



# Variable-fidelity expected improvement method for efficient global optimization of expensive functions

Yu Zhang<sup>1</sup> · Zhong-Hua Han<sup>1</sup> · Ke-Shi Zhang<sup>1</sup>

Received: 12 July 2017 / Revised: 5 November 2017 / Accepted: 8 March 2018 / Published online: 5 May 2018  
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

The efficient global optimization method (EGO) based on kriging surrogate model and expected improvement (EI) has received much attention for optimization of high-fidelity, expensive functions. However, when the standard EI method is directly applied to a variable-fidelity optimization (VFO) introducing assistance from cheap, low-fidelity functions via hierarchical kriging (HK) or cokriging, only high-fidelity samples can be chosen to update the variable-fidelity surrogate model. The theory of infilling low-fidelity samples towards the improvement of high-fidelity function is still a blank area. This article proposes a variable-fidelity EI (VF-EI) method that can adaptively select new samples of both low and high fidelity. Based on the theory of HK model, the EI of the high-fidelity function associated with adding low- and high-fidelity sample points are analytically derived, and the resulting VF-EI is a function of both the design variables  $\mathbf{x}$  and the fidelity level  $l$ . Through maximizing the VF-EI, both the sample location and fidelity level of next numerical evaluation are determined, which in turn drives the optimization converging to the global optimum of high-fidelity function. The proposed VF-EI is verified by six analytical test cases and demonstrated by two engineering problems, including aerodynamic shape optimizations of RAE 2822 airfoil and ONERA M6 wing. The results show that it can remarkably improve the optimization efficiency and compares favorably to the existing methods.

**Keywords** Efficient global optimization · Surrogate model · Variable-fidelity optimization · Expected improvement · Kriging

## 1 Introduction

### 1.1 Background

The efficient global optimization method (EGO) (Jones et al. 1998) based on kriging model and expected improvement (EI) has received much attention and gained a great success in engineering design optimizations (Forrester et al. 2008; Liu et al. 2017; Queipo et al. 2005; Wang and Shan 2007), where high-

fidelity (hi-fi) and expensive numerical simulations are often employed. It generally includes the following steps. First, a limited number of sample points in the design space are chosen by the method of design of experiments (DoE) (Giunta et al. 2003) and an expensive numerical analysis, such as computational fluid dynamics (CFD) or computational solid dynamics (CSD), is executed to get the functional responses; then an initial kriging model is built based on the sampled data. Second, new sample point(s) is selected by maximizing the expected improvement (EI) function calculated through the predicted functional values of the cheap-to-evaluate kriging model and its mean-squared error (MSE). Third, the newly selected sample point(s) is evaluated by the expensive numerical analysis and the functional response(s) is augmented to the sampled data. Fourth, the kriging model is rebuilt and this adaptive sampling process is repeated until the resulting sample-point sequence converges to the global optimum. The EGO method soon got popularity in engineering design after its birth, for it enables the searching of global optimum with much less number of expensive function evaluations than the methods such as genetic algorithm (GA) (Holland 1975) and particle swarm optimization algorithm (PSO) (Kennedy and Eberhart 1995).

---

✉ Zhong-Hua Han  
hanzh@nwpu.edu.cn

Yu Zhang  
momozhang@126.com

Ke-Shi Zhang  
zhangkeshi@nwpu.edu.cn

<sup>1</sup> National Key Laboratory of Science and Technology on Aerodynamic Design and Research, School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, People's Republic of China

The advent of the EGO method greatly inspired the research and development of a type of optimization method called surrogate-based optimization (SBO) (Forrester and Keane 2009; Han and Zhang 2012; Koziel et al. 2013). The SBO method can be viewed as a generalized EGO method, by using different types of surrogate models and infill-sampling criteria (ISC) (Forrester and Keane 2009; Jones 2001; Liu et al. 2012), without being limited to kriging and EI only. Despite the success of SBO, it is still suffering from the “curse of dimensionality” (Koch et al. 1999; Shan and Wang 2010). To tackle this problem, many researchers have devoted themselves to developing more accurate surrogate models, towards more efficient global optimization. Two kinds of methods are concerned. One is gradient-enhanced surrogate model (Han et al. 2017) with the cheap gradients computed by adjoint method, and the other is variable-fidelity model (VFM) (Park et al. 2017) using auxiliary low-fidelity (low-fi), cheaper function to assist the prediction of hi-fi, expensive function. This article is mainly concerned with the VFM method, and the use of cheap gradients to enhance the prediction is beyond the scope of this article.

Generally, the VFM methods can be classified into three categories. The *first* type is the correction-based method that corrects the low-fi model using a bridge function (or scaling function) to approximate the hi-fi function (Alexandrov et al. 1998, 2001). The correction can be multiplicative (Haftka 1991; Chang et al. 1993), additive (Choi et al. 2004, 2009), or hybrid (Gano et al. 2005; Han et al. 2013). The *second* type is the space mapping (Bakr et al. 2001; Leifsson et al. 2016). The design space of the low-fi function is distorted to cause its optimal point to match that of the hi-fi function (Robinson et al. 2008; Viana et al. 2014). The *third* type is variable-fidelity kriging such as cokriging (Kennedy and O’Hagan 2000; Han et al. 2012; Forrester et al. 2007) and HK (Han and Görtz 2012). The cokriging was originally proposed in geostatistics community (Journal and Huijbregts 1978) and then extended to deterministic computer experiments by Kennedy and O’Hagan (2000). Cokriging introduces the assistance of cheap low-fi function by constructing the cross covariance between the low- and hi-fi functions. The newly developed HK model (Han and Görtz 2012) directly takes the kriging of the cheap low-fi function as the model trend of the kriging for the expensive hi-fi function, and the difficulty associated with constructing the cross covariance as that in cokriging is avoided. The HK model turns out to be as simple and robust as the correction-based method and as accurate as cokriging method. Moreover, it provides a more reasonable mean-squared error (MSE) estimation than any of the existing kriging and cokriging methods. Recently, the HK model received much attention in engineering design such as uncertainty analysis in CFD (Palar and Shimoyama 2017), multi-objective optimization (Ha et al. 2014), aerodynamic shape optimization (Zhang et al. 2015), and structural optimization (Courier et al. 2016).

With the development of VFM, there is a trend that the VFO has become more and more attractive for engineering

design problems where numerical simulations of varying fidelity are available. The literatures in the research area, published since 2000, are summarized in Table 1, in chronological order. We found that most of the literatures are dealing with low-dimensional optimization problems (< 10 dimensions) and most of the ISC methods used are essentially for local optimization. To enable global optimization, EI or a combination of MSP (minimizing surrogate prediction) and MSE have been used. When the EI method proposed by Jones et al. (1998) for single-fidelity optimization is directly applied to a variable-fidelity kriging model such as HK or cokriging model, only hi-fi samples can be chosen to update the variable-fidelity surrogate model. To improve the optimization efficiency, there is a strong need for adaptively selecting the samples of both low fidelity and high fidelity. However, currently, the theory and method of infilling low-fi samples for the improvement of hi-fi function of interest are still a blank area.

## 1.2 Objectives

This article is inspired by the motivation of developing an adaptive method that can select the new samples of both low fidelity and high fidelity to update the variable-fidelity surrogate model towards the global optimum of hi-fi function. The objectives of developing such a method are as follows:

- 1) The first objective is to solve the problem associated with how to adaptively infill both low-fi and hi-fi samples. A variable-fidelity surrogate model is generally built through a number of observed low-fi samples and a few observed hi-fi samples (Han and Görtz 2012). After the initial variable-fidelity surrogate model is built, it can be improved not only by adding a hi-fi sample but also by adding a low-fi sample. But the question is how to select the new sample points? One way, the traditional way, is to use a fixed number of low-fi samples during the design, while hi-fi samples are repetitively added based on the ISC. It is, of course, not the most efficient way, since it is always questionable that how many low-fi samples would be sufficient. The other way, and a potentially better way, is that both low-fi and hi-fi samples are adaptively added. This motivates us to develop a theory and method that the infilling of low-fi and hi-fi samples are both towards the improvement of hi-fi function.
- 2) The second objective is to develop a more efficient global optimization method based on numerical simulations of varying fidelity. Literature survey shows that most of the VFO methods reported by current literatures are essentially for local optimization. The reason is that it is not economic or possible to build a sufficiently accurate global surrogate model for a multi-dimensional optimization problem. Therefore, there is a strong need to extend the

**Table 1** Summary of literatures about variable-fidelity surrogate-based optimization

Paper	Variable-fidelity surrogate model	Infill-sampling criterion	Engineering test case	No. of design variables
Alexandrov et al. 2001	First-order model management	Minimizing surrogate prediction (MSP)(local opt.)	Aerodynamic wing design	2
Gano et al. 2005	Correction-based VFM	MSP (local opt.)	Multi-element airfoil design	3
Huang et al. 2006	Cokriging	Augmented EI (global opt.)	Metal-forming process design	2
Forrester et al. 2007	Cokriging	EI (global opt.) (Only hi-fi sample is infilled)	Aerodynamic wing design	4
Robinson et al. 2008	Space mapping	MSP (local opt.)	Aerodynamic wing design	10
Choi et al. 2009	Correction-based VFM	MSP (local opt.)	Aircraft conceptual design	23
Koziol and Lefifsson 2013	Space mapping	MSP (local opt.)	Aerodynamic airfoil design	11
Jo et al. 2016	Cokriging & correction-based VFM	MSP and maximizing RMSE with dynamic fidelity indicator (global opt.)	Multi-element airfoil design	4
Courrier et al. 2016	Correction-based VFM, Cokriging & HK	EI (global opt.) (Only hi-fi sample is infilled)	Shrink disk design	3
Benamara et al. 2016	Gappy-POD VFM	Maximizing the error estimator (local opt.)	Aerodynamic airfoil design	2
Cai et al. 2017	Multi-fidelity radial-basis function	R-square, relative maximum absolute error and RMSE (local opt.)	Axle design in propeller	2
Kim Y et al. 2017	HK	EI (global opt.) (Only hi-fi sample is infilled)	Film hole array design	9

standard EI method, proposed by Jones et al. (1998), to a VFO. The only method that serves this target is the augmented EI method proposed by D Huang and T T Allen et al. (2006) and the improved version in 2014 (Reisenthel and Allen 2014). However, the performance of augmented EI method is dependent on the empirical setting of free parameters. This motivates us to develop a variable-fidelity EI method which is free of empirical parameters and mathematically more general.

