

## Research Article

# Kriging-Based Space Exploration Global Optimization Method in Aerodynamic Design

Wei Zhang,<sup>1</sup> Zhenghong Gao ,<sup>1</sup> Chao Wang,<sup>2</sup> and Lu Xia <sup>1</sup>

<sup>1</sup>Northwestern Polytechnical University, Xi'an 710072, China

<sup>2</sup>Beijing Institute of Electronic System Engineer, Beijing 100000, China

Correspondence should be addressed to Lu Xia; [xialu@nwpu.edu.cn](mailto:xialu@nwpu.edu.cn)

Received 10 September 2022; Revised 17 December 2022; Accepted 22 December 2022; Published 2 February 2023

Academic Editor: Fangzhou Fu

Copyright © 2023 Wei Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For complicated aerodynamic design problems, the efficient global optimization method suffered from the defect of the incorrect portrayal of the design space, resulting in bad global convergence and efficiency performance. To this end, a Kriging-based global optimization method, named the Kriging-based space exploration method (KSE), was proposed in this paper. It selected multiple promising local minima and classified them into partially and fully explored minima in terms of the fitting quality of the surrogate model. Then, the partially explored minima would be furtherly exploited. During the local search, an enhanced trust-region method was adopted to make deep exploitation. By combining local and global searches, the proposed method could improve the fitting quality of the surrogate model and the optimization efficiency. The KSE was compared to other global surrogate-based optimization methods on 12 bound-constrained testing functions with 2 to 16 design variables and 2 aerodynamic optimization problems with 24 to 77 design variables. The results indicated that the KSE generally took fewer function evaluations to find the global optima or reach the target value in most test problems, holding better efficiency and robustness.

## 1. Introduction

In aerodynamic shape design, Computational fluid dynamics (CFD) solvers and optimization algorithms were essential. The optimization algorithms have been divided into two categories: gradient-free and gradient-based methods. Generally, the gradient-based methods needed sufficient gradient calculations and inherently converged to local minima, which was not suitable for complex multimodal problems [1]. The gradient-free methods, which did not need gradient information, could be easily incorporated into existing frameworks and have been found in many engineering applications [2, 3]. The gradient-free methods could be divided into several different groups: biology-based, physics-based, social-based, music-based, chemical-based, sport-based, swarm-based, and hybrid methods which were combinations of these [4, 5]. Since the optimization field was expansive, suitable gradient-free methods would be problem dependent. However, the gradient-free methods would need massive calls and simulations for the high-fidelity design requirements, leading to an enormous computational burden. Therefore, the efficient global

optimization method (EGO) [6] has greatly acquired the research and development of the global optimization method called surrogate-based optimization (SBO) [7–12], which had the potential to uncover the unconventional aircraft configurations in the aerodynamic design. For problems with multimodality, the accuracy for the surrogate model largely depended on the capability of the surrogate model in capturing the nonlinearity of the problems. Ignoring this would result in suboptimality or infeasibility of optimal solutions [13–16]. The EGO method soon acquired a reputation for its flexibility to imitate complicated functions for it had a random error estimate between the true and the surrogate functions to quantitatively measure the uncertainty in prediction [3, 17].

The most attractive characteristic of the EGO method was that it could meet two requirements: the exploitation of local minima (local search ability) and the exploration of uncertain areas (global search ability). The balance of global and local search ability largely relied on the predicted standard error. However, the numerical experiments showed that the EGO did not perform very well on multimodal problems and it also suffered from the curse of

dimensionality on problems with more than 30 design variables [18–20]. Thus, it would take a considerable amount of computational cost to improve the surrogate model accuracy [18] for complicated aerodynamic design problems with high multimodality and dimensionality [21]. Therefore, enhancing and balancing these two abilities was essential to improve optimization efficiency [22]. Various extensions and modifications of EGO have been developed and applied to engineering design problems. Sun et al. [23] proposed two different multiobjective EGO algorithms, where the infill criterion formed by the “Kriging believer” and “multiple good local optima” strategy was adopted to compromise exploitation and exploration. By extending the generalized EGO algorithm and adopting a constant liar strategy, Li et al. [24] used a Kriging-based unconstrained global optimization method to obtain multiple infill sample points. Xing et al. [25] proposed a P-EI method, combined with the effective design domain simplification technology, to search multiple peaks of the expected improvement (EI) function. Amine Bouhlel et al. [26] proposed a regional extreme criterion combined with the partial least squares method (KPLS(+K) model) to deal with the high-dimensional problems.

Besides, there were many meaningful works to enhance and balance the global and local searches for surrogate-based global optimization: Wang and Simpson [27] applied the fuzzy  $K$ -means clustering method on many cheap sample points to reduce the design space. A stochastic RBF method for the global optimization of expensive functions was proposed by Regis and Shoemaker [28], which improved the Gutmann-RBF method by varying the size of the subdomain in different iterations [29]. Villanueva et al. [30] used the  $K$ -means clustering method to partition initial design space into different subregions and explored them separately by multiagent systems. Long et al. [31] combined a set of intelligent space exploration strategies with TR-ARSM for global optimization. In addition, the multistart approach has also proved to be an efficient global optimization strategy. Regis and Shoemaker [18] introduced a QUasi-multistart Response Surface (AQUARS) locating the promising regions of the objective function by performing local searches around the local minima of the surrogate model, and it bounced between the local minima to obtain the global optima. Dong et al. [32] combined the multistart approach with space reduction strategy, introducing the multistart space reduction method. The MSSR reduced space size automatically based on the current best sample point. Eriksson et al. [33] proposed the TURBO algorithm which applies multiple local surrogate models to approximate the objective function and uses an implicit bandit method to perform a principled global allocation of samples across these models. Bartoli et al. [34] proposed a superefficient global optimization coupled with a mixture of expert (SEGOMOE) method, where an enrichment strategy based on a mixture of experts and adaptive surrogate models was applied to tackle multimodal problems with complex constraints. Diouane et al. [35] proposed a trust-region-like EGO method (TREGO), which alternated between traditional EGO steps and local steps within a trust region. To deal with the constrained black-box problems, some

researches adopted different infill criteria to figure out the feasible regions during the optimization process [36–38], and some researches integrated surrogate model with evolutionary algorithms to deal with expensive constraints [39–41], which could dramatically improve the optimization efficiency and convergence on constrained optimization problems.

These researches indicated that the design space reduction approach presented to be an efficient optimization method for multimodal and computationally expensive black-box problems. However, there were two main arguments significantly affecting the optimization stability, convergence, and efficiency:

- (a) The local minima of the surrogate model would change significantly as more sample points were evaluated, and the new subspaces were generally centered at a single best point. Thus, it was difficult to select suitable local minima to construct promising subspaces. Once it was trapped in some deceptive regions, it would cost plenty of computational effort to escape and search for other possibilities [32]
- (b) Some algorithms tended to search for new sample points around the well-explored local minima, leading to wasted computing resources, and would ignore the potential regions where better sample points might exist

The Kriging-based space exploration method (KSE) was proposed to address these problems by using multiple subregion local search strategy and the enhanced trust-region method in this paper. It combined the global search over the entire space and local search in the subspaces to balance and enhance the global exploration ability and local exploitation, thereby addressing problem (a). During the local searches, not all local minima need to be further exploited. The local minima were classified into fully and partially explored minima in terms of root mean square error (RMSE). And only the local minima with large RMSE needed further exploitation, during which an enhanced trust-region method was proposed to emphasize the optimization of the partially local minima, thereby addressing problem (b). Besides, to overcome the limitations of the original EGO, the infill criterion would be turned into minima prediction (MP) [6] to make further exploitation when all local minima were fully explored and the improvement of the EI criterion was lower than the preset threshold.

The proposed KSE method was compared to alternative global optimization methods on 12 test functions with 2 to 16 design variables and 2 aerodynamic design problems with 24 to 77 design variables. The numerical results indicated that the KSE method generally used fewer function evaluations to find the global minima on most test problems compared to alternative methods. Moreover, it could use fewer function evaluations to reach a reasonably good solution to aerodynamic design problems compared to alternatives. Hence, the KSE was very promising for the global optimization of engineering design.

The rest of the paper was organized as follows: part 2 introduced the background while part 3 focused on the framework of the proposed method. Part 4 and part 5 described the computational experiments and their results as well as the related discussion. Finally, the last part provided a summary and the conclusions.

## 2. Methodology

During the last decade, the EGO method [6] has become one of the most popular methods for global optimization. The EGO method was a sequential planning strategy that used a Gaussian process-based surrogate model, e.g., Kriging, and a variance-based merit function or criterion to generate candidate solutions. Therefore, the theories of EGO were briefly depicted in this section.

**2.1. The Kriging Model.** The Kriging was first proposed in the field of geostatistics by Krige [42] and Matheron [43]. After Sacks et al. [44] applied it to computer-based engineering design, the Kriging model became popular in engineering design and optimization.

In the Kriging model, the values of the black-box function  $f$  were assumed to be the outcomes of a stochastic process. In particular, it was assumed that the function value  $f(x)$  was a realization of a random variable  $Y(x)$  which was normally distributed with mean  $\mu$  and variance  $\sigma^2$ . For any two points  $x_i$  and  $x_j$ , the correlation between  $Y(x_i)$  and  $Y(x_j)$  could be modeled as

$$\text{Corr}[Y(x_i), Y(x_j)] = \exp\left(-\sum_{k=1}^d \theta_k |x_{ik} - x_{jk}|^{p_k}\right), \quad (1)$$

where  $1 \leq p_k \leq 2$  and  $\theta_k > 0$  ( $k = 1, \dots, d$ ) were hyperparameters of the model, and they provided more degree of freedom for the predictor to imitate the behavior of multimodal functions. Let  $y_1 = f(x_1), \dots, y_n = f(x_n)$  and  $y = (y_1, \dots, y_n)^T$  be the vector function values. Fitting the surrogate model through the  $(x_1, y_1), \dots, (x_n, y_n)$  was equivalent to finding the maximum likelihood estimates (MLEs) of the parameters  $\mu, \sigma^2, \theta_1, \dots, \theta_d, p_1, \dots, p_d$ . Then, the Kriging predictor at a new point  $x^*$  was presented as

$$\hat{y}(x^*) = \hat{\mu} + r^T R^{-1} (y - J_{n \times 1} \hat{\mu}), \quad (2)$$

where  $\hat{\mu} = J_{n \times 1}^T R_y^{-1} y / J_{n \times 1}^T R_y^{-1} J_{n \times 1}$  and  $r = (\text{Corr}[Y(x^*), Y(x_1)], \dots, \text{Corr}[Y(x^*), Y(x_n)])^T$ . The uncertainty at  $x^*$  could be given by mean squared error (MSE):

$$s^2(x) = \hat{\sigma}^2 \left[ 1 - r^T R^{-1} r + \frac{(1 - r^T R^{-1} r)^2}{J_{n \times 1}^T R^{-1} J_{n \times 1}} \right], \quad (3)$$

where  $\hat{\sigma}^2 = (y - J_{n \times 1} \hat{\mu})^T R^{-1} (y - J_{n \times 1} \hat{\mu}) / n$ . The first term represented the uncertainty due to the correlation of the sample points, and the second term described the uncertainty of

the estimated parameter  $\mu$ . For more details, Jones [21] gave a taxonomy of surrogate-based optimization methods.

**2.2. The EGO Method.** The EGO method belonged to the so-called ‘‘two-stage’’ method: the first step consisted of fitting the surrogate model and estimating the associated hyperparameters. The second step consisted of using this surrogate model instead of the true function and searching for promising sample points. It used (4) to find new sample points.

$$\text{EI} = \begin{cases} (f_{\min} - \bar{y}(x)) \Phi\left(\frac{f_{\min} - \bar{y}(x)}{s(x)}\right) + s(x) \phi\left(\frac{f_{\min} - \bar{y}(x)}{s(x)}\right), & \text{if } s(x) > 0, \\ 0, & \text{if } s(x) = 0. \end{cases} \quad (4)$$

Here,  $f_{\min}$  was the current best objective function value. The  $\Phi$  and  $\phi$  were the cumulative density function and the standard normal probability density functions.

For an inequality constraint function as  $g(x) > g_{\text{limit}}$ , the constraint function was also modeled by a Gaussian process. And this model could be used to calculate the probability that the constraint was met. So, the ‘probability of feasibility’ for each constraint  $g(x)$  would be calculated in (5):

$$P[g(x) > g_{\text{limit}}] = \frac{1}{\hat{s}(x) \sqrt{2\pi}} \int_0^{\infty} e^{-[(G(x) - g_{\text{limit}}) - \hat{g}(x)]^2 / (2\hat{s}^2(x))}, \quad (5)$$

where  $G(x)$  was a normally distributed random variable with mean value  $\hat{g}(x)$  and variance value  $\hat{s}^2(x)$  and  $G(x) - g_{\text{limit}}$  measured the feasibility of new points. Provided that the constraint and objective functions were all independent models, Sasena et al. [45] proposed the constrained expected improvement (CEI) method to handle constraints by multiplying the expected improvement of the objective function and the probability of feasibility as

$$\text{CEI} = \text{EI} \times P[g(x) - g_{\text{limit}}]. \quad (6)$$

Locatelli [46] proved that the EGO method generally took plenty of function evaluations to find the global optima. During the EI-based global optimization, the  $s(x)$  was usually slightly lower than its true error value [47]. Unless the uncertainty of the current region became relatively low, the underestimation would cost many iterations to explore the local spaces, especially for some deceptive regions with lower estimated standard deviation, causing slower convergence rate.

**2.3. Trust-Region Method.** The trust-region method (TRM) [48] has proven to be an efficient space reduction method for its good performance in optimization [49]. It identified the promising space, where the actual global optima probably exist, in light of how well the current surrogate model predicted the improvement of the objective function. In TRM, the current best sample point was chosen as the center of the new space. The trust-region factor  $r_k$  and trust-region

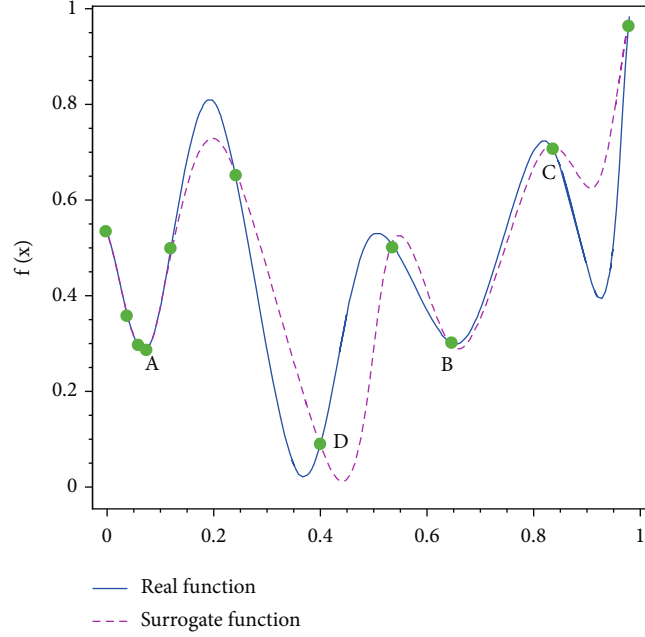


FIGURE 1: The sketch of local minima.

radius  $\delta_k$  were computed through

$$r_k = \frac{f(x_{k-1}) - f(x_k)}{f(x_{k-1}) - \tilde{f}(x_k)} = \frac{\Delta f}{f_{k-1} - \tilde{f}_k}, \quad (7)$$

$$\delta_k = \begin{cases} c_1 \|x_k - x_{k-1}\|, & r_k < r_1, \\ \min(c_2 \|x_k - x_{k-1}\|, \delta_{\max}), & r_k > r_2, \\ \|x_k - x_{k-1}\|, & r_2 \geq r_k \geq r_1, \\ \delta_{\min}, & r_k < 0. \end{cases} \quad (8)$$

The  $f(x_k)$  and  $\tilde{f}(x_k)$  were the true and predicted values, respectively.  $r_k$  could measure the improvement of objective function and accuracy of surrogate model at the same time through  $\Delta f$  and  $f_{k-1} - \tilde{f}_k$ . If  $\Delta f > 0$ , the current surrogate model was capable of decreasing the true objective value, and then, the trust radius  $\delta_k$  was updated according to (8) [49], where the typical values of  $r_1, r_2, c_1, c_2$  were used, namely,  $r_1 = 0.10, r_2 = 0.75, c_1 = 0.75$ , and  $c_2 = 1.25$ . The  $\delta_{\max}$  was the upper bound of the trust radius, which was usually set to 0.5, and  $\|\cdot\|$  indicated the  $L_1$  Euclidean norm. Otherwise, the current surrogate model could not improve the objective value for its bad fitting quality. In this case, the minimum trust radius  $\delta_{\min}$  was used, which usually lies in the interval  $[0.01, 0.1]$ . Then, TRM was forced to exploit a small neighborhood around the present best point. The TRM allowed optimization to focus on a small region around the current point with dynamic bounds, and it was very efficient in local search, avoiding the prior knowledge of constructing local space.

