, Chicago, IL, USA.

Jin R, Chen W, and Simpson TW (2001), "Comparative studies of metamodelling techniques under multiple modelling criteria", *Structural and Multidisciplinary Optimization*, Vol. 23, pp. 1-13.

Jonson ME, Moore LM, and Yalvisaker D (1990), "Minimax and maximin distance designs", *Journal of Statistical Planning and Inferences*, Vol. 26, pp. 131-48.

Kalagnanm JR and Diwekar UM (1997), "An Efficient Sampling Technique for off_line quality control", *American Statistical Association and the American Society for Quality Control Technometrics*, Vol. 39, pp.
Lin JG (2006), "Modeling Test Responses by Multivariable Polynomials of Higher Degrees", *SIAM Journal on Scientific Computing*, Vol. 28, pp. 832-67.
Marrel A, Iooss B, Veiga S, and Ribatet M (2011), "Global sensitivity analysis of stochastic computer models with joint metamodels", *Statistics and Computing*, Vol. 22, pp. 833-47.
Mckay MD, Beckkman RJ, and Conover WJ (1979), "A comparison of three methods for selceting values of input variables in the analysis of output from a computer code", *Technometrics*, Vol. 21, pp. 266-94.
Montgomery DC (2007), *Design and Analysis of Experiments*, 6th (ed.), Post & Telecom Press, Beijing.
Morris MD and Mitchell TJ (1995), "Exploratory design for computational experiments", *Journal of Statistical Planning and Inferences*, Vol. 43, pp. 381-402.
Myers RH and Montgomery D (1995), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley and Sons, Inc., Toronto.
Owen AB (1992), "Orthogonal arrays for computer experiments, integration and visualiztion", *Statistica Sinca*, Vol. 2, pp. 439-52.
Pronzato L and Müller WG (2011), "Design of computer experiments: space filling and beyond", *Statistics and Computing*, Vol. 22, pp. 681-701.
Qu X, Venter G, and Haftka RT (2004), "New formulation of a minimum-bias central composite experimental design and Gauss quadrature", *Structural and Multidisciplinary Optimization*, Vol. 28, pp. 231-42.
Romero VJ, Slepoy R, Swiler LP, and Giunta AA (2005) Error estimation approaches for progressive response surfaces - more results. In *46th AIAA/ASME/ASCE/AHS/ASC Structural Dynamics and Materials Conference*. (ed.), Vol. pp. 269-89, Austin, TX.
Romero VJ, Swiler LP, and Giunta AA (2004), "Construction of response surfaces based on progressive-lattice-sampling experimental designs with application to uncertainty propagation", *Structural Safety*, Vol. 26, pp. 201-19.
Rutherford BM, Swiler LP, Paez TL, and Urbina A (2005) Response Surface (Meta-model) Methods and Applications. In (ed.), Vol. SAND2005-7933C, pp. SANDIA Nat. Lab.,
Shan S and Wang GG (2010), "Metamodeling for high dimensional simulation-based design problems", *Journal of Mechanical Design*, Vol. 132, pp. 051009:1-11.
Simpson TW, Lin DKJ, and Chen W (2001a), "Sampling Strategies for Computer Experiments Design and Analysis", *International Journal of Reliability and Applications*, Vol. 2, pp. 209-40.
Simpson TW, Mauery TM, Korte JJ, and Mistree F (2001b), "Kriging models for global approximation in simulation-based multidisciplinary design optimization", *AIAA Journal*, Vol. 39, pp. 2233-41.
Viana FAC, Venter G, and Balabanov V (2009), "An algorithm for fast optimal Latin hypercube design of experiments", *International Journal for Numerical Methods in Engineering*, Vol. 82, pp. 135-56.
Wang GG and Shan S (2007), "Review of metamodeling techniques in support of engineering design optimization", *Journal of Mechanical Design*, Vol. 129, pp. 370-80.
Wong T-T, Luk W-S, and Heng P-A (1997), "Sampling with Hammersley and Halton Points", *Journal of Graphics Tools*, Vol. 2, pp. 9-24.
Wu J, Luo Z, Zhang N, and Zhang Y (2014a), "A new sampling scheme for developing metamodels with the zeros of Chebyshev polynomials", *Engineering Optimization*, Vol. pp. 1-25.
Wu J, Luo Z, Zhang Y, and Zhang N (2014b), "An interval uncertain optimization method for vehicle suspensions using Chebyshev metamodels", *Applied Mathematical Modelling*, Vol. pp.
Wu J, Luo Z, Zhang Y, Zhang N, and Chen L (2013a), "Interval uncertain method for multibody mechanical systems using Chebyshev inclusion functions", *International Journal for Numerical Methods in Engineering*, Vol. 95, pp. 608-30.
Wu J, Zhang Y, Chen L, and Luo Z (2013b), "A Chebyshev interval method for nonlinear dynamic systems under uncertainty", *Applied Mathematical Modelling*, Vol. 37, pp. 4578-91.