## 2 Variable-fidelity optimization based on hierarchical kriging and standard EI

The aim of this study is to solve the following constrained optimization problem based on a hi-fi expensive numerical analysis with assistance of a low-fi but cheaper numerical analysis

$$\begin{aligned}
 & \text{Min. } y(\mathbf{x}) \\
 & \text{w.r.t. } \mathbf{x}_{low} \leq \mathbf{x} \leq \mathbf{x}_{up} \\
 & \text{s.t. } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, N_C \\
 & \text{with assistance of } y_{lf}(\mathbf{x}) \text{ and } g_{lf,i}(\mathbf{x}),
 \end{aligned} \tag{1}$$

where  $y(\mathbf{x})$  and  $g_i(\mathbf{x})$  denote the objective and constraint functions, respectively, which are evaluated by expensive, hi-fi numerical analyses;  $y_{lf}(\mathbf{x})$  and  $g_{lf,i}(\mathbf{x})$  are the low-fi objective and constraint functions, respectively, which are evaluated by cheap numerical analyses;  $N_C$  is the number of constraint functions;  $\mathbf{x}_{up}$  and  $\mathbf{x}_{low}$  are the upper and lower bounds of the design variables  $\mathbf{x}$ , respectively. Note that here we are mainly concerned with the single-objective optimization, and the extension to multi-objective optimization is beyond the scope of this article.

### 2.1 Variable-fidelity optimization framework

The optimization is solved based on HK surrogate models (Han and Görtz 2012) for the objective and constraint functions, which are built through a few expensive hi-fi samples and many cheap low-fi samples. The HK models are repetitively updated by the adaptive addition of expensive hi-fi sample points suggested by maximizing the EI function (Jones et al. 1998), until the optimum is reached. The framework of this HK-based optimization is shown in Fig. 1. The basic steps are as follows:

- Two sets of sample points are generated by using a DoE method and they are evaluated by the low- and hi-fi numerical analyses, respectively.
- Based on the low-fi sample data, a kriging model is built. Then taking it as the model trend, a HK surrogate model is built through the hi-fi samples. The HK model(s) for constraint function(s) is built in a similar manner.
- New sample point(s) is selected by maximizing the constrained EI (CEI) function through a combination of

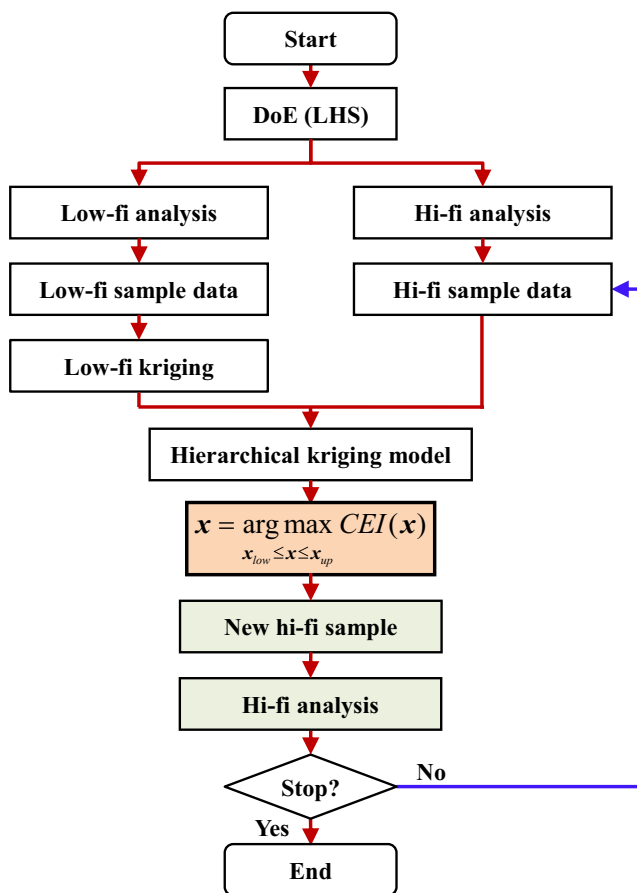


Fig. 1 Flowchart of VFO based on hierarchical kriging model and the standard EI method

- GA, Hooke and Jeeves, and BFGS optimization methods, and then evaluated by the expensive, hi-fi numerical analysis. When BFGS method is used, the gradients of EI function are computed by a central difference based on the HK prediction. Please note that only hi-fi sample points are chosen, and the number of low-fi samples is fixed.
- d. The newly selected sample point and its functional responses are augmented to the sampled dataset, and then the HK surrogate models are rebuilt.
  - e. The steps b–d are repeated until the termination condition is satisfied.

### 2.2 Design of experiments

Before constructing the HK model, DoE method is used to generate initial low- and hi-fi sample points. Different from other VFM methods, such as cokriging model, the hi-fi sampling sites do not need to be a subset of the low-fi sampling sites, which enables us to use DoE method for each fidelity level separately. In this article, the Latin hypercube sampling (LHS) (Giunta et al. 2003) method is adopted.

### 2.3 Surrogate modeling via hierarchical kriging

Kriging is a statistical interpolation method suggested by Krige (1951) and mathematically formulated by Matheron (1963). Kriging gained popularity in design and analysis of deterministic computer experiments after the milestone research work of Sacks et al. (1989). The HK model, proposed by Han and Görtz (2012), is an extension of the Sacks' kriging model to one that approximates the expensive hi-fi function with assistance of a cheap low-fi function. The details about the formulation of HK can be found in Appendix A.1.

### 2.4 Infill-sampling criterion of EI method

After constructing the initial HK model, it can be improved by adding new sample points using the EI method, proposed by Jones et al. (1998). EI is defined as the mathematical expectation of improvement with respect to the best solution observed so far. Through maximizing the EI function, the next update point can be identified. The details about the formulation of EI method can be found in Appendix A.2.

### 2.5 Termination conditions

Several termination criteria can be used in the HK-based optimization, such as the criteria defined based on the distance between samples and difference of their objective responses, the allowable minimum value of  $CEI_{max}$ , the thresholds about accuracy of surrogate models, and the affordable maximum number of expensive function evaluations.

## 3 Proposed method

### 3.1 Formulation of variable-fidelity EI

This article formulates a VF-EI function, which is an extension of the standard EI to consider not only the model uncertainty due to lack of a hi-fi sample but also the model uncertainty due to lack of a low-fi sample. When HK model (Han and Görtz 2012) is used, the uncertainty due to lack of a low-fi sample can be analytically derived, which might be difficult or even not possible for other kinds of models such as cokriging. This uncertainty is used to define the EI of hi-fi function due to lack of a low-fi sample. Combining this EI with the standard EI, the VF-EI is formulated as not only a function of the design variables  $\mathbf{x}$  (or spatial location) but also the function of fidelity level denoted by  $l$ .

Recall that for a HK model (see Appendix A.1 for detailed formulation), the random process of hi-fi function is assumed as

$$Y(\mathbf{x}) = \beta_0 \hat{y}_{lf}(\mathbf{x}) + Z(\mathbf{x}). \tag{2}$$

where  $\beta_0$  is a scaling factor representing how the low-fi kriging  $\hat{y}_{lf}(\mathbf{x})$  matches the hi-fi function. Based on the assumption, the resulting HK predictor and its MSE is

$$\hat{y}(\mathbf{x}) = \underbrace{\beta_0 \hat{y}_{lf}(\mathbf{x})}_{\text{trend term}} + \underbrace{\mathbf{r}^T \mathbf{R}^{-1} (\mathbf{y}_S - \beta_0 \mathbf{F})}_{\text{correction term}}$$

$$\text{MSE}\{\hat{y}(\mathbf{x})\} = s^2(\mathbf{x}) = \sigma^2 \left\{ 1.0 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \left[ \mathbf{r}^T \mathbf{R}^{-1} \mathbf{F} - \hat{y}_{lf}(\mathbf{x}) \right]^2 / (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \right\}, \tag{3}$$

where

$$\mathbf{F} = \left[ \hat{y}_{lf}(\mathbf{x}^{(1)}), \dots, \hat{y}_{lf}(\mathbf{x}^{(n)}) \right]^T \in \mathcal{R}^n. \tag{4}$$

The key issue of VF-EI is how to define the EI of hi-fi function associated with adding a low-fi sample point. As  $\beta_0 \hat{y}_{lf}(\mathbf{x})$  is its model trend, the uncertainty of the HK prediction associated with lack of a low-fi sample at  $\mathbf{x}$  is found to be  $\beta_0^2 s_{lf}^2(\mathbf{x})$ . Note that the uncertainties in  $\mathbf{F}$  are neglected and  $s_{lf}^2(\mathbf{x})$  is the uncertainty of the low-fi kriging model  $\hat{y}_{lf}(\mathbf{x})$ . Then, we extend the model uncertainty of a HK model to  $s^2(\mathbf{x}, l)$ , which is a function

of both the spatial location  $\mathbf{x}$  and fidelity level  $l$

$$s^2(\mathbf{x}, l) = \begin{cases} \beta_0^2 s_{lf}^2(\mathbf{x}) & l = 1 \text{ for low-fi level} \\ s^2(\mathbf{x}) & l = 2 \text{ for hi-fi level} \end{cases}, \tag{5}$$

where  $s^2(\mathbf{x})$  is the uncertainty of the HK model. We can assume that the prediction of the HK model for the objective function at any untried site  $\mathbf{x}$  obeys the normal distribution

$$\hat{Y}(\mathbf{x}, l) \sim \mathcal{N} \left[ \hat{y}(\mathbf{x}), s^2(\mathbf{x}, l) \right], \quad l = 1, 2, \tag{6}$$

whose mean is the prediction of the HK model  $\hat{y}(\mathbf{x})$ .

With the assumption above, we redefine the statistical improvement w.r.t. the best-observed hi-fi objective function  $y_{\min}$  so far as:

$$I_{vf}(\mathbf{x}, l) = \max \left( y_{\min} - \hat{Y}(\mathbf{x}, l), 0 \right), \tag{7}$$

which is also a function of both the spatial location  $\mathbf{x}$  and fidelity level  $l$ . Then the VF-EI can be derived in a similar manner as the standard EI by Jones et al. (1998) does:

$$EI_{vf}(\mathbf{x}, l) = \begin{cases} \left( y_{\min} - \hat{y}(\mathbf{x}) \right) \Phi \left( \frac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x}, l)} \right) + s(\mathbf{x}, l) \phi \left( \frac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x}, l)} \right), & \text{if } s(\mathbf{x}, l) > 0 \\ 0, & \text{if } s(\mathbf{x}, l) = 0 \end{cases}. \tag{8}$$

It is worth to note that the VF-EI only relates to the improvement of the hi-fi function. The VF-EI for the hi-fi level,  $EI_{vf}(\mathbf{x}, l=2)$  indicates the expected improvement of the hi-fi function if a hi-fi sample point is added, which is similar to the standard EI method. But the VF-EI for the low-fi level,  $EI_{vf}(\mathbf{x}, l=1)$ , is the expected improvement of the hi-fi function when a low-fi sample is added. Please note that here we are not concerned with the expected improvement of low-fi function predicted by a low-fi kriging model. Rather, the infillings of both low-fi and hi-fi samples are hi-fi-optimum oriented, which is a unique property of the proposed method.