### 3. Proposed Method

**3.1. Classification of Local Minima.** From the discussion above, the EGO method attempted to explore the entire design space which would require numerous computational resources. Once it focused on some deceptive local minima, it was hard for it to escape from local minima and explore other potential regions in time. An unbalanced approach to exploitation and exploration usually caused a poor depiction of the design space, as well as the convergence and efficiency of optimization.

To overcome these issues, this paper proposed the Kriging-based space exploration method (KSE). The multiple subregion local search strategy was used to explore the local regions in the entire design space, improving the depiction of the design space and surrogate model accuracy. One of the main aspects of this strategy was to select reasonable potential regions to make deep exploitation. It classified the local minima of the surrogate model into fully and partially explored local minima, from which the partially explored minima were selected to make further exploitation. The classification of partially explored local minima was relayed on the depiction of local regions, defined by the root mean square error (RMSE) of local regions. The poor fitting quality of the surrogate model would lead to substantial fluctuations in design space, resulting in misleading results for the algorithm. And a good depiction of the entire design space, which relied on depicting vicinities of local minima, could improve the fitting quality of the surrogate model. Thus, the main purpose of multiple subregion local search strategy was to make the proper depiction of potential local regions and provided efficient information for global optimization.

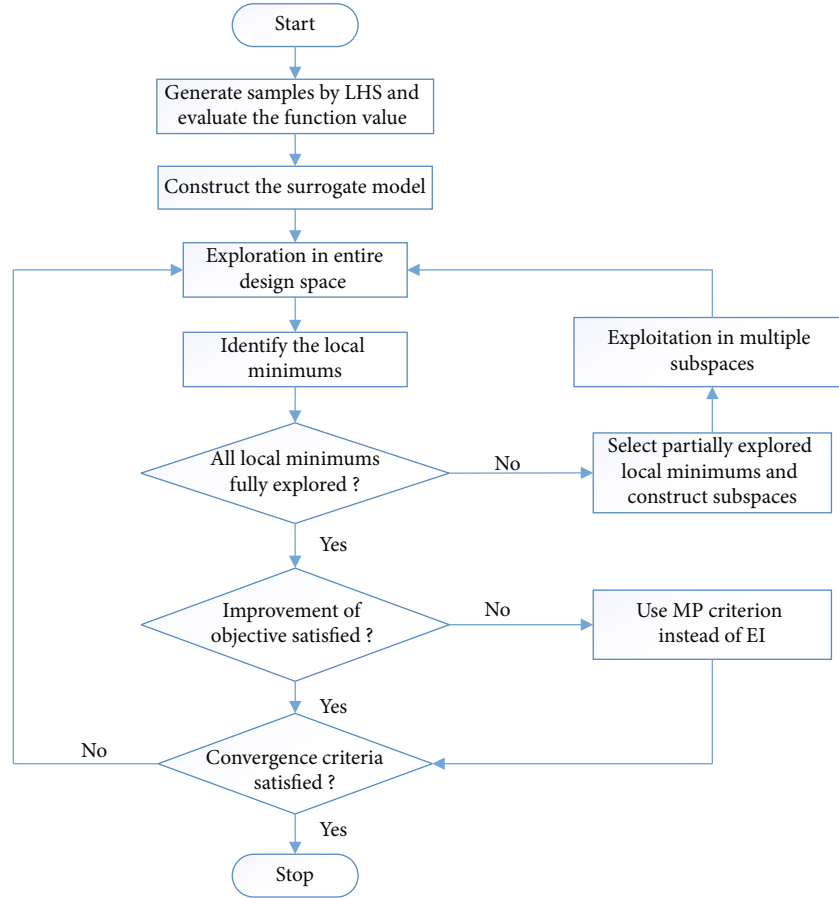


FIGURE 2: Flowchart of the KSE optimization process.

Let design space  $D = [a, b] \subseteq \mathbb{R}^d$  and  $f : D \rightarrow \mathbb{R}^d$  be any function from  $D$  to  $\mathbb{R}^d$ , whose local regions were nearly linear. Suppose  $x^*$  was a local minima of the objective function  $f$  over  $D$ ; then, there existed radius  $\rho > 0$ , so that  $f(x^*) \leq f(x)$  for all  $x \in L(x^*, \rho) \cap D = \{x \in D : \|x - x^*\| \leq \rho\}$ . Similarly, the  $\hat{f}(x_s^*) \leq \hat{f}(x)$  for the surrogate model with local minima  $x_s^*$  over local space  $\tilde{L}(x_s^*, \rho)$ . For a well-performed surrogate model, the  $x^*$  would be consistent with the  $x_s^*$  and the trend of  $\tilde{L}(x_s^*, \rho)$  in surrogate model would be similar to that of  $L(x^*, \rho)$  in objective function.

Otherwise, it would need more sample points to depict the neighborhood properly, like  $C$  in Figure 1. Under this circumstance, the root mean square error (RMSE) of sample points within the local region was used to decide whether the corresponding local minima were fully explored or not, and the partially explored local minima would need further exploration. During these processes, the exploitation of partially explored local minima of the surrogate model was used to improve the fitting quality of the surrogate model in local space, and the exploitation of partially explored local minima of the objective function was used to search for better points. Local minima for the surrogate model did not always correspond to a local minima for the objective function. However, the deep exploitation in the vicinity of local minima of the surro-

gate model could eliminate the “fake local minima” and obtain the true local minima for the objective function. Since the local minima with better performance were more likely to be close to the global optima, the further exploitation of the local minima of objective function would make the algorithm toward the global optima more efficient. In combination, the global exploration for the target value would be promoted.

The above conditions could not guarantee that the surrogate model would achieve a good approximation to the objective function  $f$  in some neighborhoods of local minima. However, if  $f$  had a relatively simple surface in the neighborhood of  $x^*$  (e.g., if it was nearly linear), then the possibility that  $f$  was well approximated by the surrogate model in that local region would be increased by those conditions. Thus, the fully explored local minima of the surrogate model were more likely to be the local minima of the objective function, as shown in Figure 1.

**3.2. The Enhanced Trust-Region Method.** Although the traditional trust-region method (TRM) has shown the advantages of improving the optimization efficiency, it might not make sufficient search in local regions since the space radius of new design space was largely dependent on the previous best point (the radius from  $\|x_k - x_{k-1}\|$  was linearly dependent),



**The KSE for Global Optimization**

**Input:** (1) original design space  $D = [X \text{ max}, X \text{ min}] \subset \mathbb{R}^d, d \geq 1$ ;

(2) global exploration iteration number  $m\_global$ ;

(3) cluster number  $N$ , cluster radius  $r_{req}$ ;

(4) maximum iteration number  $Max\_k$ ;

(5) FunctionCriterion: EI function;

(6) threshold of RMSE:  $RMSE_{threshold}$

**Output:** the best solution as global optima  $(x^*, y(x^*))$ ;

**Begin**

Choose an initial set of sample points  $X_D = [x_1^D, \dots, x_n^D] \in D$  by LHS sampling method.

Evaluate  $X_D$  to obtain the corresponding response  $Y_D = [y_1^D, \dots, y_n^D]$ , and store them in the database  $S$ .

While  $k < Max\_k$

for  $i = 1 : m\_global$

Employ EGO algorithm with FunctionCriterion;

Store the optimized solution  $x$  and its function response value  $y$  in the database  $S$ ;

$k = k + 1$ ;

end

Select  $N$  clusters from database  $S$ , and identify the cluster center  $X^* = [x_1^*, \dots, x_N^*]$ , other sample points within the distance  $r_{req} = 0.01 \sqrt{|X \text{ max} - X \text{ min}|^2}$ , would be classified into the corresponding cluster  $S^* = \{S_1^*, \dots, S_N^*\}$ . The distance between sample points with cluster center was computed by Euclidean distance.

for  $j = 1 : N$

Sort all sample points  $X_c^* = [x_1, \dots, x_n]$  of  $S_j^*$  in terms of their function response value  $Y_c^* = [y_1, \dots, y_n]$  in ascending order;

Use Cross-validation to obtain the predicted value  $Y_c^* = [y_1, \dots, y_n]$  of all sample points;

$RMSE = 1/n \sum_{i=1}^n (y_i - \tilde{y}_i)$ ;

if  $RMSE \leq RMSE_{threshold}$

Store  $x_i^*$  into fully explored local minima dataset  $X^f = [x_1^f, \dots, x_n^f]$ ;

else

Store  $x_i^*$  into partially explored local minima dataset  $X^p = [x_1^p, \dots, x_n^p]$ ;

end

end

if  $X^p \neq \emptyset$

for  $i = 1 : n$

Define local space  $V_{TR}$  centered at  $x_i^p$  by the ETRM algorithm;

Employ EGO algorithm with FunctionCriterion in  $V_{TR}$ ;

Store the optimized solution  $x$  and its function response value  $y$  in the database  $S$ ;

$k = k + 1$ ;

end

else

Select current best sample point value  $f_{best}$ .

$C_{EI} = \text{ComupteFunctionCriterion}(EI_{max}, f_{best})$

if  $C_{EI} \geq 0.01$

FunctionCriterion  $\leftarrow$  EI criterion;

else

FunctionCriterion  $\leftarrow$  MP criterion;

end

end

end

**End**

ALGORITHM 1: The encoding type of KSE.

which would lead to premature of the optimization. To avoid this shortcoming, this paper proposed an enhanced trust-region method (ETRM) combining the independent/dependent strategy with the stochastic perturbation strategy, as shown below.

$$D_{new} = D_{independent} \cup D_{dependent} \cup D_{stochastic}, \quad (9)$$

$$A_{EI} = \frac{\max(f_{min} - f(x_{k-1}), 0)}{EI}, \quad (10)$$

$$r_k = \begin{cases} \min(\gamma_1 r_{k-1}, r_{max}), & A_{EI} \geq \eta, \\ r_{k-1}, & 0 < A_{EI} < \eta, \\ \max(\gamma_2 r_{k-1}, r_{min}), & A_{EI} = 0, \end{cases} \quad (11)$$

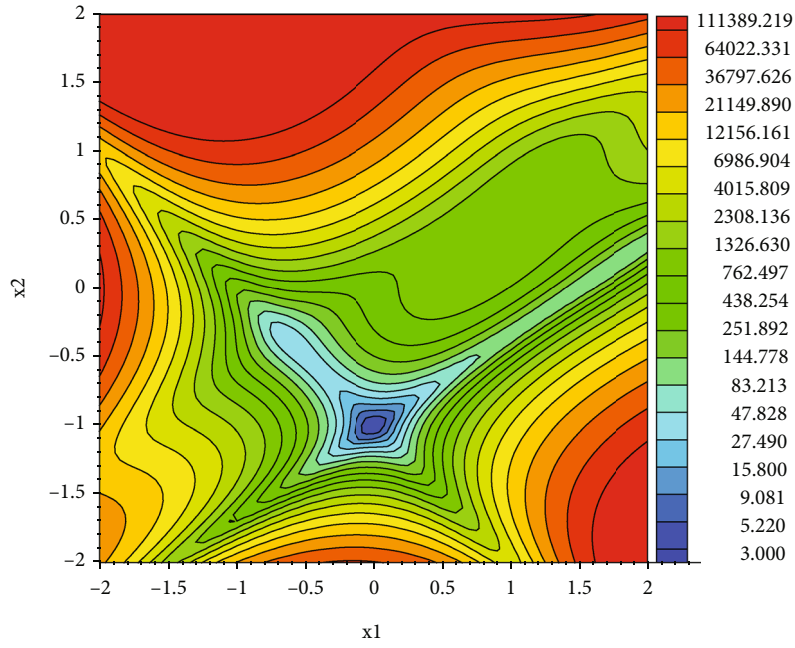


FIGURE 3: Original GP function.

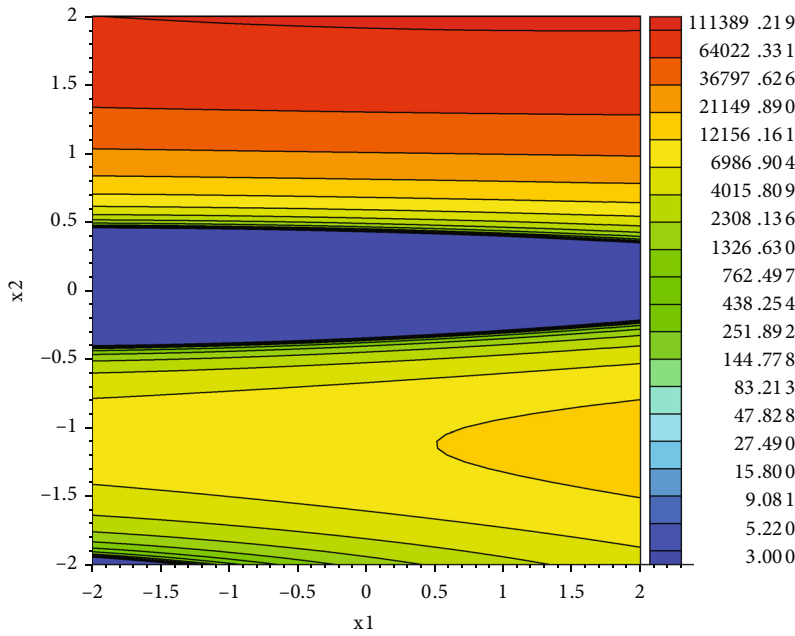


FIGURE 4: Kriging prediction with initial sample points.

$$D_{\text{independent}} = [x_{\min} - r_k \|b - a\|, x_{\min} + r_k \|b - a\|], \quad (12)$$

$$D_{\text{dependent}} = [x_{\min} - r_k(x_{\min} - x_{\min,i-1}), x_{\min} + r_k(x_{\min} - x_{\min,i-1})],$$

$$x_{\min} \in X_{\min}, X_{\min} = [x_{\min,1}, \dots, x_{\min,i}], \quad (13)$$

$$P(k) = 1 - \frac{\ln(k)}{\ln(N_{\max})},$$

$$I_{\text{pertube}} := \{i : \omega_i < P(k), \omega_i \in [0, 1], i = 1, \dots, d\}, \quad (14)$$

$$\delta_k = I_{\text{pertube}}(r_{\max} \|b - a\| \cdot \text{rand}(d, 1)),$$

$$D_{\text{stochastic}} = [x_{\min} - \delta_k, x_{\min} + \delta_k].$$

Here, the  $D_{\text{new}}$  was the new local space defined by the

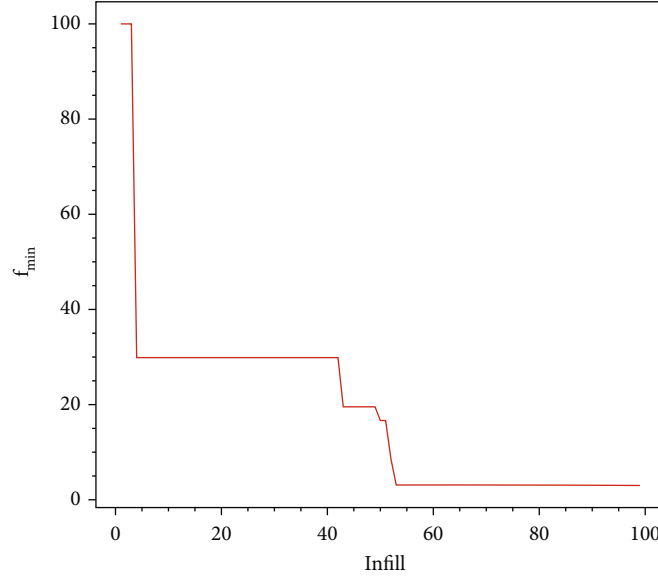


FIGURE 5: Convergence history of the objective function.