For a unconstrained optimization, the new sample point  $\mathbf{x}$  and its fidelity level  $l$  can be determined by solving the following sub-optimization problem

$$\mathbf{x}, l = \underset{\mathbf{x}_{low} \leq \mathbf{x} \leq \mathbf{x}_{up}, l=1,2}{\text{argmax}} EI_{vf}(\mathbf{x}, l). \tag{9}$$

If the new sample is found to be a low-fi one ( $l = 1$ ), which means that adding a low-fi sample will improve the hi-fi objective function most, it will be evaluated by the cheaper analysis. Otherwise, the new sample will be evaluated by the expensive analysis ( $l = 2$ ).

During the design, we found that the VF-EI will automatically tend to add more low-fi samples, which is a preferred feature for the consideration of improving the overall optimization efficiency. And the addition of low- and hi-fi samples will be in an alternative manner, along with the change of predicted

correlation between low- and hi-fi functions and the ratio of two kinds of uncertainties mentioned above,  $s_{lf}^2/s^2$ . For a HK model (see eq. (3)), supposing that the low-fi function can correctly capture the trend of the hi-fi function, the uncertainty of HK model, due to lack of a hi-fi sample (correction term in HK), will be generally much smaller than that due to lack of a low-fi sample (trend term in HK). As a result, maximizing the VF-EI will automatically lead to the more addition of low-fi samples. On the contrary, if the low-fi function correlates poorly with the hi-fi function, the  $\beta_0$  will be very small (close to zero), and maximizing the VF-EI will tend to add hi-fi samples since adding a low-fi sample has a little contribution to the improvement of the hi-fi function. On the other hand, the addition of low- or hi-fi samples is also controlled by the ratio of uncertainties  $s_{lf}^2/s^2$ . Besides the process variance,  $s_{lf}^2/s^2$  is related to the distance to the observed sample. At a point where the low-fi function is undersampled, the VF-EI may tend to add low-fi samples, simply because the process variance of the hi-fi HK model will be much smaller than that of the low-fi kriging. And at a point where low-fi function is densely sampled, the VF-EI may switch to add hi-fi samples.

### 3.2 Constraint handling

For a constrained optimization, the HK model for the constraint function,  $\hat{g}(\mathbf{x})$ , is also built, and we also assume the

prediction of the HK model for the constraint function at any site is a normal distribution

$$G(\mathbf{x}, l) \sim N\left[\hat{g}(\mathbf{x}), s_g^2(\mathbf{x}, l)\right], \quad l = 1, 2, \tag{10}$$

where  $s_g^2(\mathbf{x}, 1)$  and  $s_g^2(\mathbf{x}, 2)$  are the uncertainties of constraint function due to lack of low- and hi-fi samples, respectively. Therefore, the probability of satisfying the constraint is

$$P[G(\mathbf{x}, l) \leq 0] = \Phi\left(\frac{-\hat{g}(\mathbf{x})}{s_g(\mathbf{x}, l)}\right). \tag{11}$$

Eventually, the constrained VF-EI can be expressed as

$$CEI_{vf}(\mathbf{x}, l) = EI_{vf}(\mathbf{x}, l) \cdot P[G(\mathbf{x}, l) \leq 0]. \tag{12}$$

For an optimization problem with multiple constraints, the corresponding constrained VF-EI is

$$CEI_{vf}(\mathbf{x}, l) = EI_{vf}(\mathbf{x}, l) \cdot \prod_{i=1}^{N_C} P[G_i(\mathbf{x}, l) \leq 0], \tag{13}$$

where  $N_C$  is the number of constraints. It can be seen that in the above equation, the constrained VF-EI is a function of both the design variable  $\mathbf{x}$  and the fidelity level  $l$ .

Through maximizing the constrained VF-EI, the spatial location of new sample point  $\mathbf{x}$  and fidelity level of next numerical evaluation  $l$  can be determined by

$$\mathbf{x}, l = \underset{\mathbf{x}_{low} \leq \mathbf{x} \leq \mathbf{x}_{up}, l=1,2}{\operatorname{argmax}} CEI_{vf}(\mathbf{x}, l). \tag{14}$$

### 3.3 Discussion about the difference with augmented EI

By literature survey, we found that there is a similar but different method called augmented EI, which is proposed by Huang et al. (2006) and formulated based on cokriging model (Kennedy and O’Hagan 2000). The formulation of augmented EI is

$$EI_{aug}(\mathbf{x}, l) = EI(\mathbf{x}) \cdot \alpha_1 \cdot \alpha_2, \quad l = 1, 2, \tag{15}$$

where

$$\alpha_1 = \operatorname{corr}[\hat{y}_l(\mathbf{x}), \hat{y}_2(\mathbf{x})], \quad \alpha_2 = t_2/t_l, \tag{16}$$

where  $\alpha_1$  is a cross-correlation coefficient between the low- and hi-fi models and  $\alpha_2$  is the cost ratio between a single hi-fi numerical simulation and a single low-fi numerical simulation. As the cross-correlation coefficient  $\operatorname{corr}[\hat{y}_l(\mathbf{x}), \hat{y}_2(\mathbf{x})]$  could be not easy to calculate for other kinds of surrogate models, an alternative formulation is given in (Reisenthel and Allen 2014)

$$\operatorname{corr}[\hat{y}_l(\mathbf{x}), \hat{y}_2(\mathbf{x})] = s_l(\mathbf{x}) / \left( \left| \hat{y}_2(\mathbf{x}) - \hat{y}_l(\mathbf{x}) \right| + s_2(\mathbf{x}) \right). \tag{17}$$

Although the motivations and ideas are similar, the VF-EI method proposed in this article is different from the

augmented EI method. The main difference is that the VF-EI is derived analytically, but the augmented EI is somehow heuristic and empirical. As a result, VF-EI is free of empirical parameters and the addition of low- and hi-fi samples is fully adaptive. In contrast, the performance of the augmented EI is dependent on the proper use of low-fi function and empirical setting of  $\alpha_1$  and  $\alpha_2$ . From this point of view, we could infer that the VF-EI can be potentially better than or at least as good as the augmented EI. In fact, this hypothesis is confirmed by the analytical function test cases and engineering design problems provided in Sections 4 and 5.

### 3.4 Implementation of proposed VF-EI

In this article, all the optimization test cases are performed based on an in-house optimization code called ‘‘SurroOpt’’ (Han 2016a, b). It is a surrogate-based generic optimization code that can be used to efficiently solve arbitrary single and multi-objective (Pareto front), unconstrained, and constrained optimization problems. The proposed VF-EI method and augmented EI were implemented in ‘‘SurroOpt.’’ Algorithm 1 is given as an implementation of the VFO and the corresponding framework is sketched in Fig. 2. The difference with the conventional optimization framework based on the standard EI method (see Fig. 1) is that both low-fi and hi-fi samples are added to update the HK model towards the global optimum of the hi-fi function,

#### Algorithm 1: Procedure of VF-EI based optimization

---

```

0:   procedure
1:   set initial low- and hi-fi sample points,  $J = 0$ 
2:   evaluate the samples by low-fi and hi-fi analyses, respectively
3:   while  $J < J_{max}$ 
4:     construct the low-fi kriging model  $\hat{y}_{lf}$ 
5:     construct the HK model  $\hat{y}_{hf}$ 
6:     find  $\mathbf{x}_{lf}^* = \operatorname{argmax} CEI_{vf}(\mathbf{x}, l = 1)$ 
7:     find  $\mathbf{x}_{hf}^* = \operatorname{argmax} CEI_{vf}(\mathbf{x}, l = 2)$ 
8:     if  $EI_{vf}(\mathbf{x}_{lf}^*, 1) > EI_{vf}(\mathbf{x}_{hf}^*, 2)$  then:
9:        $\mathbf{x}_{new} = \mathbf{x}_{lf}^*, l = 1$ 
10:    else
11:       $\mathbf{x}_{new} = \mathbf{x}_{hf}^*, l = 2$ 
12:    endif
13:    if  $(l = 1)$  then
14:      evaluate  $\mathbf{x}_{new}$  using low-fi analysis, go to 4
15:    else
16:      evaluate  $\mathbf{x}_{new}$  using hi-fi analysis,  $J = J + 1$ , go to 5
17:    endif
18:  end while
19:  end procedure

```

---

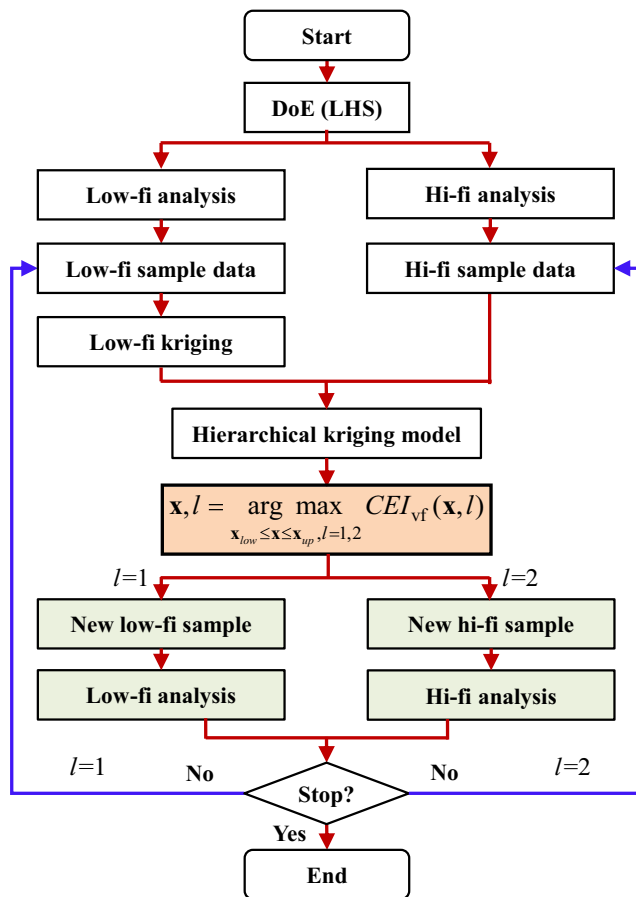


Fig. 2 Flowchart of VFO based on hierarchical kriging and the proposed VF-EI method

### 4 Test cases of analytical functions

#### 4.1 An illustrative example

In this sub section, a one-dimensional test function (Forrester et al. 2007) is used to verify the correctness of the proposed VF-EI method. The mathematical model is

$$\begin{aligned} \min \quad & y = (6x-2)^2 \sin(12x-4) \\ & y_{lf} = 0.5y + 10(x-0.5)-5. \end{aligned} \tag{18}$$

w.r.t.  $x \in [0, 1]$

Please note that the theoretical optimal solution is at  $x^* = 0.75725$  with optimal functional value of  $f(x^*) = -6.020740$ .

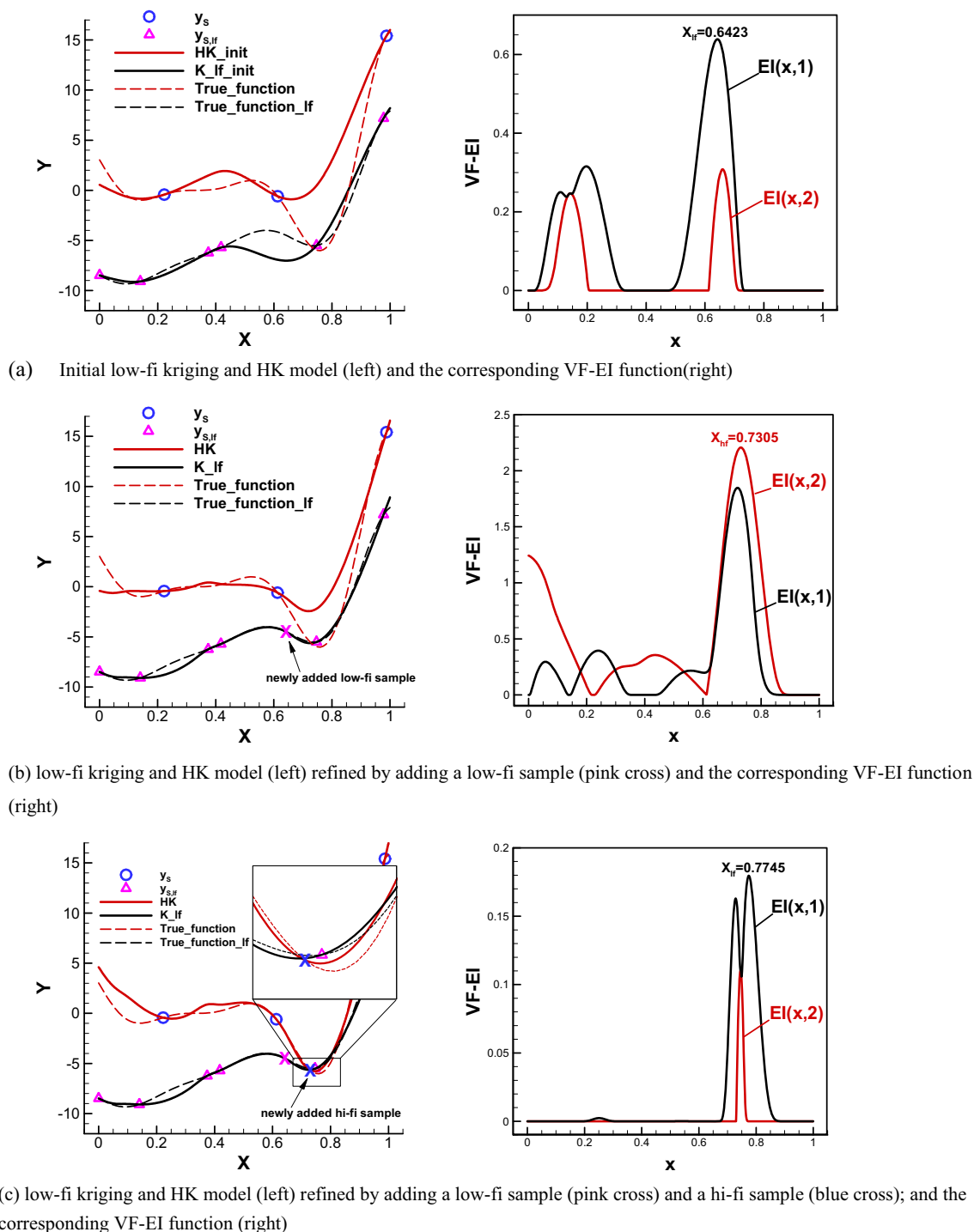
The refining process of the HK model by using the proposed VF-EI is sketched in Table 2 and Fig 3. *First*, the initial HK model is built with six low-fi and three hi-fi sample points ( $S_{lf} = \{0.0, 0.1402, 0.3742, 0.4180, 0.7469, 0.9777\}$ ,  $S = \{0.2223, 0.6128, 0.9877\}$ ), shown in Fig. 3a. Figure 3 (a, left) shows the low-fi kriging and HK models compared with the true functions. Although the low-fi kriging is not accurate enough, the trend is correctly mapped to three hi-fi

samples. Figure 3 (a, right) is the distribution of the VF-EI function through the whole design space. It can be seen that the maximum of  $EI(x, 1)$  in the low-fi level is much higher than the maximum of  $EI(x, 2)$  in the hi-fi level, which implies that for the 1st update cycle, we should add a low-fi sample at location of  $x = 0.6423$ . *Second*, the low-fi kriging and HK models are rebuilt with the initial samples and the newly added low-fi sample (pink cross symbol), as shown in Fig. 3 (b, left). The low-fi kriging model is getting more accurate and the resulting HK model features the similar trend with the true hi-fi function. As the maximum of  $EI(x, 2)$  in hi-fi level is larger, shown in Fig. 3 (b, right), a hi-fi sample point at  $x = 0.7305$  is added at the 2nd cycle. *Then*, the HK model is updated and becomes more accurate than the previous cycle. However, the exact optimal location is still not precisely reached, as shown in Fig. 3c. *Then*, two low-fi sample points are added into the sample dataset at the 3rd and 4th cycles and the low-fi kriging and HK models are both getting more and more accurate near the optimum, as shown Fig. 3d, e. *Finally*, a hi-fi sample point is added at a site that is very closing to the optimal location and the updating process terminates, see in Fig. 3f. The refining process of the HK model by using the augmented EI is sketched in Fig. 4. Here, we assume that the cost ratio  $t_2/t_1$  is 4. We find that only hi-fi samples are added during the optimization iterations. This is because the low- and hi-functions have a big difference in functional value, which results in a small value of correlation coefficient  $\alpha_1$ , consequently, and only hi-fi samples points are added during the entire design (see Table 3).

In this test case, only 5 hi-fi samples, including 3 initial samples, are used for VF-EI method to successfully find the global optimum (see Table 2), with assistance of 9 low-fi samples, which in turn verifies the correctness and effectiveness of the proposed VF-EI method. It is shown that the VF-EI method can adaptively determine both the location and fidelity level of a new design point, and the addition of both low- and hi-fi samples are in an alternative manner and hi-fi optimum oriented. During the updating process, the low-fi sample points are firstly used to depict the trend of the hi-fi function near the optimal solution; and then the hi-fi sample points are added to correct the absolute value until the hi-fi global optimum is reached.

#### 4.2 Other test functions

To demonstrate the VF-EI method and make a reasonable comparison with the augmented EI method, six analytic functions with the number of dimensions in the range from 1 to 5 are tested. Among them, the test functions 5 and 6 are from the literature (Huang et al. 2006), in which the augmented EI was proposed. The expressions of adopted functions are shown in



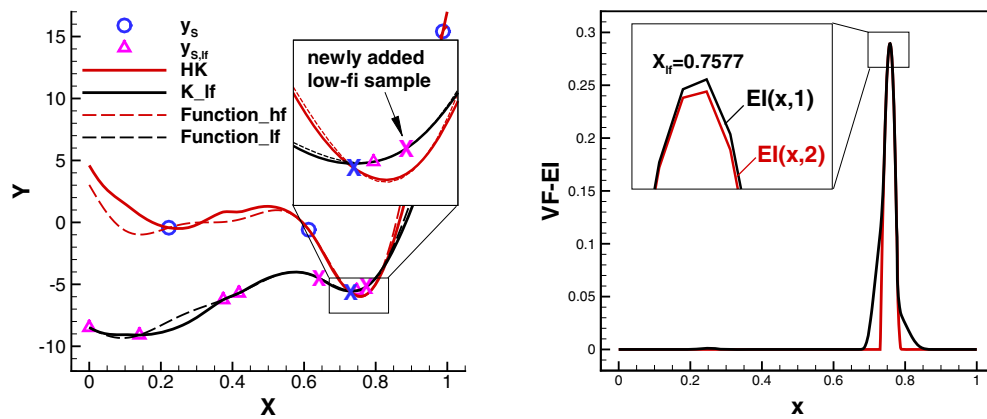
**Fig. 3** Refinement process of VFO of a 1-d analytical function based on HK surrogate model and the proposed VF-EI method (fidelity level of added samples: low-hi-low-low-hi)

Table 4, and the definitions of analytical test cases are shown in Table 5.

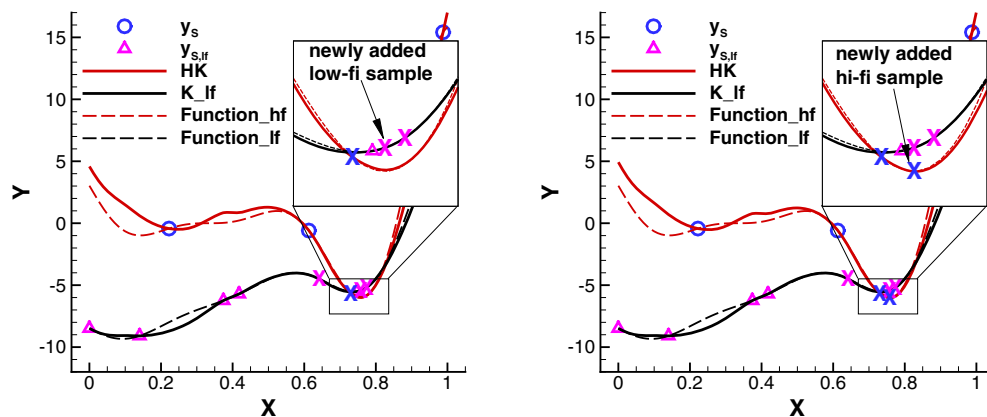
The test cases include the following: (1) case 1–3, all of the hi-fi functions are the 1-d function from (Forrester et al. 2007) with different low-fi functions. For Forrester

1a, the low-fi function has a different optimal location with the hi-fi function; the Forrester 1b has the same trend but large difference in functional response with the hi-fi function; the Forrester 1c has a totally opposite trend with the hi-fi function; (2) case 4 is a 2-d function





(d) low-fidelity kriging and HK model (left) refined by adding two low-fidelity samples (pink cross) and a high-fidelity sample (blue cross); and the corresponding VF-EI function (right)



(e) low-fidelity kriging and HK model refined by adding three low-fidelity samples (pink cross) and a high-fidelity sample (blue cross)

(f) low-fidelity kriging and HK model refined by adding three low-fidelity samples (pink cross) and two high-fidelity samples (blue cross)