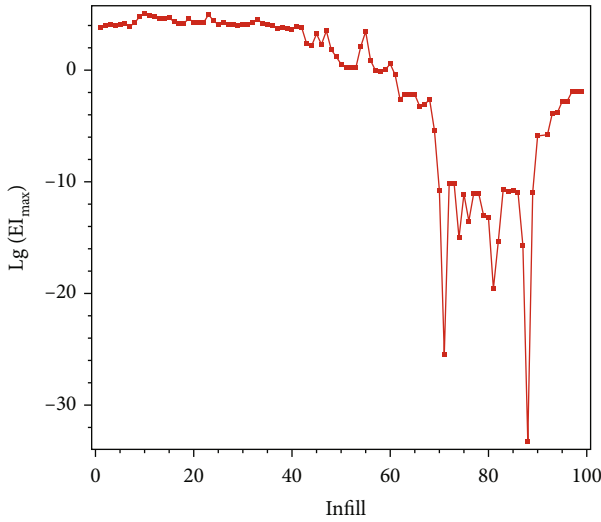


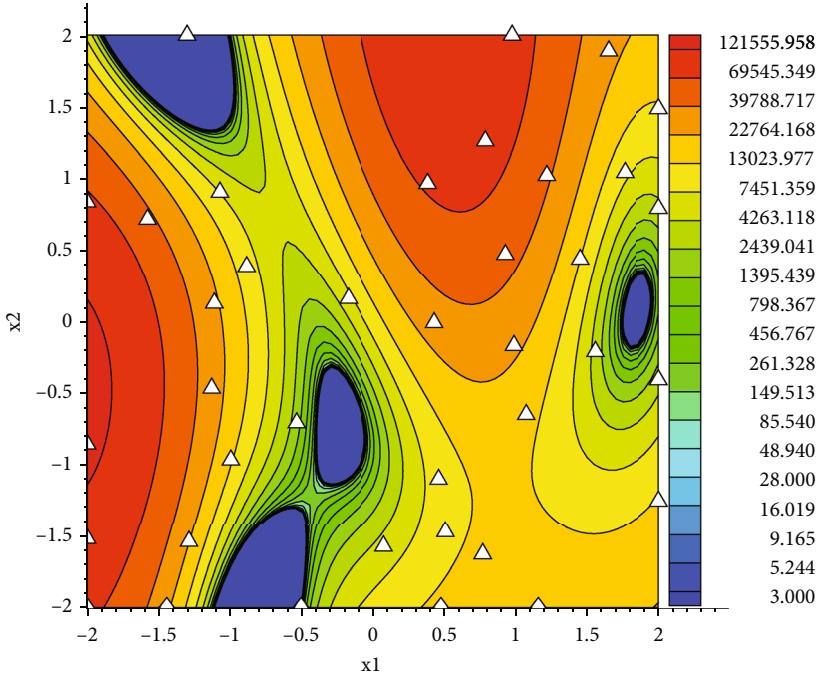
FIGURE 6: Convergence history of max EI.

$D_{\text{independent}}$ ,  $D_{\text{dependent}}$ , and  $D_{\text{stochastic}}$ , which were computed by different strategies, as shown in (9). The space radius of the ETRM changed in terms of the actual improvement of EI ( $A_{\text{EI}}$ ) while the original TRM adopted the predicted improvement based on the surrogate model as shown (8). The trust factor  $r_k$  was computed through (10), where typical values of  $r_{\text{min}} = 0.04$ ,  $r_{\text{max}} = 0.08$ ,  $\gamma_1 = 1.2$ , and  $\gamma_2 = 0.8$  were used, respectively. If  $A_{\text{EI}} \geq \eta$ , where the  $\eta$  was the threshold for the  $A_{\text{EI}}$ , the EI criterion was capable of decreasing the objective function value; the space radius would be expanded according to (11). If  $0 < A_{\text{EI}} < \eta$ , the space radius would keep unchanged. If  $A_{\text{EI}} = 0$ , the surrogate model was unable to obtain better points in current space; then, a conservative space radius would be selected. (12) used the independent strategy from the current best point  $x_{\text{min}}$  and the entire

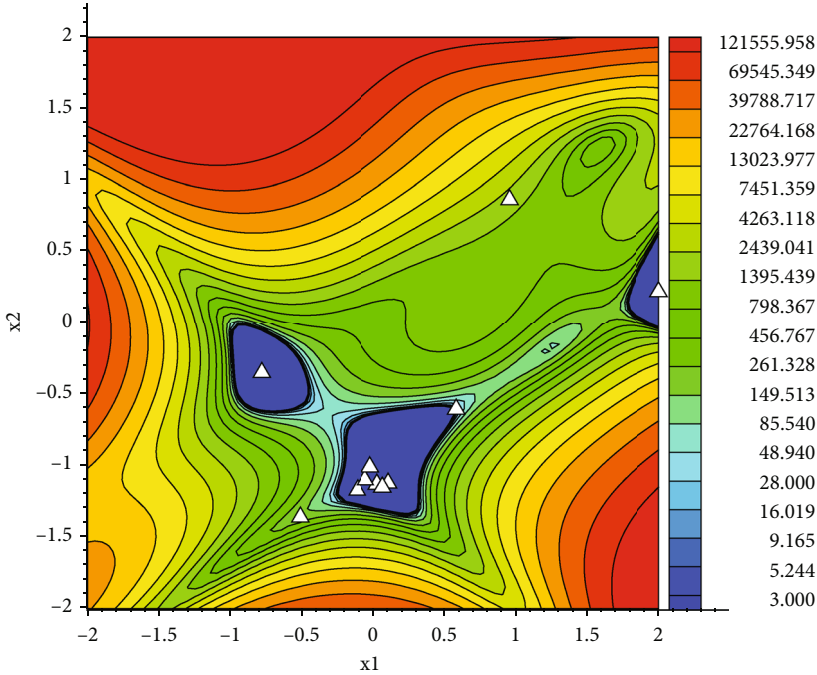
design space to construct a uniform box-shaped trust region, where  $a$  and  $b$  were the lower and upper bound of the entire design space, respectively, and  $\| \cdot \|$  was the Euclidean norm. The purpose of this strategy was to reduce the influence of the previous sample points and make the optimization more independent to find diverse sample points in each iteration. However, the dependent information ( $x_{\text{min}} - x_{\text{min},i-1}$ ) from the previous best point  $X_{\text{min}} = [x_{\text{min},1}, \dots, x_{\text{min},i}]$  could maintain the efficient dimension information which would help the algorithm toward the global optima quickly. Thus, the dependent information was used to improve the dimension setting of the trust region, which was indicated by the vector direction from the previous best point  $x_{\text{min},i-1}$  to the current best point  $x_{\text{min}}$ , as shown in (13). Besides, to make a full exploration of the local region, a stochastic strategy for trust region was presented by (14). Here, the  $P(k)$  was the probability of perturbing any dimensions of design space, and  $N_{\text{max}}$  was the maximum number of iterations. For the iteration  $k$ , the number of the dimension to be perturbed was a random variable with a binomial ( $d, P(k)$ ) distribution with a mean number of perturbed dimensions  $d \times P(k)$ , where  $d$  was the number of all dimensions in the design space. In (14), the  $P(k)$  was decided by a strictly decreasing function and would decrease as the optimization carries on. The  $I_{\text{pertube}}$  was the set of the index of the perturbed dimension decided by  $P(k)$  and  $\omega_i$ , where  $\omega_i$  was a uniform random number generated in the range  $[0, 1]$ , and it would be different in each iteration to maintain the diversity of design space. Then, the  $D_{\text{stochastic}}$  was computed by the trust radius  $\delta_k$  and the current best point  $x_{\text{min}}$ . The use of a stochastic strategy for constructing a trust region was expected to be helpful for problems with many local minima and design variables.

At the initial stage of optimization, the distribution of  $x_{\text{min}}$  could be very sparse and their positions would change significantly in different iterations since the surrogate model accuracy would not be very precise. Under this condition, the dependent strategy could make use of useful dimension



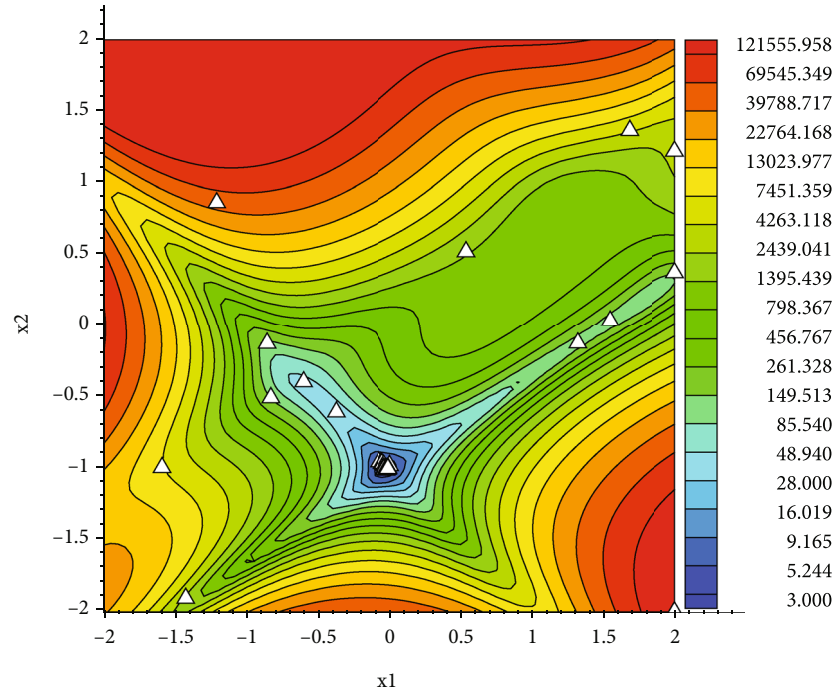


(a) The first stagnation process (iteration 4 to 44)

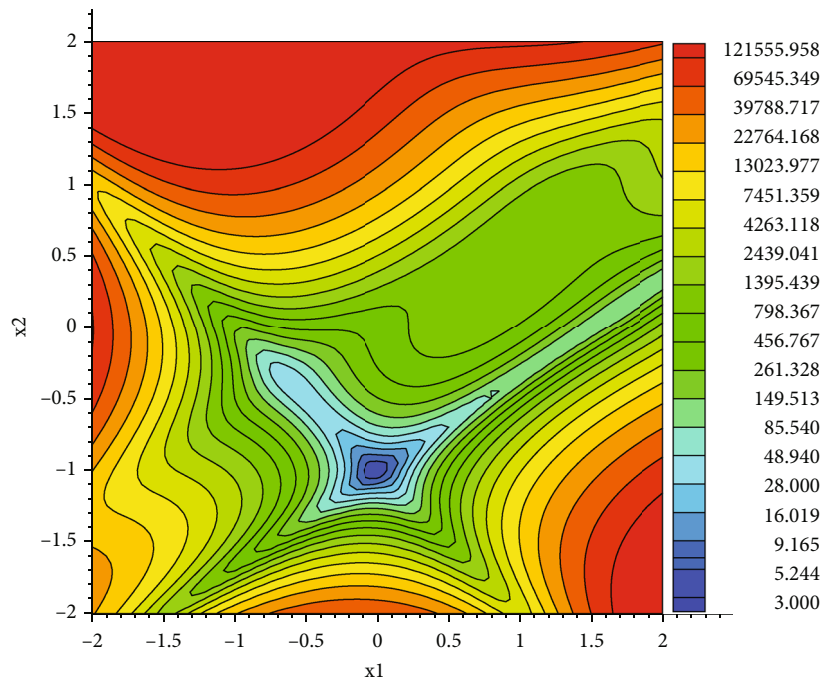


(b) The efficient searching process (iteration 45 to 53)

FIGURE 7: Continued.



(c) The second stagnation process (iteration 54 to 99)



(d) The final surrogate model

FIGURE 7: The infill process of the EGO method.

information to accelerate the convergence speed to better points. And the stochastic strategy would help the algorithm to search for multiple diverse sample points by perturbing a large fraction of the dimensions. As the progress of optimization, the distribution of  $x_{\min}$  would be more centralized and the distance between the best points would be smaller. By perturbing a progressively smaller fraction of dimension,

the vicinity of the current best point  $x_{\min}$  was less likely to be disturbed. At present, the trust region would be largely dependent on the independent strategy, which ensured that the neighborhood of  $x_{\min}$  would be focused throughout the process of optimization.

The ETRM measured the improvement of the objective function and combined an independent/dependent strategy

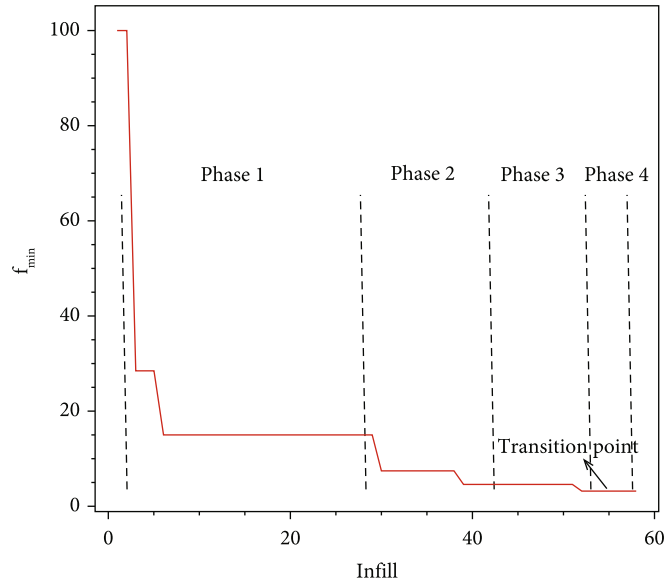


FIGURE 8: Convergence history of the objective function.

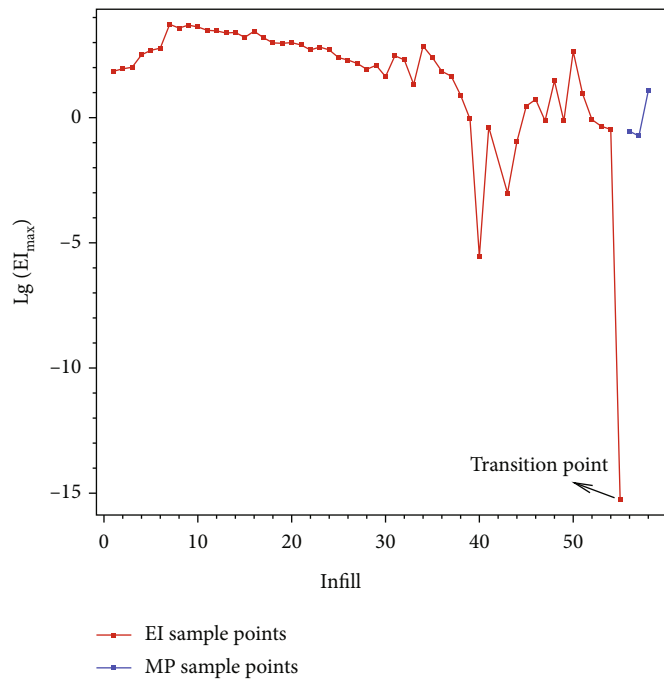


FIGURE 9: Convergence history of max EI.

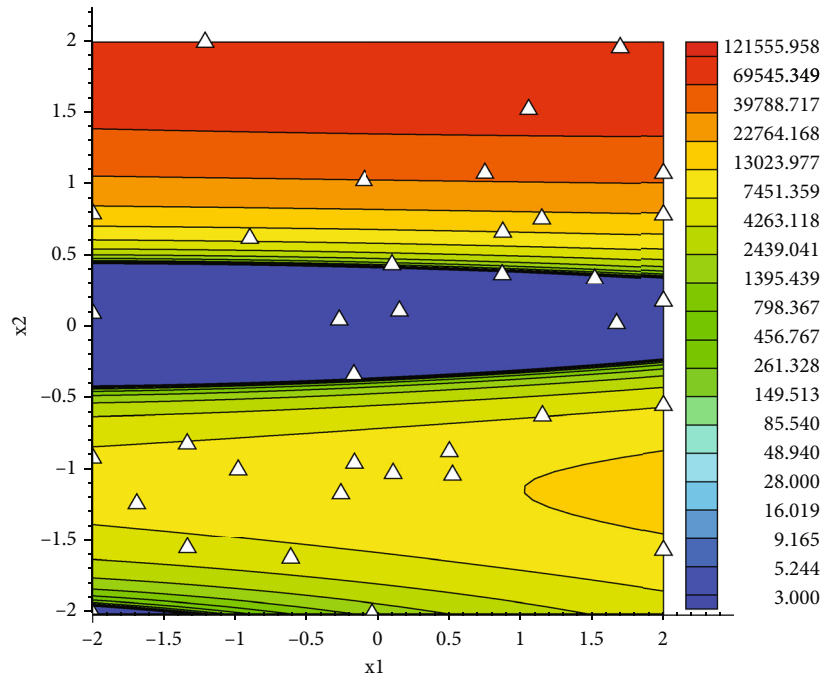
with the stochastic perturbation strategy to construct the promising local regions. These strategies could deal with different situations across the optimization process, and in combination, they could preserve the diversity of design space and prevent the premature of optimization, improving the efficiency and convergence of optimization.

**3.3. Detail of Proposed Method.** The proposed KSE method consisted of four parts: (1) search effective samples and select local minima; (2) classify local minima into fully and partially explored local minima; (3) exploit the partially explored local minima; and (4) combine global exploration

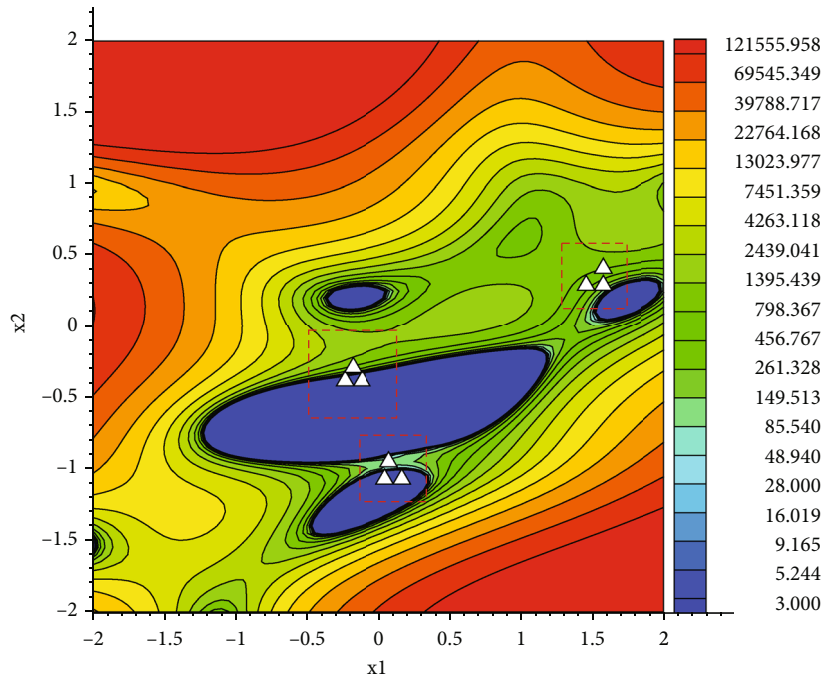
and local exploitation. The flowchart of the method is illustrated in Figure 2, and the following were the detailed steps of this method.

*Step 1* (initial design). In the beginning, Latin hypercube sampling (LHS) method [50] was adopted to generate initial sample points.

*Step 2* (global search by EI criterion). The EGO method based on DACE, where the Gaussian kernel function was adopted, was used to explore the entire design space, during which the particle swarm optimization (PSO) algorithm was

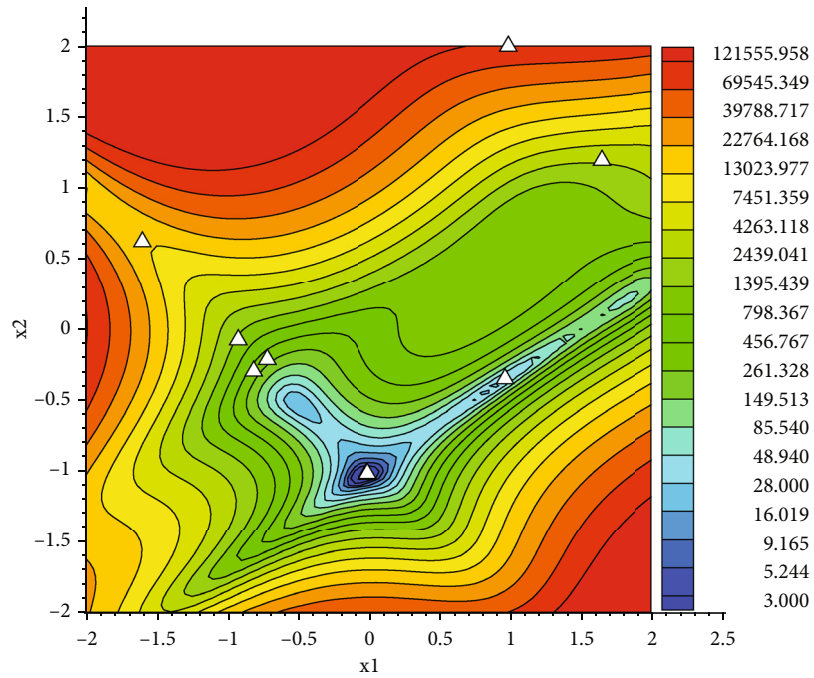


(a) Iterations 1 to 36

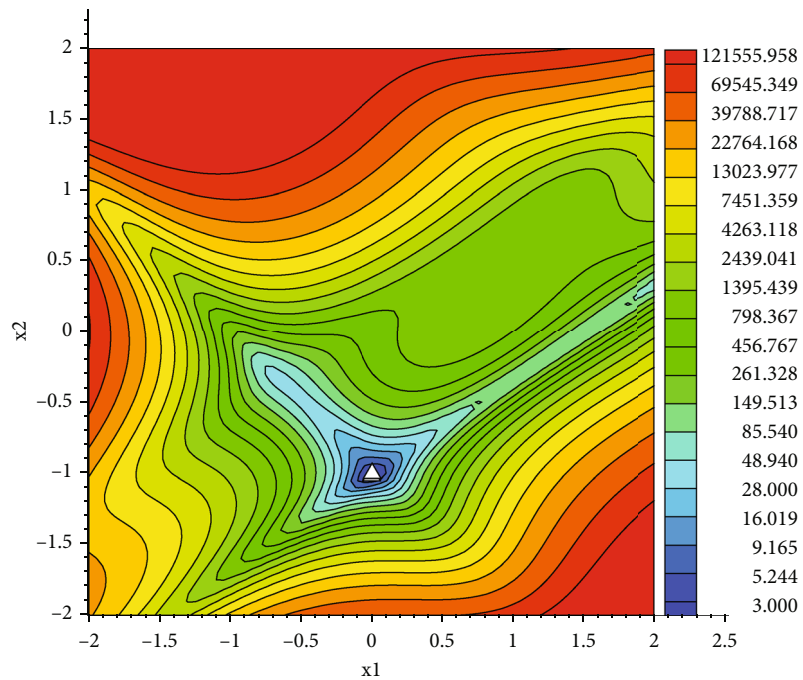


(b) Iterations 37 to 45

FIGURE 10: Continued.



(c) Iterations 46 to 55



(d) Iterations 55 to 58

FIGURE 10: The infill process of the KSE method.

adopted as an optimizer. As for the problems with multiple constraints, the constrained expected improvement criterion (CEI) [45] would be adopted to handle constraints in KSE.

*Step 3* (identify the local minima). The sample points with good performance were selected as local minima and corresponding sample cluster.

*Step 4* (select the partially explored local minima). Classify the local minima into fully and partially explored local minima.

*Step 5* (multiple subregion local search). Deep exploitation of all partially explored local minima, based on the ETRM, would be carried out to improve the surrogate accuracy. After all local minima have been explored to their fullest



TABLE 1: Bound-constrained benchmark problems for global optimization.

Test functions	Dim	Domain	No. of local min	Global min value
Branin	2	$[-5, 10] \times [0, 15]$	3	0.3980
Goldstein-Price	2	$[-2, 2]^2$	4	3.0000
Hartman3	3	$[0, 1]^3$	4	-3.8628
Shekel5	4	$[0, 10]^4$	5	-10.1532
Hartman6	6	$[0, 1]^6$	6	-3.3223
SC	2	$[-2, 2]^2$	6	-1.0316
Shekel7	4	$[0, 10]^4$	7	-10.4029
Shekel10	4	$[0, 10]^4$	10	-10.5364
Shubert	2	$[-10, 10]^2$	>10	-186.7309
Trid10	10	$[-100, 100]^{10}$	—	-210.0000
Sphere10	10	$[-5.12, 5.12]^{10}$	—	$\leq 0.001$
F16	16	$[-1, 1]^{16}$	—	25.8750

extent, the algorithm would proceed to Step 2 until no partially explored local minima could be identified.

*Step 6* (switch criterion function). At the last stage of optimization, the MP criterion would be adopted instead of EI criterion based on the improvement of the EI criterion. The objective function was set as (15), and the function transition factor  $C_{EI}$  measured the improvement of the EI criterion, as shown in (16). The  $f_{\min}$  was the objective value of the current best sample point, and  $EI_{\max}$  was the improvement of the EI criterion. In this paper, the  $C_{EI}$  was set as 0.01, which meant if the  $EI_{\max}$  was less than 1% of  $f_{\min}$ , the objective function turned into the MP criterion.

$$\text{Objective} = \begin{cases} \text{EI criterion,} & C_{EI} \geq 0.01, \\ \text{MP criterion,} & \text{else,} \end{cases} \quad (15)$$

$$C_{EI} = \frac{EI_{\max}}{f_{\min}}. \quad (16)$$

## 4. Illustrative Examples

*4.1. Example of the EGO Method.* To illustrate the limitations of the traditional EGO method, the Goldstein-Price function [51] was used as an example. This 2- $d$  function was usually evaluated on the  $[-2, 2]^2$ , with 3 local minima. And the global optima was located at  $(0, -1)$  with the value of 3.000. Figure 3 describes the original GP function while Figure 4 shows the surrogate space constructed by the Kriging model with initial  $(d + 1) * 2$  sample points [18], where  $d$  was the dimension. It demonstrated that the initial surrogate mode could not achieve a good approximation of the objective function.

The convergence history of the GP function is shown in Figure 5, where two obvious stagnations happened. The first stagnation was in the global exploration process, during which

the maximal EI value was very large and the optimizer tended to explore the entire design space and improve the surrogate accuracy, as presented in Figure 6. During the late stages of optimization, the optimizer searched the neighborhood of the global optima. Since the objective values of sample points in that local space were close to that of the global optima, the EI criterion showed very little improvement and the optimizer tended to stochastic search. Thus, the optimization was trapped again.

The sample infilling process is presented in Figure 7, in which the contour was the space constructed by the Kriging model, and white triangles were the infilled sample points. At the beginning of optimization, the fitting quality of the Kriging model was not that good and it would take a large number of sample points to explore the entire design space, during which the first stagnation happened, as shown in Figure 7(a). After that, the surrogate accuracy was improved, and the optimizer would turn into an efficient searching process during which the objective value was efficiently decreased. Figure 7(b) presents that the surrogate model achieved a much better approximation to the objective function, and sample points with good results have been obtained. However, at the later stage of the optimization, the first term of the EI function would become small since the difference between the prediction value of unknown points and the current best point was very small, and the second term of the EI function would dominate the optimization trend. Thus, the optimizer tended to stochastic search in unknown regions, during which the EI criterion could not make substantial improvement. Therefore, the optimization was trapped again, as shown in Figure 7(c). After 45 sample points were updated, the global optima would be found at the 99th iteration. Figure 7(d) presents the final surrogate model.

The process above demonstrated that the EGO method explored the entire design space at the price of plenty of computational costs, for it took a large number of sample points to improve the fitting quality of the surrogate model in the entire design space and get rid of deceptive regions. Besides, the EI criterion tended to be maximized in the regions close to the bounds of the design space and far away

TABLE 2: Parameter setting of KSE method.

Method	Parameter	Value
KSE	Cluster number $N$	5
	Objective function transition factor $C_{EI}$	0.01
	Distance coefficient	0.01
MLSL-SQP	Reduced sample coefficient $\xi$	0.5
	Fixed number in critical distance $\rho$	4
	Number of local minima $w$	10
	Termination tolerance for current design variable	0.001
	Termination tolerance for objective function value	0.0001
	Maximum number of iteration	500
MPS	Mode-pursing sample points $n_p$	The number of dimension
	Model fitting tolerance coefficient $c_d$	0.01
	Diff tolerance $\varepsilon_r$	0.00001
	Speed control factor $r$	1
MADS	Mesh size update parameter $w^-$	-1
	Mesh size update parameter $w^+$	1
	Fixed rational number $\tau$	4
	Initial mesh size parameter $\Delta_0^m$	1
	Initial poll size parameter $\Delta_0^p$	1
EGO	Iteration number $N$	800
	Population number pos	Ten times the dimension

TABLE 3: Average number of function evaluations (over 30 trials) to get a relative error of &lt;1%.

Problems	KSE		MLSL-SQP		MPS		MADS		EGO		The best-published results from the comparison of TR-ARSM, MSSR, AQUARS, and TRIKE-Restart Algorithm	
	Results	$N_e$	Results	$N_e$	Results	$N_e$	Results	$N_e$	Results	$N_e$		
Branin	22.7 (3.69)	30	24.4 (1.52)	30	57.3 (2.15)	30	31.5 (1.35)	30	40.8 (2.33)	30	25.6	TRIKE-restart
Goldstein-Price	43.6 (4.31)	30	56.8 (9.63)	30	132.4 (6.25)	30	63.6 (4.77)	30	98.5 (10.77)	30	36.2	TR-ARSM
Hartman3	23.4 (2.21)	30	42.3 (3.36)	30	>38.9(-)	27	41.7 (2.55)	30	42.1 (3.88)	30	26.5	TRIKE-restart
Shekel5	164.9 (33.3)	30	>368.9(-)	25	>446.3(-)	5	213.5 (37.1)	30	>139.9(-)	9	202.4	TRIKE-restart
Hartman6	131.1 (11.4)	30	157.1 (12.6)	30	>500(-)	0	144.6 (13.62)	30	154.2 (9.63)	30	54.5	MSSR
SC	21.6 (3.46)	30	37.5 (2.77)	30	32.3 (5.66)	30	26.4 (2.11)	30	49 (6.37)	30	22.5	MSSR
Shekel7	144.3 (22.5)	30	>362.3(-)	22	>137.6(-)	13	168.5 (19.5)	30	>500 (30)	0	152.7	AQUARS
Shekel10	137.5 (9.62)	30	>414.2(-)	17	>455.7(-)	7	136.8 (11.2)	30	>500 (30)	0	121.1	AQUARS
Shubert	124.5 (11.3)	30	237.1 (21.7)	30	>500(-)	0	143.7 (12.6)	30	>437.5 (22)	8	122.9	MSSR
Trid10	114.7 (13.5)	30	189.4 (11.4)	30	125.2 (5.74)	30	145.6 (4.55)	30	>500 (30)	0	142.4	MSSR
Sphere10	107.6 (9.85)	30	172.5 (8.96)	30	131.5 (9.68)	30	177.6 (11.12)	30	>500 (30)	0	115.4	MSSR
F16	135.2 (8.57)	30	165.1 (7.3)	30	>500(-)	0	159.7 (8.55)	30	337.4 (15.9)	30	137.7	MSSR

from the previous sample points, which made it inefficient for finding the global optima. To overcome these problems, this paper proposed a Kriging-based space exploration method.

*4.2. Example of the Proposed Method.* To better demonstrate the KSE method, the optimization process of the Goldstein-Price function was shown below. Figures 8 and 9 present the convergence history of the objective function value and the

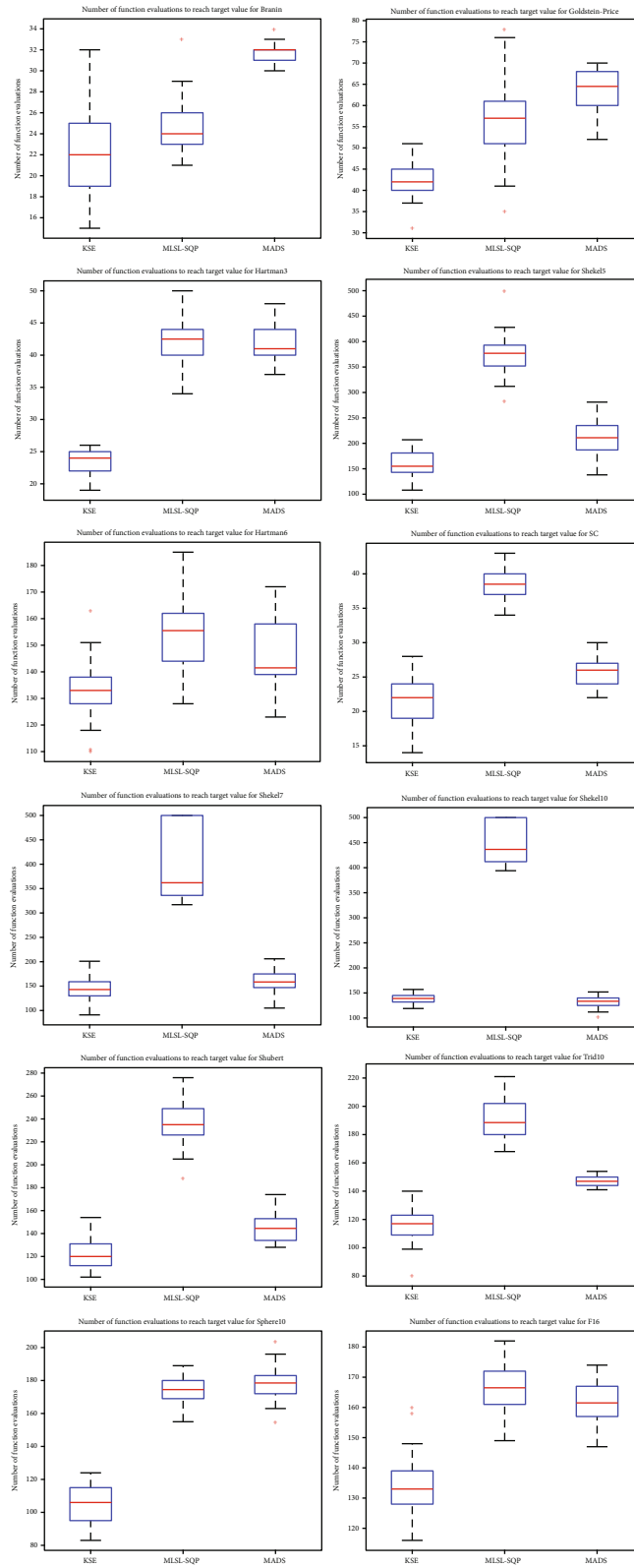


FIGURE 11: Box plot of the number of function evaluations to get within 1% of the target value.