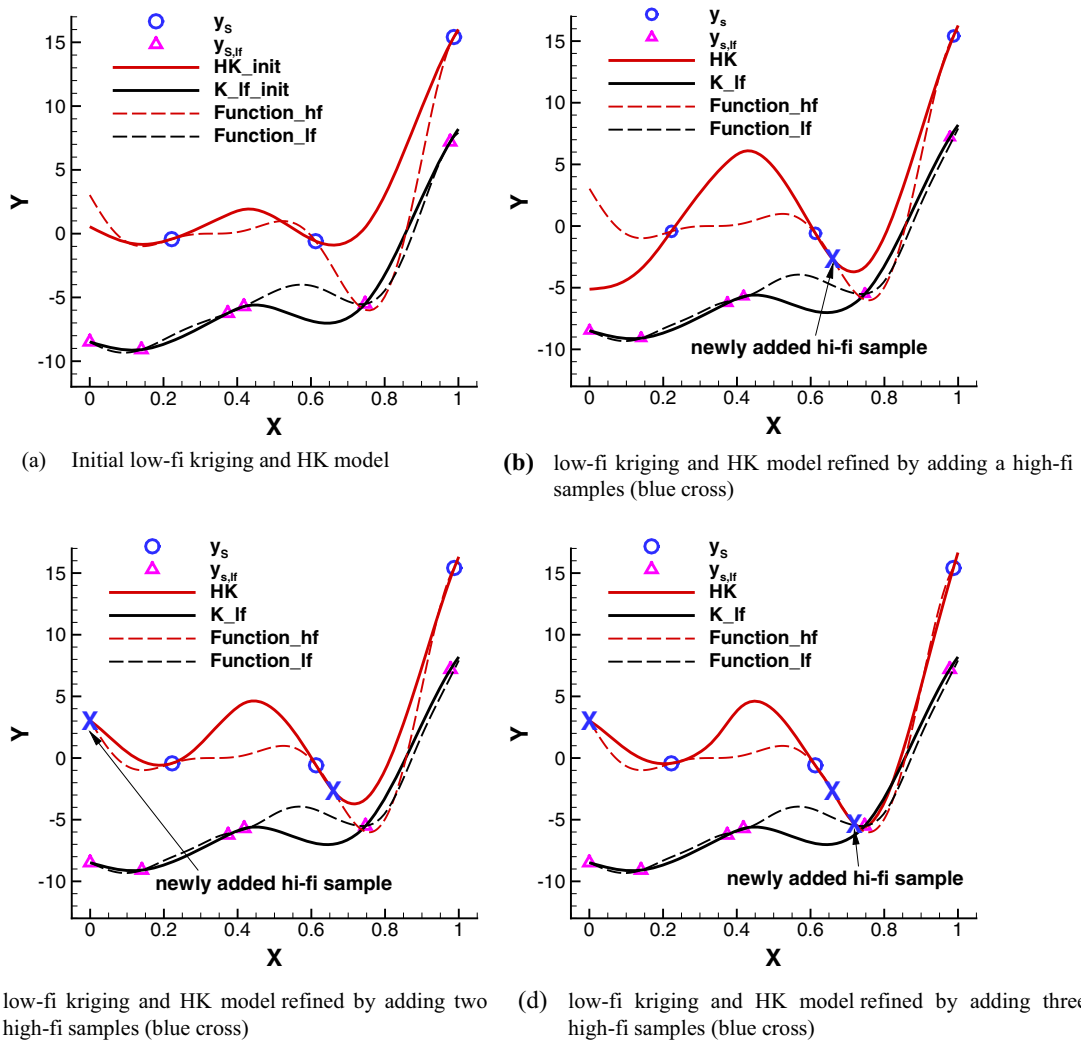
Fig. 3 (continued)

test case from Ref (Gano et al. 2005), where the low-fidelity function is obtained by adding linear and nonlinear noise factors to the high-fidelity function. The low-fidelity function has similar trend and close optimal location with the high-fidelity function; (3) case 5 is a 3-d ‘‘Hartman 3’’ function and the low-fidelity function is defined as adding systematic errors ‘‘MA 3’’ multiplied by 7.6 to the high-fidelity functions. As a large deviation is introduced in the low-fidelity function, there is a big difference in functional responses between low- and high-fidelity functions; (4) case 6 is a 5-d ‘‘Ackley 5’’ function, and the low-fidelity function is defined as adding systematic errors ‘‘MA 5’’ multiplied by 0.74 to the high-fidelity function, which can retain the main trend of the high-fidelity function. The 1-d and 2-d function cases, which can be visualized, are shown in Fig. 5.

In all of the three test cases, the initial numbers of low- and high-fidelity sample points are 3 and 6 times of the number of dimensions and chosen by LHS respectively. The optimization using

standard EI, augmented EI, and VF-EI are all repeated 10 times to consider the influence of the randomness due to initial sampling and the sub-optimization algorithms (when GA is used). The average test results are shown in Table 6.

From the test results, we can observe that, *first*, the global optima are successfully found in all of the cases, which validate the applicability of HK model-based optimization. *Second*, it is shown that the VF-EI can improve the optimization efficiency in all of the cases, in terms of the number of high-fidelity sample points, compared with that using the standard single-fidelity EI method. Although the low-fidelity function has a poor correlation with high-fidelity function (case 3), the performance of VF-EI is at least as good as the standard EI method. *Third*, as for the augmented EI, it only performs well in case 4, in which the low-fidelity function has similar trend and close optimal location with that of the high-fidelity function. In case 6, it does improve the optimization efficiency but not as good as our VF-EI method.



**Fig. 4** Refinement process of VFO of a 1-d analytical function based on HK surrogate model and the augmented EI method (fidelity level of added samples: hi-hi-hi)

However, it does not work well in other analytical function cases and even worse than the standard EI in case 5. Please note that the similar result of case 5 was presented in the literature (Huang et al. 2006). We observed that unsatisfactory performance of the augmented EI method

is due to the big difference in functional responses or optimal locations between low- and hi-fi functions.

In conclusion, the proposed VF-EI method can adaptively infill the low- and hi-fi sample points in an alternative manner, and is proved to be effective in all of the presented optimization cases. In contrast, the performance of the augmented EI method is problem dependent.

**Table 2** Refinement process of the VF-EI based optimization of the 1-d illustrative test case

Updating cycle	Added sample	Fidelity level	Best-observed hi-fi objective function
1	$x = 0.6423$	1 (low-fi)	-0.5903
2	$x = 0.7305$	2 (hi-fi)	-5.6696
3	$x = 0.7745$	1 (low-fi)	-5.6696
4	$x = 0.7577$	1 (low-fi)	-5.6696
5	$x = 0.7580$	2 (hi-fi)	-6.0205

**Table 3** Refinement process of the augmented EI based optimization of the 1-d illustrative test case

Updating cycle	Added sample	Fidelity level	Best-observed hi-fi objective function
1	$x = 0.6597$	2 (hi-fi)	-2.6835
2	$x = 0.0000$	2 (hi-fi)	3.0272
3	$x = 0.7190$	2 (hi-fi)	-5.3345
4	$x = 0.7529$	2 (hi-fi)	-6.0110

**Table 4** Test functions

Name and source	Description of functions
Forrester 1 (Forrester et al. 2007)	$f(x) = (6x-2)^2 \sin(12x-4)$ $x \in [0, 1]$ $x^* = 0.75725, f^* = -6.020740$
Forrester 1a	$f_{1f}(x) = 0.5f(x) + 10(x-0.5) - 5$
Forrester 1b	$f_{1f}(x) = f(x) - 5$
Forrester 1c	$f_{1f}(x) = [6(x+0.2) - 2]^2 \sin[12(x+0.2) - 4]$
Gano 2 (Gano et al. 2005)	$f(\mathbf{x}) = 4x_1^2 + x_2^3 + x_1x_2$ $g(\mathbf{x}) = 1/x_1 + 1/x_2 - 2$ $x_i \in [0.1, 10], \text{ for } i = 1, 2$ $\mathbf{x}^* = (0.8846146, 1.1500039), f^* = 5.668365$
Gano 2a	$f_{1f}(\mathbf{x}) = 4(x_1 + 0.1)^2 + (x_2 - 0.1)^3 + x_1x_2 + 0.1$ $g_{1f}(\mathbf{x}) = 1/x_1 + 1/(x_2 + 0.1) - 2 - 0.001$ $x_i \in [0.1, 10], i = 1, 2$
Hartman 3 (Hartman 1973)	$f(\mathbf{x}) = -\sum_{i=1}^4 c_i \exp\left[-\sum_{j=1}^3 \alpha_{ij} (x_j - p_{ij})^2\right]$ <p>where <math>\alpha_{ij} = \begin{bmatrix} 3 &amp; 10 &amp; 30 \\ 0.1 &amp; 10 &amp; 35 \\ 3 &amp; 10 &amp; 30 \\ 0.1 &amp; 10 &amp; 35 \end{bmatrix}</math>, <math>c_i = \begin{bmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{bmatrix}</math>, <math>p_{ij} = \begin{bmatrix} 0.3689 &amp; 0.1170 &amp; 0.2673 \\ 0.4699 &amp; 0.4387 &amp; 0.7470 \\ 0.1091 &amp; 0.8732 &amp; 0.5547 \\ 0.03815 &amp; 0.5743 &amp; 0.8828 \end{bmatrix}</math></p> $0 \leq x_i \leq 1, \text{ for } i = 1, 2, 3, \text{ and } j = 1, \dots, 4$ $N_{local} > 1, N_{global} = 1$ $\mathbf{x}^* = (0.114, 0.556, 0.852), f^* = -3.8627$
Ackley 5 (Ackley 1987)	$f(\mathbf{x}) = -a \exp\left[-b \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right] - \exp\left[\frac{1}{n} \sum_{i=1}^n \cos(cx_i)\right] + a + \exp(1)$ $a = 20; b = 0.2; c = 2\pi$ $-2.0 \leq x_i \leq 2.0, \text{ for } i = 1, \dots, 5$ $N_{local} > 1, N_{global} = 1$ $\mathbf{x}^* = (0, \dots, 0), f^* = 0.0$
MA 3 (McDaniel and Ankenman 2000)	$f(\mathbf{x}) = 0.585 - 0.324x_1 - 0.379x_2 - 0.431x_3 - 0.208x_1x_2 + 0.326x_1x_3$ $+ 0.193x_2x_3 + 0.225x_1^2 + 0.263x_2^2 + 0.274x_3^2$
MA 5 (McDaniel and Ankenman 2000)	$f(\mathbf{x}) = 0.585 - 0.00127x_1 - 0.00113x_2 - 0.00663x_3 - 0.0129x_4 - 0.00611x_5$ $+ 0.00526x_1x_4 + 0.0106x_1x_5 - 0.000626x_2x_4 - 0.00310x_2x_5$ $- 0.00724x_4x_5 - 0.00096x_3^2 - 0.0124x_4^2 - 0.0101x_5^2$

## 5 Application to engineering design problems

### 5.1 Drag minimization of RAE2822 airfoil in transonic viscous flow

#### 5.1.1 Problem statement

The VF-EI method is applied to a benchmark problem of airfoil design, defined by the AIAA aerodynamic design optimization discussion group (ADODG) (Zhang et al. 2016),

which is the drag minimization of a RAE 2822 airfoil, subject to area, and pitching moment constraints, at a freestream Mach number of 0.734, lift coefficient of 0.824, and Reynolds number of  $6.5 \times 10^6$ . It can be written as a standard nonlinear programming problem

$$\begin{aligned} \min \quad & C_d \\ \text{s.t.} \quad & C_l = 0.824 \\ & C_m \geq -0.092, \\ & \text{Area} \geq \text{Area}_0 \end{aligned} \quad (19)$$

**Table 5** Definition of variable-fidelity test cases

Number	Hi-fi function	Low-fi function	Cost <sub>hf</sub>	Cost <sub>lf</sub>
1	Forrester 1	Forrester 1a	1.0	0.25
2	Forrester 1	Forrester 1b	1.0	0.25
3	Forrester 1	Forrester 1c	1.0	0.25
4	Gano 2	Gano 2a	1.0	0.25
5	Hartman 3	Hartman 3 + MA 3 × 7.6	1.0	0.5
6	Ackley 5	Ackley 5 + MA 5 × 0.74	1.0	0.2

where  $C_l$ ,  $C_d$ , and  $C_m$  is the lift, drag, and pitching moment coefficient respectively and Area is the airfoil cross-sectional area normalized by the chord length.

For convenience, in our study, the problem is reformulated as following:

$$\begin{aligned}
 & \min C_d \\
 & s.t. \quad C_l \geq 0.824 \\
 & \quad C_m \geq -0.092, \\
 & \quad Area \geq Area_0
 \end{aligned} \tag{20}$$

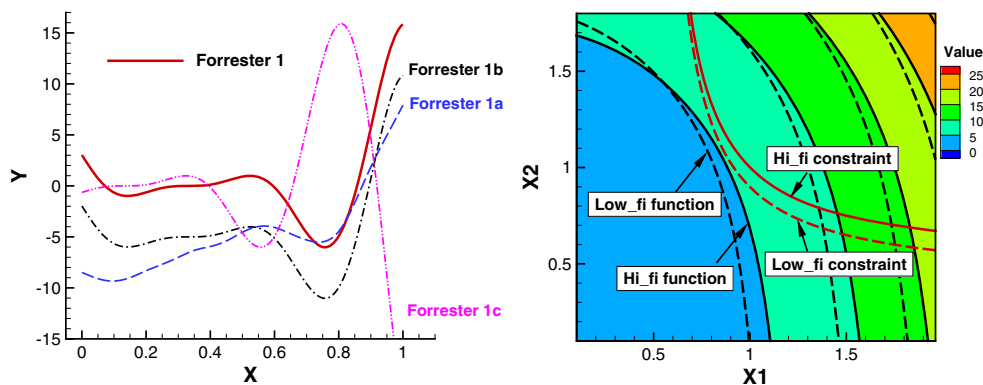
where the angle of attack  $\alpha$  is fixed at 2.8795 degree, at which the  $C_l$  of baseline RAE 2822 is 0.824. Please note that in the following text, the drag coefficient will be given in terms of drag count (cts) with  $1 \text{ cts} = 1 \times 10^{-4}$ .

**5.1.2 Determination of low- and hi-fi CFD models**

The high- and low-fi analysis models are defined as CFD simulations solving the same governing flow equations (Reynolds-averaged Navier-Stokes equations) with different convergence criteria and computational meshes of varying resolutions. The flow solver used in this paper is an in-house solver called “PMNS2D” (Han et al. 2007; Xie et al. 2008). According to the grid study shown in Fig. 6, we choose the grid with 73,728 cells as the hi-fi model, and the grid with 18,432 cells as the low-fi model

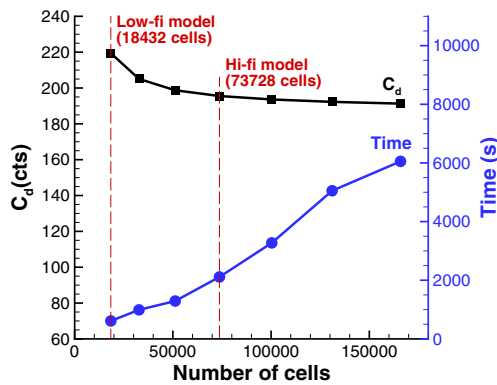
**Table 6** Comparison of VFOs with the standard EI, augmented EI, and the proposed VF-EI for analytical test cases

Test case	ISC methods	Averaged No. of added sample points		Difference with optima*
		Low-fi sample	Hi-fi sample	
Case 1 (1d)	EI	/	10	0.0002
	Augmented EI	2.5	8	0.00006
	VF-EI	18.9	6	0.00005
Case 2 (1d)	EI	/	8	0.001
	Augmented EI	1.4	8	0.0003
	VF-EI	21.8	5	0.00006
Case 3 (1d)	EI	/	8	0.00006
	Augmented EI	1.9	8	0.00005
	VF-EI	10.6	7	0.00005
Case 4 (2d)	EI	/	43	0.015
	Augmented EI	25.8	18	0.0016
	VF-EI	28	20	0.0012
Case 5 (3d)	EI	/	38	0.0025
	Augmented EI	2.8	47	0.002
	VF-EI	14	12	0.003
Case 6 (5d)	EI	/	71	0.1
	Augmented EI	129.4	36	0.1
	VF-EI	132	20	0.1



**Fig. 5** Sketch of the low- and hi-fi functions for case 1 to 4 (left: case 1–3; right: case 4)

(see in Fig. 7). The low-fi CFD simulation terminates after 200 iterations and the hi-fi CFD simulate stops until the density residual has dropped 5.5 orders in magnitude, as shown in Fig. 8. As a result, the computational cost ratio of a hi-fi simulation to a low-fi simulation is approximately 100, which means that the cost of 100 low-fi CFD simulations is identical to that of a single hi-fi CFD simulation.



**Fig. 6** Determination of low- and hi-fi CFD models by grid convergence study for baseline RAE2822 airfoil ( $Ma = 0.734$ ,  $\alpha = 2.8795^\circ$ ,  $Re = 6.5 \times 10^6$ )

### 5.1.3 Results

The airfoil is parameterized by class-shape transformation (CST) method (Kulfan 2008), and 18 design variables are utilized to describe the airfoil. After parameterizing the baseline RAE2822 airfoil, the design space is obtained by expanding the initial parameters by 1.5 times and narrowing it by half.

Here, we compare VF-EI, augmented EI, and standard EI-based VFO with the single-fidelity optimization. 5 hi-fi and 100 low-fi samples are selected by DoE to construct the initial HK model. To increase the confidence of the statistical analysis, the optimization based on each method is repeated 10 times. Figure 9 shows the comparison of the shapes and pressure coefficient distributions of baseline and optimal airfoils. As one can see that the location of maximal thickness is moved backward, compared with the baseline RAE2822 airfoil, and the leading edge radius is reduced. Consequently, the shock intensity is decreased sharply.

Figure 10 shows the averaged convergence histories of the objective function. The vertical bars represent the

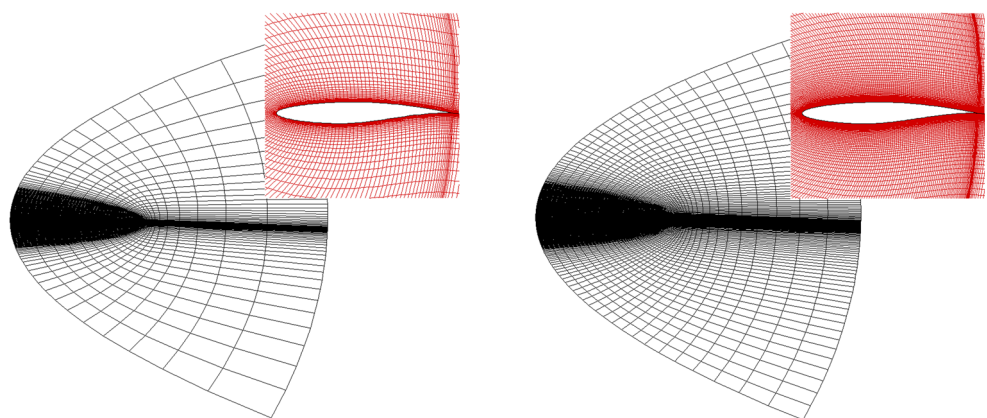
standard deviation of the objective function caused by repeating the optimizations. Please note that the convergence history is plotted in a manner that shows the variation of the best hi-fi objective function observed so far versus the number of hi-fi functional evaluations only. It is obvious that both the VFOs with the proposed VF-EI, augmented EI, and the standard EI perform much better than the single-fidelity optimization, and both the VF-EI and augmented EI are dramatically more efficient than the standard EI. For using the VF-EI, although more low-fi samples are used, the total computational cost is reduced dramatically, since a low-fi CFD is much cheaper than a hi-fi CFD. Table 7 shows the comparison of optimization results using four different methods. It can be seen that the VF-EI-based method offers the best design on average and smallest standard deviation. Table 8 is for the comparison of optimization efficiency of using four different methods. We can see that when 25 hi-fi sample points are used for the VFO with the standard EI method, the drag is reduced by 36.03%. In contrast, only 13 hi-fi samples (including 5 initial hi-fi samples) are needed to achieve the same level of drag reduction when the VF-EI method is used, which means that the VF-EI is almost twice faster than the standard EI method.

## 5.2 Drag minimization of ONERA M6 wing in transonic inviscid flow

### 5.2.1 Problem statement

This case is the drag minimization of a ONERA M6 wing at a freestream condition of  $Ma = 0.8395$ ,  $\alpha = 3.06^\circ$ , subject to four constraints on the maximum thickness-to-chord ratio and one constraints on the lift coefficient. The wing is parameterized by four control sections with the planar shape being fixed (see Fig. 11). Each of the control section is parameterized by five-order CST method, resulting in 12

**Fig. 7** Sketch of C type grids for low-fi (left) and hi-fi (right) CFD computations for baseline RAE2822 case



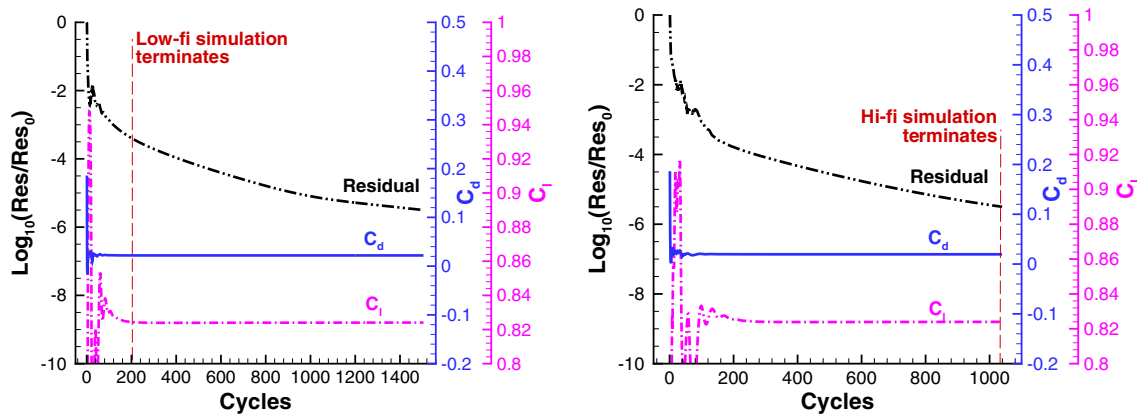


Fig. 8 Convergence histories of low-fi (left) and hi-fi (right) simulations for the baseline RAE2822 airfoil ( $Ma = 0.734$ ,  $\alpha = 2.8795^\circ$ ,  $Re = 6.5 \times 10^6$ )

variables for each section and 48 design variables in total. It can be written as a standard nonlinear programming problem:

$$\begin{aligned}
 &\min C_D \\
 &s.t. \quad 1) C_L \geq C_{L,0} \\
 &\quad \quad 2) t_1 \geq t_{1,0} \quad 3) t_2 \geq t_{2,0} \\
 &\quad \quad 4) t_3 \geq t_{3,0} \quad 5) t_4 \geq t_{4,0}
 \end{aligned} \tag{21}$$

where  $t$  denotes the maximum thickness-to-chord ratio of the control sections, and its subscript is the index of the control sections.