$C_{EI}$ , respectively. There were mainly four phases during the optimization process. Phase 1 was the global exploration process based on the EI criterion. However, during this process, the objective value had little improvement after many

iterations due to the bad fitting quality of the Kriging model. Then, three local minima were found and multiple subregions of local searches started to make deep exploitation, during which the objective value decreased a lot, as shown

TABLE 4: Average number of function evaluations (over 30 trials) to get a relative error of &lt;1%.

Problems	$\langle N = 5, C_{EI} = 0.01 \rangle$	$\langle N = 3, C_{EI} = 0.01 \rangle$	$\langle N = 7, C_{EI} = 0.01 \rangle$	$\langle N = 5, C_{EI} = 0.005 \rangle$	$\langle N = 5, C_{EI} = 0.02 \rangle$	$\langle N = 5, C_{EI} = \text{null} \rangle$
Branin	22.7 (3.69)	21.2 (2.52)	24.5 (1.04)	23.9 (2.97)	22.2 (3.01)	24.6 (1.94)
Goldstein-Price	43.6 (4.31)	44.8 (6.31)	45.8 (5.15)	44.9 (3.17)	45.4 (3.69)	45.7 (3.74)
Hartman3	23.4 (2.21)	47.8 (2.38)	39.9 (3.44)	22.7 (4.19)	26.6 (3.22)	46.9 (4.21)
Shekel5	164.9 (33.3)	179.5 (42.8)	166.7 (35.18)	167.5 (31.09)	166.7 (32.91)	188.4 (33.22)
Hartman6	131.1 (11.4)	166.4 (12.18)	141.2 (13.55)	133.3 (12.07)	132.8 (13.67)	169.6 (15.71)
SC	21.6 (3.46)	39.8 (4.11)	26.7 (2.19)	23.9 (2.64)	22.4 (3.01)	39.6 (3.52)
Shekel7	144.3 (22.5)	187.7 (25.23)	158.4 (29.45)	147.6 (23.11)	146.4 (26.71)	189.3 (23.21)
Shekel10	137.5 (9.62)	192.3 (10.74)	145.8 (9.75)	141.6 (10.44)	145.9 (11.09)	201.8 (12.73)
Shubert	124.5 (11.3)	177.6 (12.08)	146.1 (11.70)	129.1 (13.88)	127.6 (10.76)	184.4 (11.03)
Trid10	114.7 (13.5)	118.8 (12.77)	115.3 (13.8)	119.7 (13.01)	121.3 (13.69)	132.3 (15.01)
Sphere10	107.6 (9.85)	106.8 (10.18)	105.2 (12.33)	106.4 (9.24)	110.4 (8.35)	122.6 (8.11)
F16	135.2 (8.57)	147.5 (10.12)	134.7 (8.44)	133.5 (9.69)	139.6 (7.75)	157.1 (11.24)

in phase 2. After that, the global exploration started again to search the global optima, which was demonstrated as phase 3. During this process, the improvement of the EI criterion was very little. To speed up the convergence of optimization, the objective function was transited into the MP criterion, and the global optima was soon found, shown as phase 4.

The update of sample points is shown in Figure 10. Figure 10(a) presents the global exploration based on the EI criterion. Then, three partially explored local minima were selected to make further exploitation by multiple subregion local search strategy. To perform better exploitation, the ETRM algorithm was adopted to restrict optimization in a small region with dynamic bounds, where the red-dashed rectangular were their initial design spaces, namely, as shown in Figure 10(b). The exploitation result is presented in Figure 10(c), in which the local minima disappear and the surrogate model approximated the objective function well. To detect new partially explored local minima, the global search continued and several points were updated. In Figure 10(d), all local minima were fully explored. And in terms of the value of  $C_{EI}$ , the objective function was turned into the MP criterion and it only took 3 evaluations to locate the global optima. At last, the global optima would be found at the 58<sup>th</sup> iteration.

The KSE method could enhance and balance the exploitation and exploration ability of optimization by locating potential local minima and combining multiple objective functions, improving the optimization efficiency and preventing premature optimization.

## 5. Test Functions

*5.1. Experimental Setup.* The performance of KSE was studied in the sense of global convergence capability, efficiency, and robustness on a broad set of global optimization test problems with 2 to 16 design variables. These included the seven Dixon [51] test functions (Branin, Goldstein-Price, Hartman3, Shekel5, Shekel7, Shekel10, and Hartman6), two test functions (SC and Shubert) with multiple local minima, and three test functions (Trid10, Sphere10, and F16)

with higher dimension [32]. These test problems had their unique characteristics, and in combination, they could better represent many different situations in design problems. The detailed forms of these functions are given in Table 1. The initial LHS samples with size  $2(d+1)$  were the same for all optimization algorithms on all problems. Thirty runs on each of these benchmark problems have been made.

The experiments of all test cases were performed on Intel(R) Core (TM) i7-8700 CPU @ 3.20 GHz computer. The obtained statistical results were compared with that of several well-known global optimization methods. One alternative was the efficient global optimization method (EGO) [6] as implemented in MATLAB by Viana [52], which was one of the most widely used surrogate-based global optimization methods, especially for engineering design problems. The model-pursuing sampling (MPS) method [53] was an adaptive surrogate model method producing discriminative samples toward the actual global optima. The mesh adaptive direct search (MADS) [54] algorithm was implemented as NOMADm [55], using the surrogate model to make it suitable for computationally expensive problems. A universal multistart approach called multilevel single linkage (MLSL) [56], with local optimization solver sequential quadratic programming (SQP), is implemented by MATLAB optimization toolbox [57]. Another alternative was trust-region-based adaptive response method (TR-ARSM) [49], which was an efficient space reduction method for multimodal problems. Trust-region implementation in Kriging-based optimization with expected improvement with restart strategy (TRIKE-Restart) was a trust-region-like approach where EI function was combined with trust-region and restart strategy [20]. We also compared the proposed method with two multistart methods: the A QUAsi-multistart Response Surface (AQUARS) [18], during which the initial points were obtained from the CGRBF method [29], and the multistart space reduction method (MSSR) [32], which was the combination of multistart approach and space reduction strategy. The parameter settings of these methods are summarized in Table 2. The acquired statistical results of AQUARS, TR-ARSM, TRIKE-Restart, and MSSR were

collected from the published papers due to the lack of publicly available software to implement the competitor methods.

These methods were compared in terms of the average number of function evaluations (over 30 trials) required to get a solution with relative error  $< 1\%$ . If  $f_{\text{target}}$  was the global minima value over the domain  $D$  and  $f_{\text{min}}$  was the best value obtained by an algorithm, then the relative error of the algorithm was given by

$$\text{Relative error} = \begin{cases} \frac{|f_{\text{min}} - f_{\text{target}}|}{|f_{\text{target}}|}, & f_{\text{target}} \neq 0, \\ 0.001, & f_{\text{target}} = 0. \end{cases} \quad (17)$$

**5.2. Results and Discussion.** Table 3 gives the average number of function evaluations required by each method to reach the target value with relative error  $< 1\%$  [18]. If all 30 trials got within 1% of the global optima, then the number inside the parenthesis was the standard error of the mean. The symbol “>” indicated that at least one of the trials could not detect the target value. And  $N_e$  was the number of trials that did not get within 1% of the global minima value within 500 function evaluations.

From Table 3, it was shown that the proposed KSE method performed much better than the EGO method, MPS method, and MLSL-SQP method on all test problems. The EGO method generally required more evaluations to get the target value, especially for the problems with high dimensionality and multimodality. The new sample points of EGO were more likely to be updated around the local minima which were well explored. The MPS method performed a little better than EGO on high-dimensional problems (Trid10 and Sphere10), but it was not efficient for complicated problems with higher multimodality (Shekel problems and Shubert function), since it was hard for MPS to detect the promising regions from the design space with multiple local minima. It was presented that MLSL-SQP was much better than EGO and MPS methods on 8 test problems, except for the difficult Shekel problems with deep and narrow global optima basins. However, the local searches of MLSL-SQP tended to full convergence around all local minima even if these local minima were unlikely to lead to better points. These three methods were well known for engineering design, so these results suggested that the KSE method was very promising for engineering optimization problems.

It showed that KSE could reach the target value faster than 4 alternative methods (TR-ARSM, MSSR, AQUARS, and TRIKE-Restart) on 8 test problems (Branin, SC, Hartman3, Shekel5, Shekel7, Trid10, Sphere10, and F16), and it could be compared with the best results attained by the alternative methods on 2 other problems (Goldstein-Price and Shubert). Although alternatives based on the multistart approach (MSSR and AQUARS) could perform better than the KSE on 3 test problems (Hartman6, Shekel10, and Shubert), it could not conclude that the multistart approach had

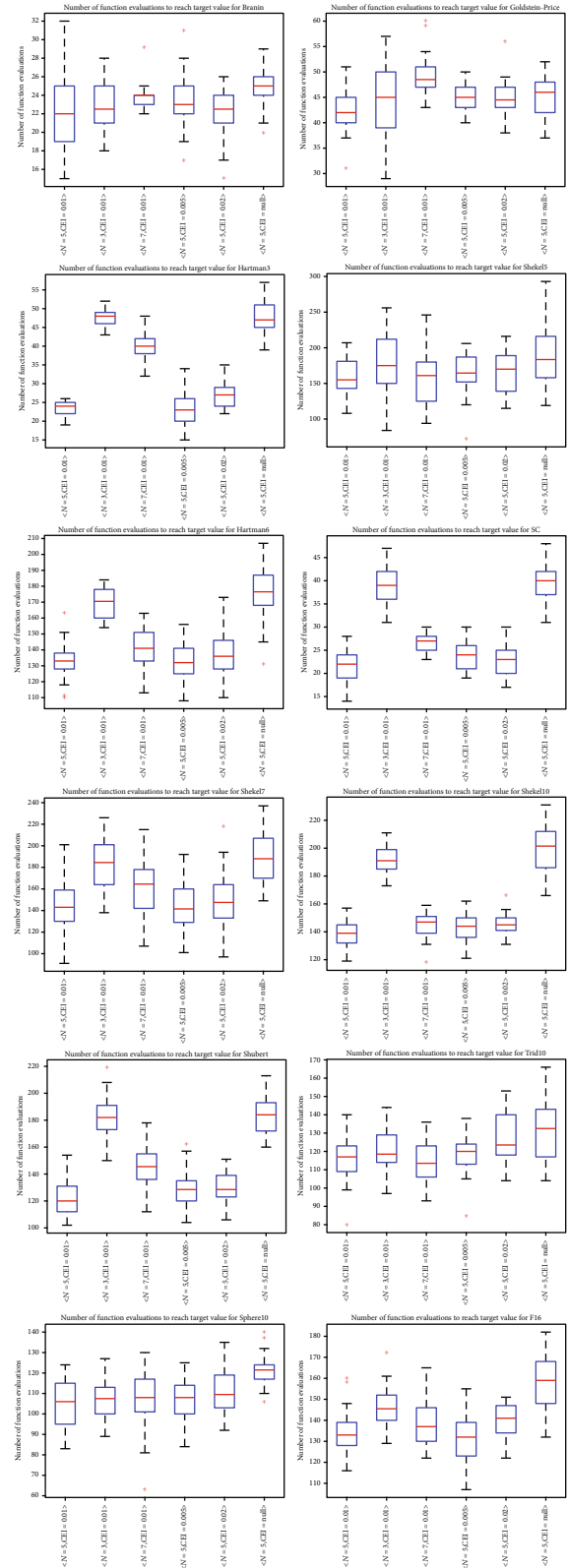


FIGURE 12: Box plot of the number of function evaluations to get within 1% of the target value.

extremely advantages over the KSE, since KSE performed much better than MLSL-SQP on all test problems. It might be that these methods put much emphasis on local search.



TABLE 5: RAE 2822 airfoil optimization problem description.

		Quantity	Lower	Upper	Scaling
Minimize	$C_d$				
Concerning	Shape (coefficients of Bernstein function)	24	-0.05	0.05	1
	Total number of design variables	24			
Subject to	$C_l$	1	0.824	0.824	1
	$C_m$	1	-0.092		10
	Area constraint	1	Area <sub>initial</sub>		10
	Total number of design variables	3			

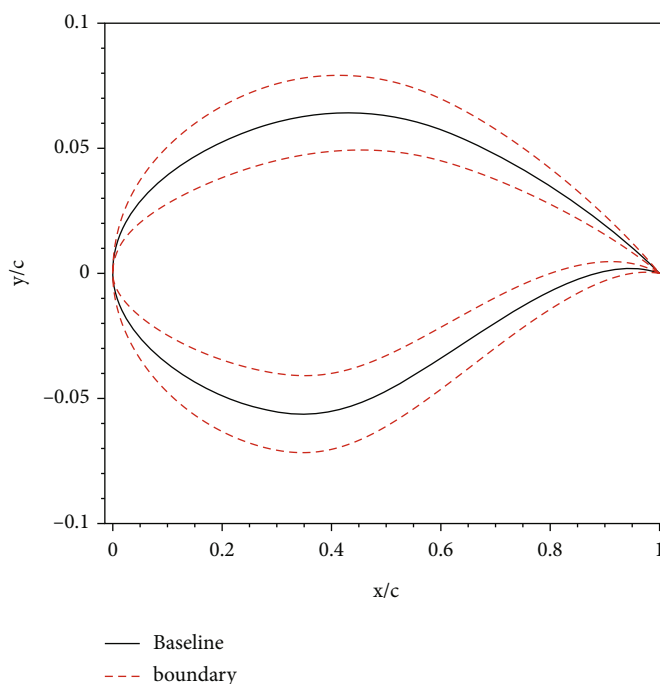


FIGURE 13: The initial design space.

Besides, the different initial LHS methods and surrogate models in the compared papers could also influence the results. Anyway, the KSE was quite robust in the test cases mentioned in Table 3. And it has shown its advantages in stability and efficiency to reach the target value in all trials on several test problems.

Figure 11 shows the box plots for the number of function evaluations required to get within 1% of the target value for the KSE and some other alternatives on all test problems. Many trials of EGO and MPS could not achieve the desired target value on many test problems, and it was impossible to create accurate box plots for these methods. Hence, they were excluded from the box plots. The box plots showed that MADS was competitive with the KSE on Shekel and Shubert problems and MLSL-SQP was much worse than the KSE on all test problems in terms of the median number of function evaluations to get within 1% of the global minima value. The distributions of the number of function evaluations were much less spread out for the KSE compared to that of other

methods on most test problems. Hence, the KSE was generally more robust and efficient than other methods.

Rather than demonstrating the superiority of the proposed algorithm over other existing algorithms, the objective of the study was to show that the proposed algorithm could be a viable alternative for computationally expensive black-box global optimization, which has been shown on a wide variety of benchmark test problems.

**5.3. Sensitivity Analysis.** In this section, the KSE was tested on various cluster numbers and function transition factors. Table 4 shows the results obtained by the test cases of applying KSE with different hyperparameters including (1)  $N = 5$ ,  $C_{EI} = 0.01$ , (2)  $N = 3$ ,  $C_{EI} = 0.01$ , (3)  $N = 7$ ,  $C_{EI} = 0.01$ , (4)  $N = 5$ ,  $C_{EI} = 0.005$ , (5)  $N = 5$ ,  $C_{EI} = 0.02$ , and (6)  $N = 5$ ,  $C_{EI} = \text{null}$ . Case (1)~case (3) were compared to study the influence of cluster number. The comparisons of the other cases as well as case (1) were to research the influence of the function transition factor, during which case (6) presented the

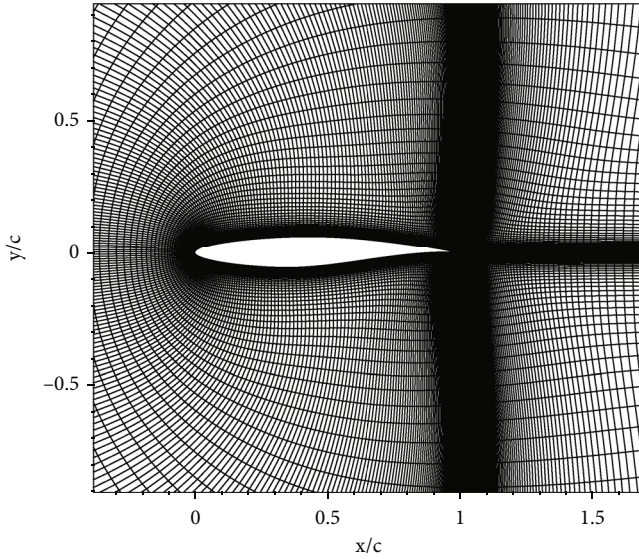


FIGURE 14: The computational grid of RAE 2822.

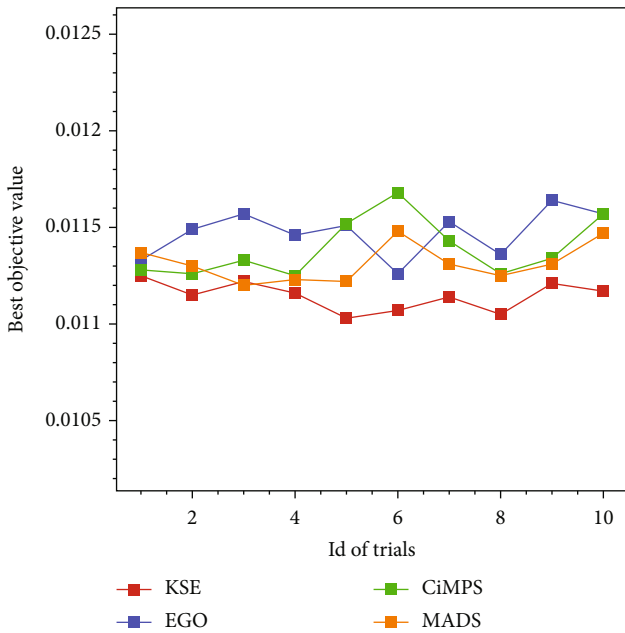


FIGURE 15: The results of ten trials of each method.

results without function transition, indicating that the objective function was always the EI criterion. The initial samples with the size of  $2(d + 1)$  were the same for all cases. Each test problem has been repeated 30 times.

From Table 4, we could know that case (1) and case (2) achieved really good results on Branin and Goldstein-Price problems and case (3) also obtained competitive results on these two test problems, which showed that the proposed method was not that sensitive to the cluster number for problems with low multimodality and dimensionality. In the era of increasing multimodality and dimensionality, a more comprehensive clustering strategy in KSE was needed

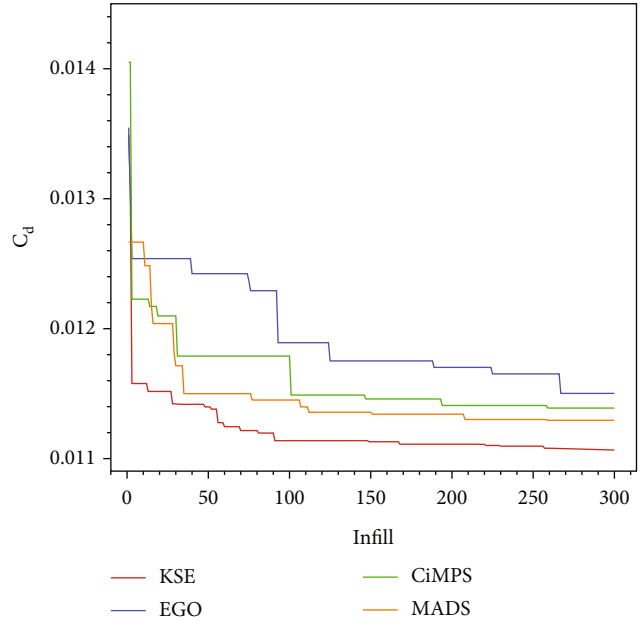


FIGURE 16: Mean convergence of best objective function values in 10 trials.

to identify the promising local minima. It was shown that case (2) obtained the best results on Hartman3, Shekel5, Hartman6, Shekel7, Shekel10, Shubert, and Trid10 test problems and competitive results on Sphere10 and F16 test problems. But case (3) also obtained the competitive results on Shekel5, SC, Shekel7, Shekel10, and Trid10 test problems.

As for the function transition factor  $C_{EI}$ , it seemed that the KSE was not sensitive to the function transition factor  $C_{EI}$  with the appropriate cluster number, like  $N = 5$ . There was little difference among the results obtained by case (1), case (4), and case (5) on all test problems, where the case (1) obtained the best results on 8 test problems, except Branin, Hartman3, Sphere10, and F16 test problems, and competitive results on Branin, Hartman3, and Sphere10 test problems. Despite this, case (6) performed worse on most test problems than the cases with a function transition factor  $C_{EI}$ . This suggested that it was important to incorporate the function transition factor as well.

Figure 12 shows the box plots for the number of function evaluations required to get within 1% of the target value of all cases. It was presented that the results obtained by case (1) were better than other cases on 8 of the 12 test problems (all except Branin, Shekel5, Trid10, and Sphere10). And the distribution of function evaluations of the case (1) was more concentrated than that of other cases, although the results of the case (4) and case (5) could be competitive on all test problems, indicating that they were not very sensitive within the current interval.

In summary, these results suggested that cluster number and function transition factor were all necessary for good performance on most test problems, and the KSE was not very sensitive to function transition factor with suitable cluster number.

## 6. Test on Benchmark Problems

This section applied the proposed KSE method to two aerodynamic optimization problems. To validate the proposed

TABLE 6: The best result comparison of different methods.

	$C_l$	$C_d$ (counts)	$C_m$	The number of objective function evaluation	CPU time (min)
Baseline	0.824	203.1	-0.0927		N/A
KSE	0.824	110.5	-0.0919	100 + 300	112.02
EGO	0.824	112.6	-0.0915	100 + 300	117.11
CiMPS	0.824	112.8	-0.0910	100 + 300	110.68
MADS	0.824	112.1	-0.0908	100 + 300	105.59

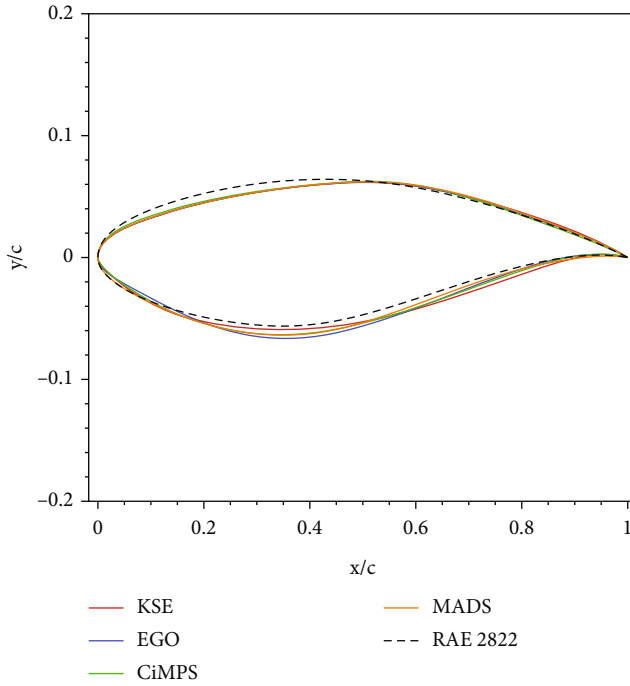


FIGURE 17: The comparison of optimized airfoil shape.

algorithm's performance, three other global optimization methods were employed for comparison. One was MSL coupled with mesh adaptive direct search (MADS) [54] algorithm, which was implemented as NOMADm [55], using the surrogate model to make it suitable for computationally expensive problems. Another one was the constraint-importance model-pursuing sampling (CiMPS), which was an extension of the MPS method developed for handling expensive constraints by Kazemi et al. [58]. The last method was the EGO method, one of the most well-known global optimization methods for engineering problems. The default or recommended tuning parameters of these methods were used. Ten runs on each problem have been made by all methods.

In this section, a high-fidelity CFD solver, CFL3D [59], was adopted to evaluate the aerodynamic characteristics of airfoils and wings. The Navier-Stokes function incorporated with shear stress transport (SST) two-equation turbulence model was used to simulate the flow field of airfoils and wings. To perform the aerodynamic design optimization, a more powerful workstation equipped with AMD Epyc 7763 64-core @ 2.45 GHz CPUs and 512 G RAM was used.

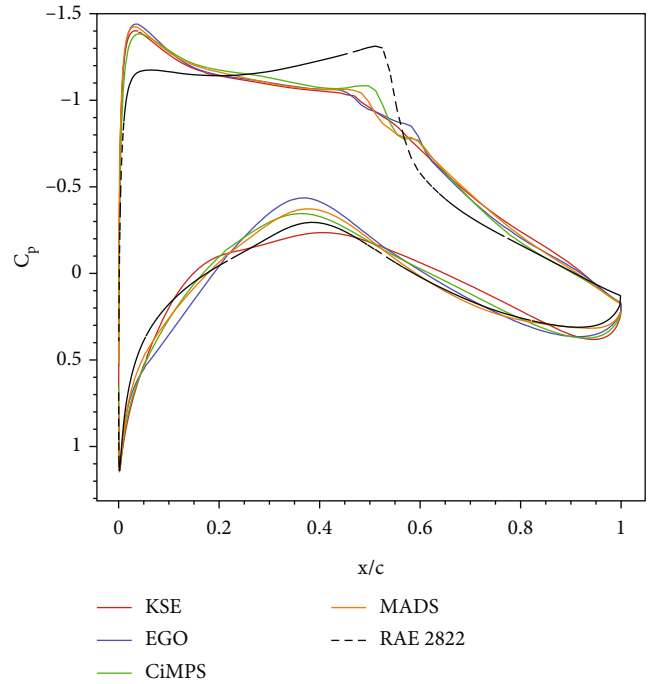


FIGURE 18: The comparison of pressure coefficient distribution.

**6.1. Case I. Transonic Airfoil Design.** Case I was provided by the AIAA aerodynamic design optimization discussion group (ADODG: the drag minimization of the RAE 2822 airfoil in transonic viscous flow). The Mach number was 0.734, and the Reynolds number was  $6.5e6$ . The RAE 2882 airfoil was optimized with a fixed life coefficient of 0.824. The pitching moment should be bigger than -0.092, and its area should not be less than that of RAE 2822. The statements of problem are summarized in Table 5. More detail on this case could be referred to in [60, 61]. The problem could be expressed as a standard nonlinear optimization problem:

$$\begin{aligned}
 & \text{Min } C_d, \\
 & \text{s.t. } C_l = 0.824, \\
 & C_m \geq -0.092, \\
 & \text{Area} \geq \text{Area}_{\text{initial}}.
 \end{aligned} \tag{18}$$

It was a standard airfoil optimization with medium multimodality [62], which could be used to test the proposed method. The baseline shape was the RAE 2822 airfoil. The design variables were the  $z$ -coordinates of the CST parameterization with 24 shape parameters in total. The design space is shown in Figure 13, and the computational grid is shown in Figure 14. Generate 100 sample points in the initial design space with the LHS method, and update 300 sample points during optimization progress.

In this case, the EGO method tried to explore everywhere of design space, leading to bad optimization convergence; the CiMPS method was very efficient in the early stage of optimization, but at early stage, the surrogate model was not stable, and it was easily trapped in local minima, leading to optimization

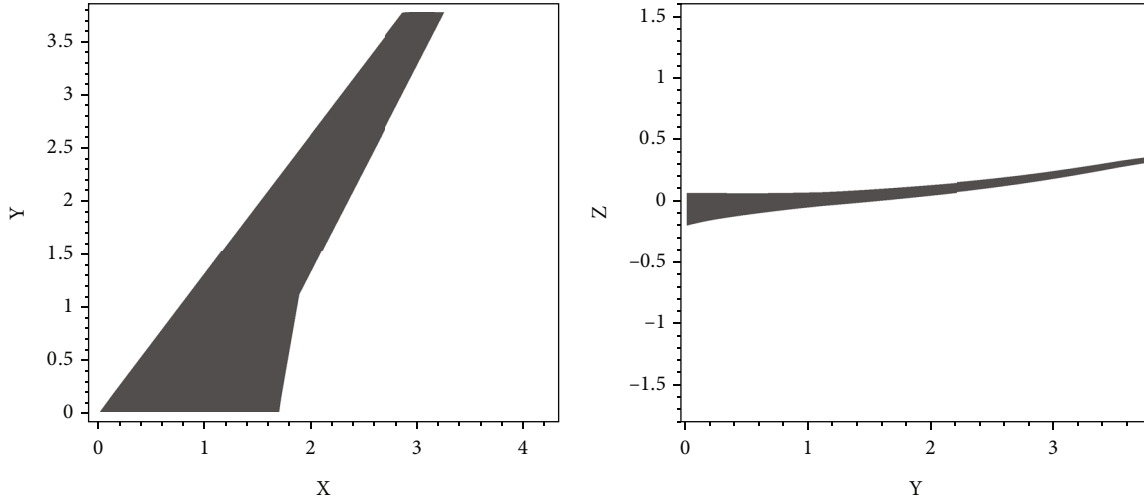


FIGURE 19: Baseline CRM wing scaled by its mean aerodynamic chord.

TABLE 7: CRM wing optimization problem description.

		Quantity	Lower	Upper	Scaling
Minimize	$C_D$				
Concerning	Twist angle	5	$-4^\circ$	$4^\circ$	1
	Sectional shape ( $z$ -coordinates of the FFD parameterization)	72	-0.05	0.05	1
	Total number of design variables	77			
Subject to	$C_L$	1	0.5	0.5	1
	$C_M$	1	-0.17		10
	Volume constraint	1	$V_{\text{baseline}}$		1000
	Thickness constraint	500	$0.25t_{\text{baseline}}$		1000
	Leading edge position	1			1000
	Trailing edge position	6			1000
	Total number of design constraints	510			

premature in this problem; the starting points could easily lie in the neighborhood of the previously found local minima, causing wasted computational cost in researching these well-explored local minima. Figure 15 presents the design results of ten trials of all methods, and most of the design results by the KSE method were better than other methods, indicating that the KSE method was more stable than the other alternatives. Figure 16 shows that the KSE took 17% iterations of MADS to reach the similar drag coefficient value and it converged to the target shape faster. Thus, the KSE method was more efficient and robust.

The best design result of each method was selected to make a comparison. Table 6 presents the result comparison of different methods which indicated that all the constraints were satisfied. Under the same limited objective function evaluations, the computational cost of the MADS was the cheapest, but the KSE obtained better design results, compared to other optimization methods. And the  $C_d$  of design result by the KSE was around 110 cts, which was similar to the result in reference [62]; therefore, it was very likely close to the global optima. Figures 17 and 18 show the comparison

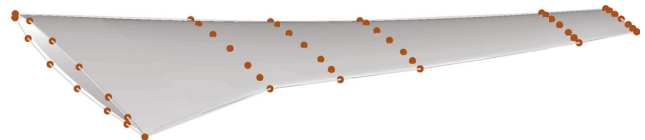
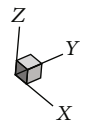


FIGURE 20: Shape design variables of FFD.

of the optimized airfoil shapes and pressure coefficient distribution. As we could see, the shape trend of all optimized airfoils was the same: the locations of maximal thickness were moved backward, compared to the baseline, and the radiuses of their leading edges were reduced. The upper

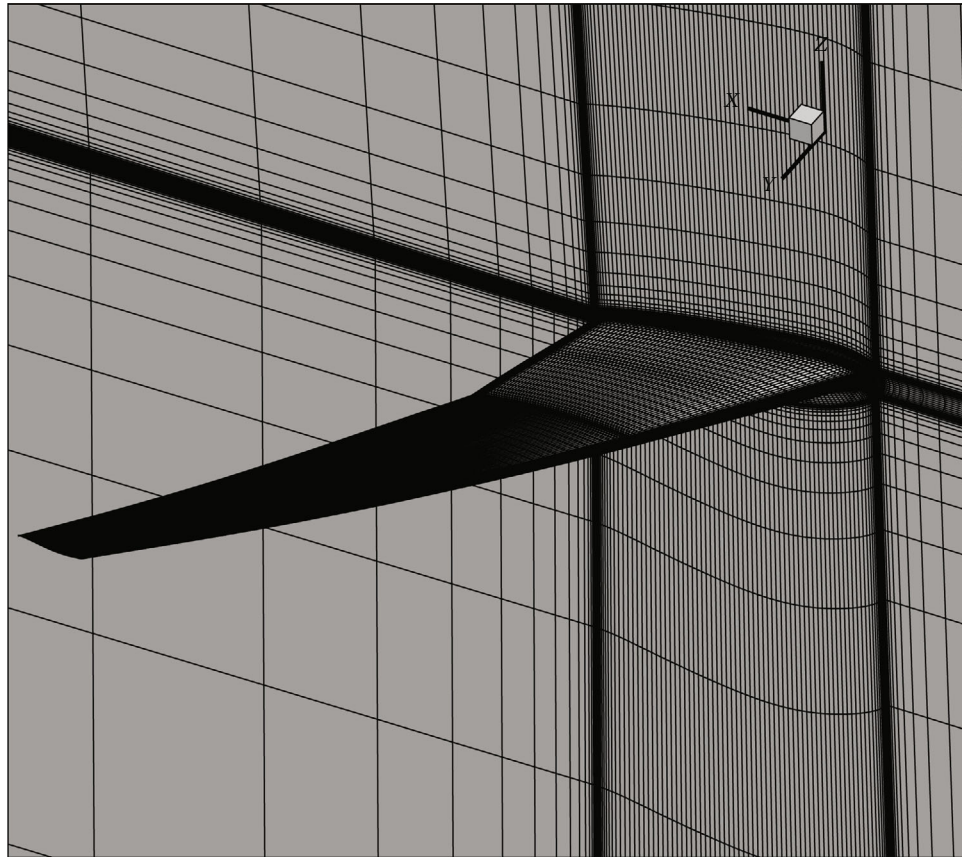


FIGURE 21: The computational grid of the CRM wing.