### 5.2.2 Determination of low- and hi-fi CFD models

The flow analyses are performed with the in-house code called PMNS3D. It solves the Euler equations to simulate the inviscid flow around the wing. With the structured grids of C-H topology generated by our in-house code, the low- and hi-fi CFD models are also defined in a similar way with the previous cases. According to the grid study shown in Fig. 12, we choose the grid with the distribution of 168 (chord-wise direction)  $\times$  36 (normal direction)  $\times$  44 (span direction) as the low-fi CFD, and the grid with the distribution of 312 (chord-wise direction)  $\times$  72 (normal direction)  $\times$  72 (span direction) as the

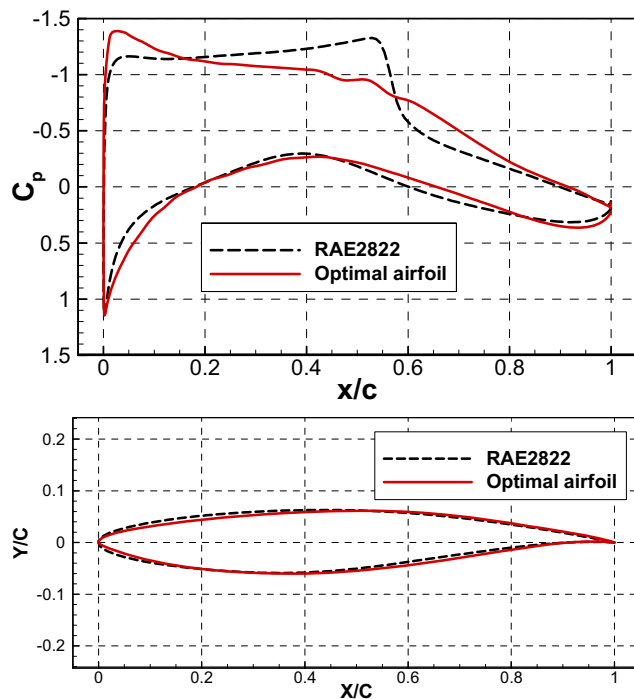


Fig. 9 Comparison of baseline RAE 2822 and optimal airfoils (upper: pressure coefficient distributions; lower: airfoil shapes,  $Ma = 0.734$ ,  $\alpha = 2.8795^\circ$ ,  $Re = 6.5 \times 10^6$ )

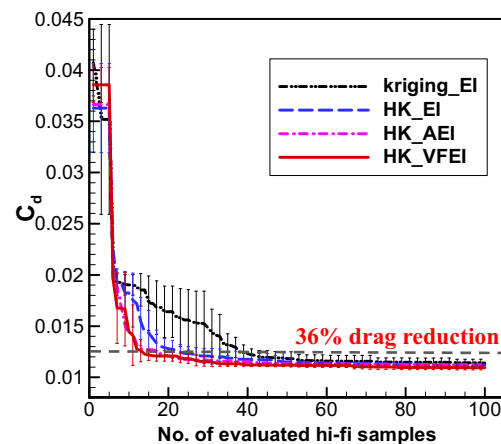


Fig. 10 Comparison of convergence histories of VFOs using proposed VF-EI, augmented EI, and the standard EI with the single-fidelity optimization based on kriging and EI (RAE2822 airfoil test case)

**Table 7** Comparison of optimization results using different methods (RAE2822 airfoil test case)

Surrogate models	Infill-sampling criterion	Optimal $C_d$ (cts)			Standard deviation (cts)
		Best	Average	Worst	
kriging	EI	110.80	113.91	118.53	3.2
HK	EI	110.08	111.83	114.17	1.1
	Augmented EI	108.65	112.23	117.94	3.1
	VF-EI	107.80	109.38	113.41	2.1

hi-fi CFD (see in Fig. 13). The low-fi CFD simulation terminates after 20 iterations and the hi-fi CFD simulate stops until the density residual has dropped 5.2 orders in magnitude, as shown in Fig. 14. For this wing case, the computational cost ratio of hi-fi CFD to low-fi CFD is approximately 10.

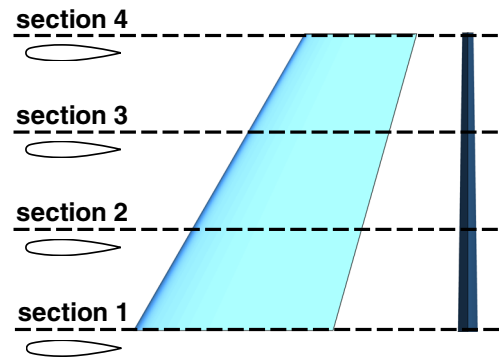
**5.2.3 Results**

The design space is obtained by expanding the CST parameters of control sections of M6 wing by 1.25 times and narrowing it by a quarter. And 24 hi-fi and 300 low-fi samples are selected by DoE to construct the initial HK model. Here, only VFO based on HK model, using VF-EI, augmented EI, and standard EI are conducted, as it is difficult for single-fidelity method to optimize the wing with 48 design variables. The optimizations are also repeated for 10 times in this case. Table 9 compares the aerodynamic performance and thickness of the control sections between the baseline and optimized wings. The drag coefficient is reduced by 29.69 cts with all of the constraints are satisfied. Figure 15 shows the comparison of pressure contour

**Table 8** Comparison of optimization efficiency of using different methods for same level of drag reduction (RAE2822 airfoil test case)

Surrogate model	Infill-sampling criterion	Average $C_d$ (cts)	$\Delta$	$N_l$	$N_h$	Equivalent to no. of hi-fi CFD
kriging	EI	123.09	- 35.98%	/	40	40
HK	EI	122.99	- 36.03%	100	25	26
	Augmented EI	122.43	- 36.32%	200	18	20
	VF-EI	122.52	- 36.29%	189	13	15

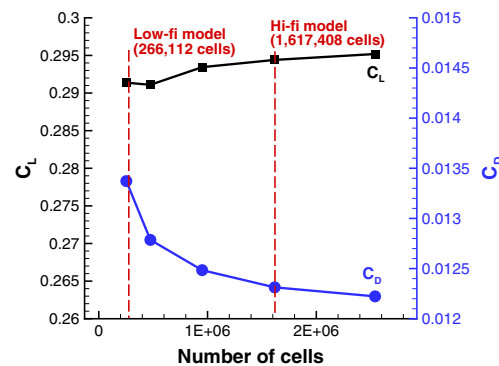
$N_l$ : number of low-fi CFD analyses,  $N_h$ : number of hi-fi CFD analyses



**Fig. 11** Parameterization of ONERA M6 wing using 4 control sections with the 48 design variables in total

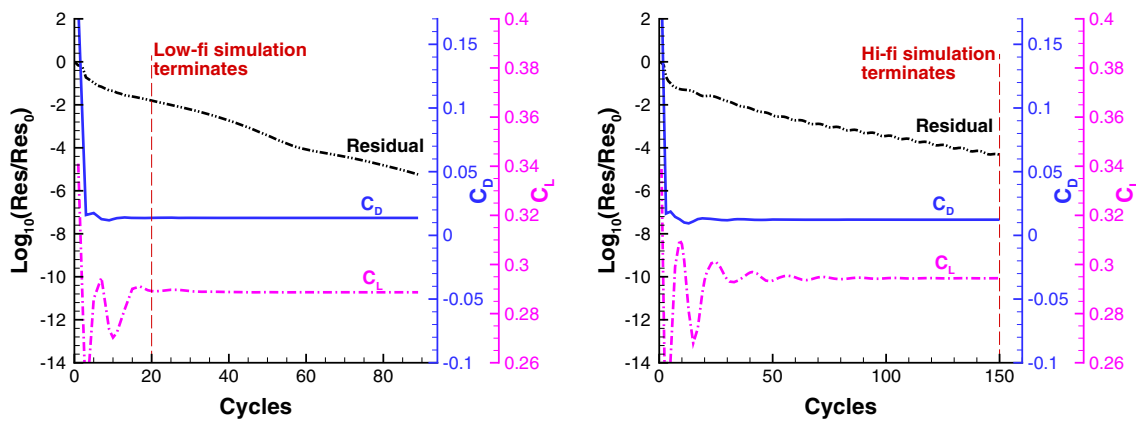
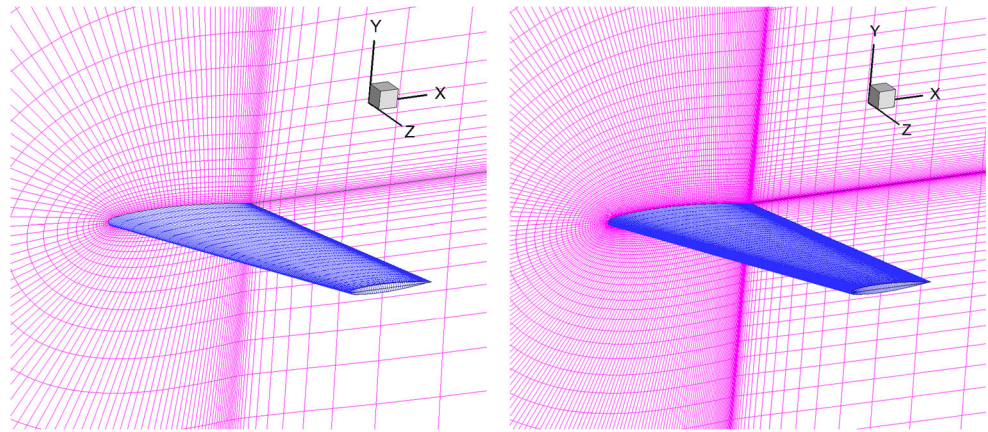
of baseline and optimized wings. The result is further checked by comparison of the pressure distributions and shape at four span-wise sections of the wing. As we can see that, the shock is weakened for the optimal shape at all of the sections.