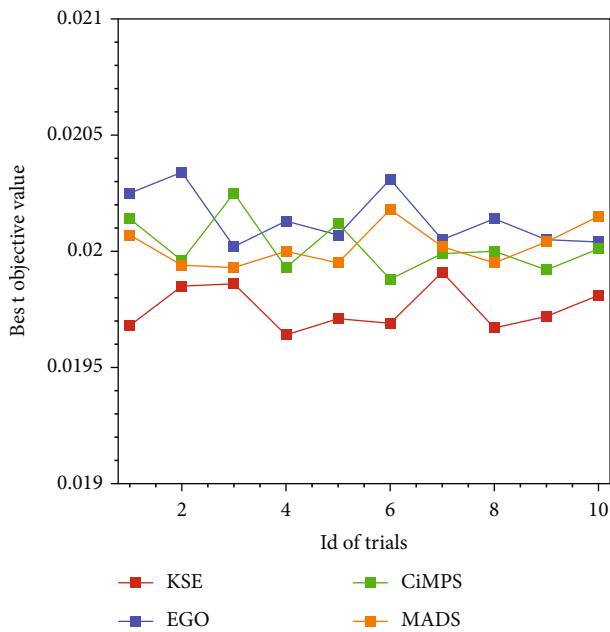


FIGURE 22: The results of ten trials of each method.

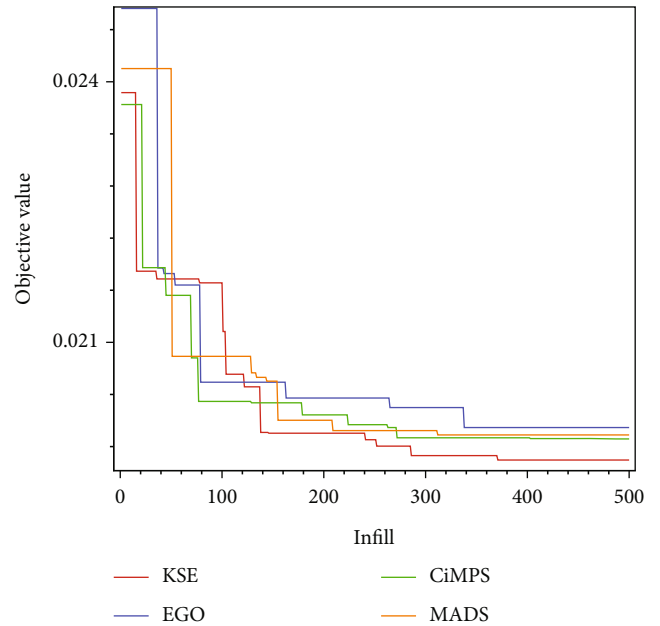


FIGURE 23: Mean convergence of best objective function values in 10 trials.

surfaces of optimized airfoils were very similar, and the difference mainly occurred on the lower surface, which was the main source of local minima. From Figure 18, we could know that the suction peaks of all optimized airfoils got

much stronger and the shock waves have been weakened compared to baseline. However, the KSE nearly eliminated the shock wave while other methods remained an obvious shock wave, leading to bad aerodynamic performance.



TABLE 8: The best result comparison of different methods.

	$C_L$	$C_D$ (counts)	$C_M$	The number of objective function evaluation	CPU time (h)
Baseline	0.50	210.8	-0.182		N/A
KSE	0.50	196.4	-0.169	200 + 500	106.30
EGO	0.50	200.2	-0.165	200 + 500	107.66
CiMPS	0.50	198.8	-0.166	200 + 500	107.53
MADS	0.50	199.3	-0.166	200 + 500	106.26

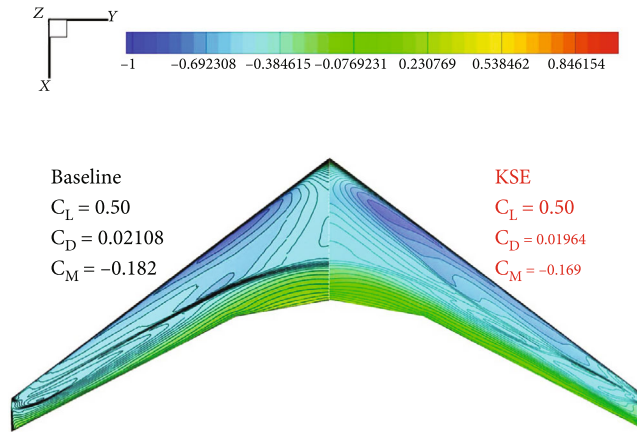


FIGURE 24: Pressure distributions of baseline and the KSE.

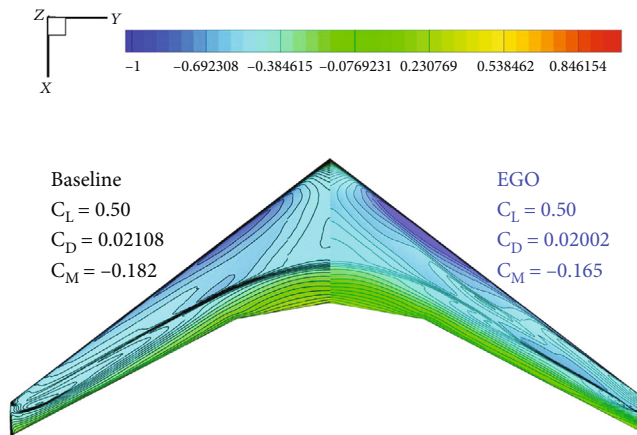


FIGURE 25: Pressure distributions of baseline and the EGO.

6.2. Case II. CRM Wing Optimization. The baseline of Case II was a wing with a blunt trailing edge selected from the CRM wing-body geometry [63], which was representative of a contemporary transonic commercial transport, similar to the Boeing 777. The geometry and specifications were also given by the ADODG, while the fuselage and tail were removed from the original CRM, and the root of the remaining wing was moved to the symmetry plane. This baseline geometry is shown in Figure 19. All coordinates were scaled by the mean aerodynamic chord (275.8 in). The resulting reference chord was 1.0, and the reference area was 3.407014. The moment reference point was located at  $(x, y, z) = (1.2077, 0.0, 0.007669)$ .

The problem statements are shown in Table 7. The objective was to minimize the drag, subject to a lift constraint ( $C_L = 0.5$ ) and a pitch moment constraint ( $C_{M_y} \geq -0.17$ ). The shape design variables were the  $z$ -coordinate movement of 72 control points on the FFD volume and 5 twist angles, which are shown in Figure 20. The control points at the trailing edge were constrained to keep the trailing edge unchanged. And the leading edge control points at the wing root were also constrained to maintain a constant incidence for the root section. There were 500 thickness constraints imposed in a 25 chord-wise and 20 spanwise grid covering the full span. Each thickness constraint was set to be no smaller than 25% of the baseline thickness at all locations. Finally, the internal volume was

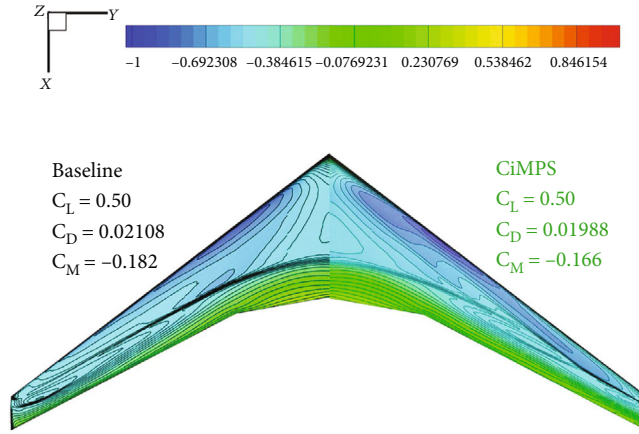


FIGURE 26: Pressure distributions of baseline and the CiMPS.

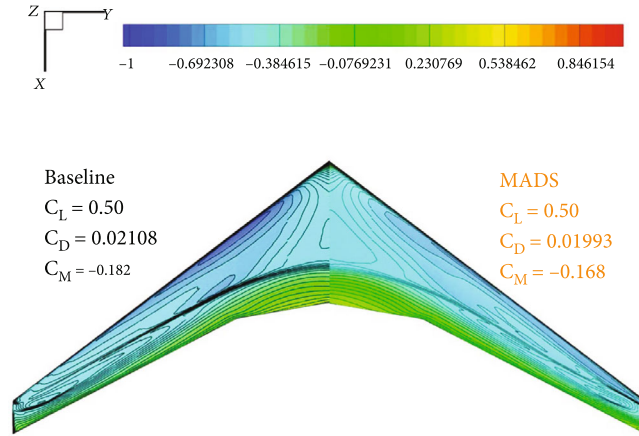


FIGURE 27: Pressure distributions of baseline and the MADS.

constrained to be greater than or equal to the baseline volume. The optimization problem was shown below.

$$\begin{aligned}
 & \text{Min } C_D, \\
 & \text{s.t. } C_L = 0.5, \\
 & C_{M_y} \geq -0.17, \\
 & t \geq 0.25t_{\text{baseline}}, \\
 & V \geq V_{\text{baseline}}, \\
 & \Delta z_{\text{TE,upper}} = -\Delta z_{\text{TE,lower}}, \\
 & \Delta z_{\text{LE,upper,root}} = -\Delta z_{\text{TE,lower,root}}.
 \end{aligned} \tag{19}$$

The computational grid is shown in Figure 21. The mesh was marched out from the surface mesh using an O-grid topology to a far-field located at a distance of 25 times the span. The cruise Mach number was 0.85 with a Reynolds number of 5 million based on the mean aerodynamic chord. The mesh generated for optimization contained 15.5 million cells. It was a complicated 3D aerodynamic optimization problem with relatively high dimensionality and multimodality [64].

200 sample points were generated in the initial design space by the LHS method, and 500 sample points were to be updated as optimization progresses. Figure 22 presents the cost-effectiveness comparison and optimized results of ten trials for each method, indicating that the KSE method generally achieved better results within default computational costs and was more efficient, compared to other methods. Figure 23 shows that the KSE held a better convergence speed to obtain the ideal design results. Table 8 summarizes the results of all methods. At the optima points, the lift coefficient and pitch moment coefficient constraints were all met. The discrepancy in the computational cost of all optimization algorithms was not that large under the same number of objective function evaluations. It was shown that the optima by the KSE were the best of all global optimization methods, and the drag coefficient has been reduced by 14.4 counts compared to the baseline. The drag coefficient of the result obtained by the CiMPS method was similar to that of the MADS method, which was slightly better than that of the EGO method.

Pressure distributions on the platforms of all optimized results are shown in Figures 24–27, during which the

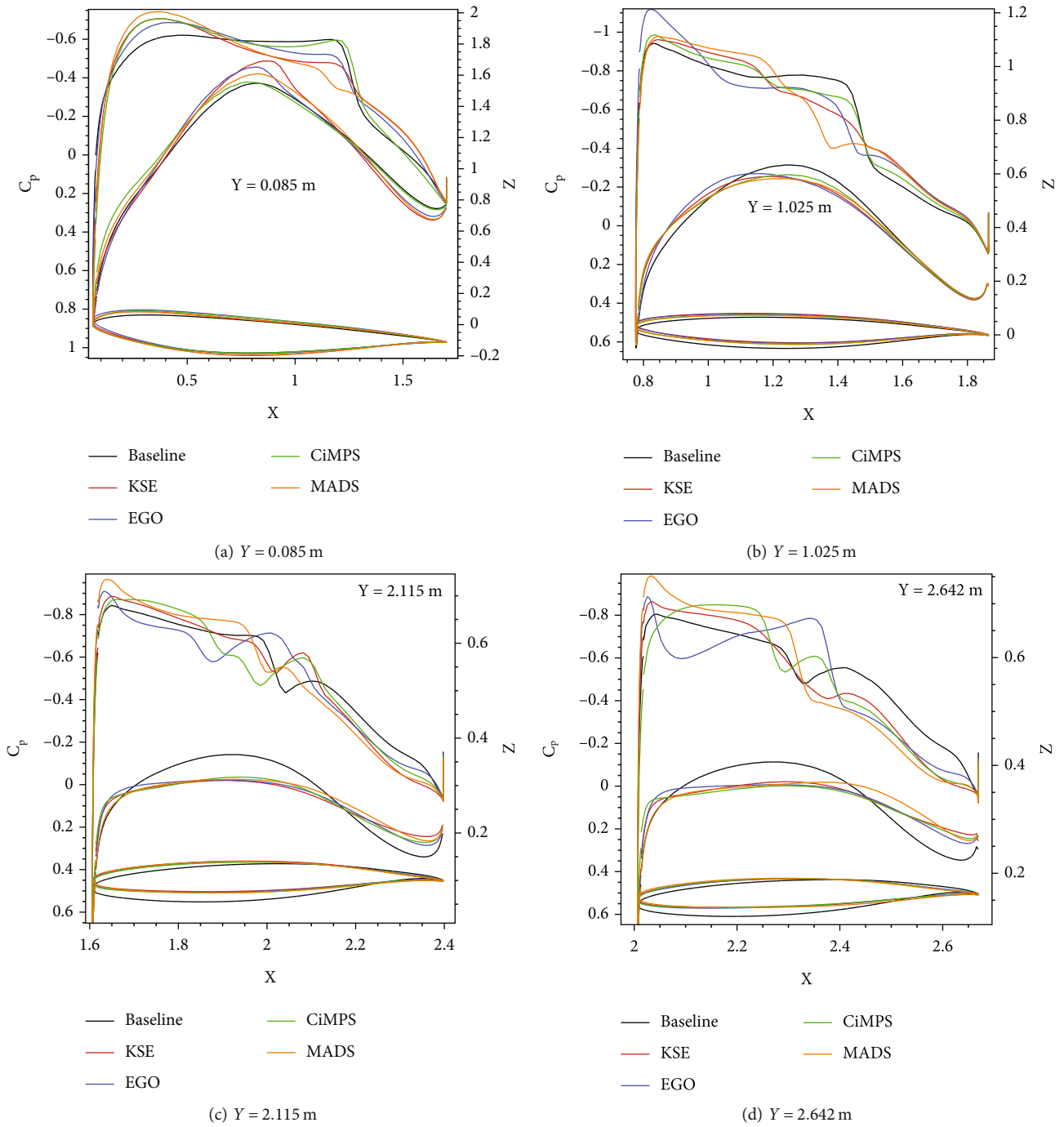


FIGURE 28: Continued.

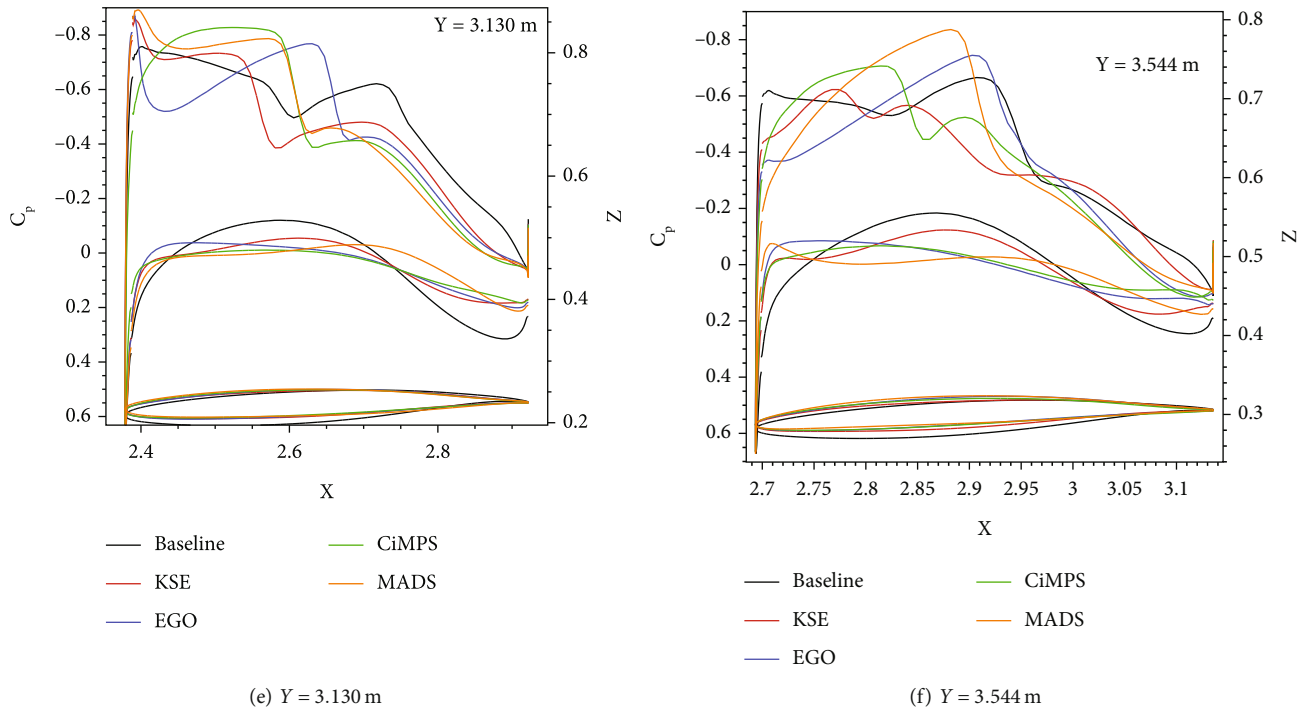


FIGURE 28: Geometry and pressure distributions at a different section of the baseline model and optimized results.

baseline was shown on the left and the optimized results were on the right, respectively. For all the optimized wing results, the lift coefficient targets were met and the pitching moment constraints were satisfied. In the optimized result of the KSE method, the shock was largely reduced, which was achieved by fine-tuning the twist distribution and airfoil shapes. The optimized results of the CiMPS and MADS methods were similar to each other, the low-pressure regions extended, which was beneficial to drag reduction, but there still existed obvious shock across the middle of the wings. As for the optimized result of the EGO method, the shock wave on the inner wing has been eliminated, while on the outer wing, the shock extended to the wingtip.

Pressure distributions on different cross-sections are compared in Figure 28. At  $Y = 0.085$  m, the thickness of the airfoil on the optimized results got thicker while airfoils on other sections get thinner, since the thicknesses were allowed to decrease to 25% of the original thickness and the wave drag was strongly influenced by decreased the thickness. But the volume was constrained to be not smaller than the baseline volume, so the optimizer decreased the thickness of the airfoils on the outer wing while increasing the thickness near the root section. Lyu et al. [63] observed similar trends in their results. Most of the optimized results eliminated the shock at  $Y = 0.085$  m, except for the results obtained by the EGO method. At  $Y = 1.025$  m,  $Y = 2.642$  m, and  $Y = 3.544$  m, the shock has been eliminated on the optimized wing of the KSE method, while for other results, the shocks were still existing. All results shared a similar trend at  $Y = 2.115$  m, and the double weak shock replaced the strong shock. And at  $Y = 3.130$  m, the shocks moved forward without obvious weakness.

## 7. Conclusion

In this paper, a Kriging-based space exploration method (KSE) was proposed for multimodal and computationally expensive black-box problems. This method combined multiple subregion local search strategy with global search to balance and enhance the exploration and exploitation ability. By multiple local search strategy, the exploitations of local minima were enhanced, which was beneficial to promoting the exploration of the entire space and improving the optimization convergence as well as efficiency. During this process, an enhanced trust-region method (ETRM) was adopted to make efficient further exploration of the promising local minima. Hence, this method was expected to perform better than traditional optimization methods.

In the numerical experiments, the KSE was compared with other global optimization methods on 12 analytical functions. The results indicated that the KSE method was robust and it generally took fewer function evaluations to get the global optima. Then, it was demonstrated for aerodynamic shape optimizations of the benchmark RAE 2822 airfoil and CRM wing with high dimensionality. Generally speaking, it was difficult for the original EGO algorithm to obtain the ideal results as a result of the curse of dimensionality. The proposed KSE could overcome the limitation of the original EGO and be more effective and stable in all optimization cases compared to other alternatives.

However, the current work still had some limitations. When dealing with high-dimensional problems, a very large amount of samples would be acquired to fit a high-quality surrogate model, and plenty of time would be cost to train the hyperparameter and construct the Kriging model, which

made KSE inefficient for enormously high-dimensional problems. Besides, KSE might be still trapped in local minima when a large number of local minima made design space appear strong nonlinearity in local regions. Therefore, our future work would focus on training the hyperparameter with fewer samples, and improving the surrogate accuracy in nonlinear local regions, to make KSE more robust with complicated problems.

### Data Availability

Code used to support the findings of this study will be made available on request.

### Conflicts of Interest

The authors declare that they have no conflict of interest.

### Acknowledgments

The authors appreciate the online OPTIP toolkits from G.G. Wang. This work was supported by the Key Laboratory Funding with the reference number 6142201200106.

### References

- [1] O. Chernukhin and D. W. Zingg, "Multimodality and global optimization in aerodynamic design," *AIAA Journal*, vol. 51, no. 6, pp. 1342–1354, 2013.
- [2] S. Shan and G. G. Wang, "Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions," *Structural and Multidisciplinary Optimization*, vol. 41, no. 2, pp. 219–241, 2010.
- [3] A. Forrester, A. Sobester, and A. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*, John Wiley & Sons, 2008.
- [4] S. Akyol and B. Alatas, "Plant intelligence based metaheuristic optimization algorithms," *Artificial Intelligence Review*, vol. 47, no. 4, pp. 417–462, 2017.
- [5] B. Alatas and H. Bingol, "Comparative assessment of light-based intelligent search and optimization algorithms," *Light & Engineering*, vol. 28, no. 6, 2020.
- [6] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global Optimization*, vol. 13, no. 4, pp. 455–492, 1998.
- [7] A. I. J. Forrester and A. J. Keane, "Recent advances in surrogate-based optimization," *Progress in Aerospace Sciences*, vol. 45, no. 1–3, pp. 50–79, 2009.
- [8] S. Koziel, S. Koziel, X.-S. Yang, and Q.-J. Zhang, *Simulation-Driven Design Optimization and Modeling for Microwave Engineering*, Imperial College Press, 2013.
- [9] C. Chevalier and D. Ginsbourger, "Fast computation of the multi-points expected improvement with applications in batch selection," in *Lecture Notes in Computer Science*, pp. 59–69, Springer, Berlin, Heidelberg, 2013.
- [10] P. I. Frazier, "A tutorial on Bayesian optimization," 2018, <https://arxiv.org/abs/1807.02811>.
- [11] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer School On Machine Learning*, pp. 63–71, Springer, 2004.
- [12] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: a review of Bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.
- [13] R. Jin, X. Du, and W. Chen, "The use of metamodeling techniques for optimization under uncertainty," *Structural and Multidisciplinary Optimization*, vol. 25, no. 2, pp. 99–116, 2003.
- [14] J. P. Kleijnen and W. C. Van Beers, "Robustness of Kriging when interpolating in random simulation with heterogeneous variances: some experiments," *European Journal of Operational Research*, vol. 165, no. 3, pp. 826–834, 2005.
- [15] A. Parnianifard, A. Azfanizam, M. Ariffin, and M. I. S. Ismail, "Comparative study of metamodeling and sampling design for expensive and semi-expensive simulation models under uncertainty," *Simulation*, vol. 96, no. 1, pp. 89–110, 2020.
- [16] A. Parnianifard, A. Azfanizam, M. Ariffin, and M. I. S. Ismail, "Crossing weighted uncertainty scenarios assisted distribution-free metamodel-based robust simulation optimization," *Engineering with Computers*, vol. 36, no. 1, pp. 139–150, 2020.
- [17] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker, "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, vol. 41, no. 1, pp. 1–28, 2005.
- [18] R. G. Regis and C. A. Shoemaker, "A quasi-multistart framework for global optimization of expensive functions using response surface models," *Journal of Global Optimization*, vol. 56, no. 4, pp. 1719–1753, 2013.
- [19] Z.-H. Han, Y. Zhang, C.-X. Song, and K.-S. Zhang, "Weighted gradient-enhanced Kriging for high-dimensional surrogate modeling and design optimization," *AIAA Journal*, vol. 55, no. 12, pp. 4330–4346, 2017.
- [20] R. G. Regis, "Trust regions in Kriging-based optimization with expected improvement," *Engineering Optimization*, vol. 48, no. 6, pp. 1037–1059, 2016.
- [21] D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *Journal of Global Optimization*, vol. 21, no. 4, pp. 345–383, 2001.
- [22] R. T. Haftka, D. Villanueva, and A. Chaudhuri, "Parallel surrogate-assisted global optimization with expensive functions—a survey," *Structural and Multidisciplinary Optimization*, vol. 54, no. 1, pp. 3–13, 2016.
- [23] G. Sun, Y. Tian, R. Wang, J. Fang, and Q. Li, "Parallelized multiobjective efficient global optimization algorithm and its applications," *Structural and Multidisciplinary Optimization*, vol. 61, no. 2, pp. 763–786, 2020.
- [24] Y. Li, S. Wang, and Y. Wu, "Kriging-based unconstrained global optimization through multi-point sampling," *Engineering Optimization*, vol. 52, no. 6, pp. 1082–1095, 2020.
- [25] J. Xing, Y. Luo, and Z. Gao, "A global optimization strategy based on the Kriging surrogate model and parallel computing," *Structural and Multidisciplinary Optimization*, vol. 62, no. 1, 2020.
- [26] M. Amine Bouhleb, N. Bartoli, R. G. Regis, A. Otsmane, and J. Morlier, "Efficient global optimization for high-dimensional constrained problems by using the Kriging models combined with the partial least squares method," *Engineering Optimization*, vol. 50, no. 12, pp. 2038–2053, 2018.



- [27] G. G. Wang and T. Simpson, "Fuzzy clustering based hierarchical metamodeling for design space reduction and optimization," *Engineering Optimization*, vol. 36, no. 3, pp. 313–335, 2004.
- [28] R. G. Regis and C. A. Shoemaker, "A stochastic radial basis function method for the global optimization of expensive functions," *INFORMS Journal on Computing*, vol. 19, no. 4, pp. 497–509, 2007.
- [29] R. G. Regis and C. A. Shoemaker, "Improved strategies for radial basis function methods for global optimization," *Journal of Global Optimization*, vol. 37, no. 1, pp. 113–135, 2007.
- [30] D. Villanueva, R. Le Riche, G. Picard, and R. Haftka, "Surrogate-based agents for constrained optimization," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, Honolulu, Hawaii, 2012.
- [31] T. Long, D. Wu, X. Guo, G. G. Wang, and L. Liu, "Efficient adaptive response surface method using intelligent space exploration strategy," *Structural and Multidisciplinary Optimization*, vol. 51, no. 6, pp. 1335–1362, 2015.
- [32] H. Dong, B. Song, Z. Dong, and P. Wang, "Multi-start space reduction (MSSR) surrogate-based global optimization method," *Structural and Multidisciplinary Optimization*, vol. 54, no. 4, pp. 907–926, 2016.
- [33] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek, "Scalable global optimization via local Bayesian optimization," *Advances in Neural Information Processing Systems*, vol. 32, pp. 5496–5507, 2019.
- [34] N. Bartoli, T. Lefebvre, S. Dubreuil et al., "Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design," *Aerospace Science and Technology*, vol. 90, pp. 85–102, 2019.
- [35] Y. Diouane, V. Picheny, R. L. Riche, and A. S. Di Perrotolo, "TREGO: a trust-region framework for efficient global optimization," 2021, <https://arxiv.org/abs/2101.06808>.
- [36] H. Akbari and A. Kazerooni, "KASRA: a Kriging-based adaptive space reduction algorithm for global optimization of computationally expensive black-box constrained problems," *Applied Soft Computing*, vol. 90, article 106154, 2020.
- [37] R. Priem, N. Bartoli, Y. Diouane, and A. Sgueglia, "Upper trust bound feasibility criterion for mixed constrained Bayesian optimization with application to aircraft design," *Aerospace Science and Technology*, vol. 105, article 105980, 2020.
- [38] H. Dong, B. Song, Z. Dong, and P. Wang, "SCGOSR: surrogate-based constrained global optimization using space reduction," *Applied Soft Computing*, vol. 65, pp. 462–477, 2018.
- [39] G. Chen, K. Zhang, X. Xue et al., "A radial basis function surrogate model assisted evolutionary algorithm for high-dimensional expensive optimization problems," *Applied Soft Computing*, vol. 116, article 108353, 2022.
- [40] A. Parnianifard, R. Chanchaoren, G. Phanomchoeng, and L. Wuttisittikulkiij, "A new approach for low-dimensional constrained engineering design optimization using design and analysis of simulation experiments," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, p. 1663, 2020.
- [41] F. Li, L. Gao, and W. Shen, "Surrogate-assisted multi-objective evolutionary optimization with Pareto front model-based local search method," *IEEE Transactions on Cybernetics*, pp. 1–14, 2022.
- [42] D. G. Krige, "A statistical approach to some basic mine valuation problems on the Witwatersrand," *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, no. 6, pp. 119–139, 1951.
- [43] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, no. 8, pp. 1246–1266, 1963.
- [44] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp. 409–423, 1989.
- [45] M. J. Sasena, P. Papalambros, and P. Goovaerts, "Exploration of metamodeling sampling criteria for constrained global optimization," *Engineering Optimization*, vol. 34, no. 3, pp. 263–278, 2002.
- [46] M. Locatelli, "Bayesian algorithms for one-dimensional global optimization," *Journal of Global Optimization*, vol. 10, no. 1, pp. 57–76, 1997.
- [47] Y. Li, J. Shi, H. Cen, J. Shen, and Y. Chao, "A Kriging-based adaptive global optimization method with generalized expected improvement and its application in numerical simulation and crop evapotranspiration," *Agricultural Water Management*, vol. 245, article 106623, 2021.
- [48] M. J. D. Powell, "On the global convergence of trust region algorithms for unconstrained minimization," *Mathematical Programming*, vol. 29, no. 3, pp. 297–303, 1984.
- [49] T. Long, X. Li, R. Shi, J. Liu, X. Guo, and L. Liu, "Gradient-free trust-region-based adaptive response surface method for expensive aircraft optimization," *AIAA Journal*, vol. 56, no. 2, pp. 862–873, 2018.
- [50] R. Jin, W. Chen, and A. Sudjianto, "An efficient algorithm for constructing optimal design of computer experiments," in *ASME 2003 International design engineering technical conferences and computers and information in engineering conference. American Society of Mechanical Engineers Digital Collection*, pp. 545–554, Chicago, Illinois, USA, 2008.
- [51] L. C. W. Dixon, "The global optimization problem. An introduction," *Toward global optimization*, vol. 2, pp. 1–15, 1978.
- [52] F. A. C. Viana, "SURROGATES toolbox user's guide, Version 2.1," 2010, <http://sites.google.com/site/felipeacviana/surrogatestoolbox>.
- [53] L. Wang, S. Shan, and G. G. Wang, "Mode-pursuing sampling method for global optimization on expensive black-box functions," *Engineering Optimization*, vol. 36, no. 4, pp. 419–438, 2004.
- [54] C. Audet and J. E. Dennis, "Mesh adaptive direct search algorithms for constrained optimization," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 188–217, 2006.
- [55] S. L. Digabel, "Algorithm 909: NOMAD: nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software*, vol. 37, no. 4, 2011.
- [56] A. H. G. Rinnooy Kan and G. T. Timmer, "Stochastic global optimization methods part II: multi level methods," *Mathematical Programming*, vol. 39, no. 1, pp. 57–78, 1987.
- [57] T. Mathworks, *Optimization Toolbox 4 User's Guide*, MathWorks Inc, 2010.
- [58] M. Kazemi, G. G. Wang, S. Rahnamayan, and K. Gupta, "Constraint importance mode pursuing sampling for continuous global optimization," in *volume 1: 36th design automation conference, Parts A and B*, pp. 325–334, Montreal, Quebec, Canada, 2010.
- [59] "NASA CFL3D Version 6 Home Page," [https://nasa.github.io/CFL3D/Cfl3dv6/cfl3dv6\\_home.html](https://nasa.github.io/CFL3D/Cfl3dv6/cfl3dv6_home.html).

- [60] G. R. Anderson, M. Nemec, and M. J. Aftosmis, "Aerodynamic shape optimization benchmarks with error control and automatic parameterization," in *53rd AIAA Aerospace Sciences Meeting*, Kissimmee, Florida, 2015.
- [61] J. Vassberg, N. Harrison, D. Roman, and A. Jameson, "A systematic study on the impact of dimensionality for a two-dimensional aerodynamic optimization model problem," in *29th AIAA Applied Aerodynamics Conference*, Honolulu, Hawaii, 2011.
- [62] C. Gao, Z. Gao, and S. Li, "Research on airfoil design space and the application of niching techniques in airfoil design," in *55th AIAA Aerospace Sciences Meeting*, Grapevine, Texas, 2017.
- [63] Z. Lyu, G. K. W. Kenway, and J. R. R. A. Martins, "Aerodynamic shape optimization investigations of the common research model wing benchmark," *AIAA Journal*, vol. 53, no. 4, pp. 968–985, 2015.
- [64] G. Streuber, D. W. Zingg, and American Institute of Aeronautics and Astronautics, "A parametric study of multimodality in aerodynamic shape optimization of wings," in *2018 Multidisciplinary Analysis and Optimization Conference*, Atlanta, Georgia, 2018.