The averaged convergence histories with standard deviation of using different optimization strategies are sketched in Fig. 16. One can see that the optimization efficiency can be improved by using the proposed VF-EI method. Table 10 is for the comparison of the optimization results using three different methods. Again, it is shown that the VF-EI-based method offers the best design on average. Table 11 is for the comparison of optimization efficiency of using three different methods for the same level of drag reduction. We can see that the VF-EI method also performs the best. When 17 hi-fi sample points are added (41 in total) for the VFO with the VF-EI method, the drag is reduced by 19.65%. As for the augmented EI and standard EI method, to achieve the same level of drag reduction,



**Fig. 12** Determination of low- and hi-fi CFD models by grid convergence study for baseline M6 wing ( $Ma = 0.8395$ ,  $\alpha = 3.06^\circ$ )

**Fig. 13** Sketch of C type grids for low-fi (left) and hi-fi (right) CFD computations for baseline M6 wing case



**Fig. 14** Convergence histories of low-fi (left) and hi-fi (right) simulations for baseline M6 wing ( $Ma = 0.8395$ ,  $\alpha = 3.06^\circ$ )

30 and 43 hi-fi sample points are need to be added. And the augmented EI method does improve the optimization efficiency, but not as good as our VF-EI method.

### 6 Summary and future works

In this article, a variable-fidelity expected improvement (VF-EI) method was proposed for efficient global

optimization of expensive-to-evaluate function with assistance of cheaper-to-evaluate function, based on a variable-fidelity model called hierarchical kriging (HK). The VF-EI method, as an extension of standard EI method, can select the new samples of both low fidelity and high fidelity to update the VFM towards the global optimum of the hi-fi function, and the infillings of both low-fi and hi-fi samples are hi-fi-optimum oriented.

**Table 9** The optimization results for ONREA M6 wing test case ( $Ma = 0.8395$ ,  $\alpha = 3.06^\circ$ )

	$C_L$	$C_D$ (cts)	$C_M$	$t_1$	$t_2$	$t_3$	$t_4$
Baseline	0.294408	123.13	-0.22356	0.098733	0.098733	0.098733	0.098733
Opt.	0.294018	95.44	-0.21996	0.098832	0.098806	0.099038	0.098734
$\Delta$	-0.00039	-27.69	-0.0036	0.49E-04	0.23E-04	0.31E-03	0.69E-06



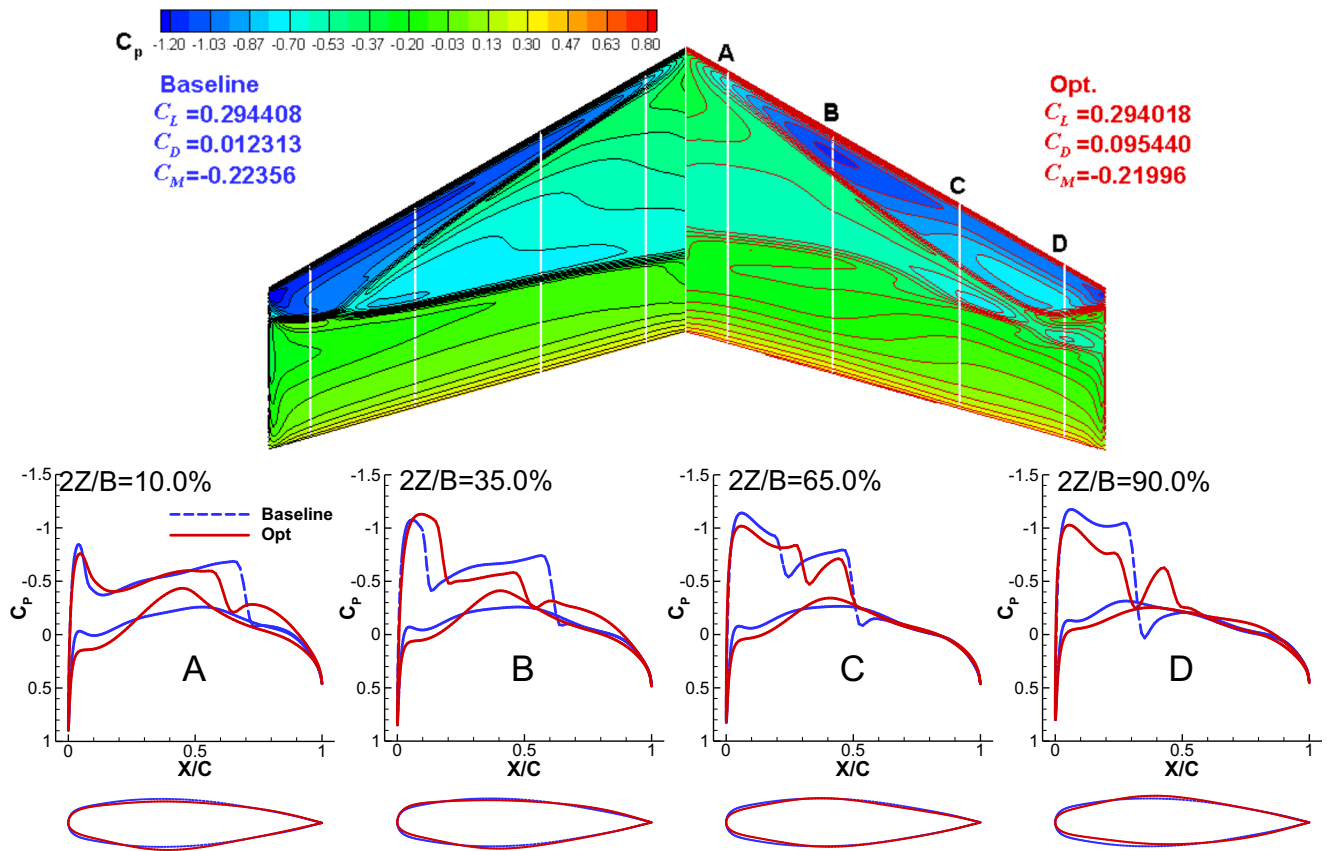


Fig. 15 Comparison of baseline ONREA M6 and the optimal wings ( $Ma = 0.8395$ ,  $\alpha = 3.06^\circ$ )

By redefining the uncertainty of the HK model, the VF-EI was analytically derived. The VF-EI refers to the expected improvement of the hi-fi function by infilling

either a low-fi or a hi-fi sample point at any untried  $x$  in the design space. The uncertainty and expected improvement of the hi-fi function due to lack of a low-fi sample was analytically derived, and then the resulting VF-EI function is formulated as a function of both the design variables  $x$  and the fidelity level  $l$ . By maximizing the VF-EI, the new sample point and its fidelity level can be determined. The VF-EI method was further extended to handle nonlinear constraints, by evaluating

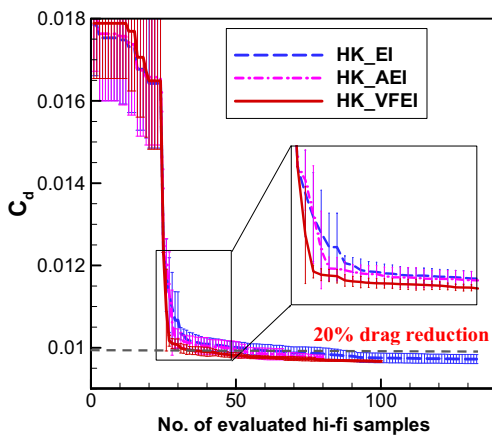


Fig. 16 Comparison of convergence histories of VFOs using proposed VF-EI, augmented EI, and the standard EI based on HK (ONREA M6 wing test case)

Table 10 Comparison of optimization results using different methods (ONREA M6 wing test case)

Surrogate models	Infill-sampling criterion	Optimal $C_d$ (cts)			Standard deviation (cts)
		Best	Average	Worst	
HK	EI	95.44	97.23	99.25	1.1
	Augmented EI	96.90	97.63	98.57	0.6
	VF-EI	96.41	96.69	96.96	0.2

**Table 11** Comparison of optimization efficiency of using different methods for same level of drag reduction (ONREA M6 wing test case)

Surrogate model	Infill-sampling criterion	Average $C_d$ (cts)	$\Delta$	$N_l$	$N_h$	Equivalent to no. of hi-fi CFD
HK	EI	98.89	- 19.69%	300	67	97
	Augmented EI	98.86	- 19.71%	308	54	85
	VF-EI	98.94	- 19.65%	338	41	75

$N_l$ : number of low-fi CFD analyses,  $N_h$ : number of hi-fi CFD analyses

the probability of satisfying the constraints using the model uncertainty associated with lack of a low-fi sample as well.

Theoretical analysis shows that, when VF-EI is used, typically the addition of low- and hi-fi samples will be in an alternative way, which is controlled by the ratio of the two kinds of uncertainties, where are due to lack of low- and hi-fi samples, respectively. Besides, the VF-EI will tend to add more low-fi samples than hi-fi samples, if the low- and hi-fi functions are highly correlated. When compared to an existing method called augmented EI, we found that the main difference is that the VF-EI is analytically derived and free of empirical parameters, but the performance of the augmented EI is dependent on the proper setting of two empirical parameters.

The VF-EI method was then verified and validated by optimization of six analytical functions, and demonstrated for aerodynamic shape optimizations of the benchmark RAE 2822 airfoil in transonic viscous flow and a 48-dimensional ONREA M6 wing in transonic inviscid flow. It was shown that the proposed VF-EI method can adaptively infill the low-fi and hi-fi samples in an alternative manner, and is proved to be effective in all of the presented optimization cases. In contrast, the performance of augmented EI method is problem dependent.

Our future work will focus on applying the VF-EI method to higher dimensional optimization, in order to examine its performance for complex engineering design problems. This is still challenging for surrogate-based optimization method, especially when the number of dimension is larger than 100. We believe that extending the two-level VFMs to multi-fidelity models and developing dedicated ISC can be a remedy to tackle this problem. On the other hand, the use of cheap gradients obtained by adjoint method or automatic differentiation will be another promising way to ameliorate the ‘‘curse of dimensionality.’’

**Acknowledgements** The author would like to thank Dr. Jun Liu and Prof. Wen-Ping Song for their valuable suggestion and discussion.

**Funding information** This research was sponsored by the National Natural Science Foundation of China (NSFC) under grant no. 11772261, Aeronautical Science Foundation of China under grant No.

2016ZA53011 and Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University under grant No. CX201801.

## Appendix

### A.1 Hierarchical kriging

For an  $m$  dimensional problem, suppose we are concerned with the prediction of an expensive-to-evaluate (and unknown) hi-fi function  $y(\mathbf{x}) : \mathcal{X}^m \rightarrow \mathcal{Y}$ , with the assistance of a cheaper-to-evaluate low-fidelity function  $y(\mathbf{x}) : \mathcal{X}^m \rightarrow \mathcal{Y}$ . Assuming that the low- and hi-fi functions are observed at  $n_{lf}$  and  $n$  sites, respectively, the sample datasets for a HK model are

$$\begin{aligned} \mathbf{S}_{lf} &= [\mathbf{x}_{lf}^{(1)}, \dots, \mathbf{x}_{lf}^{(n_{lf})}]^T \in \mathcal{X}^{n_{lf} \times m}, \mathbf{y}_{S,lf} = [y_{lf}^{(1)}, \dots, y_{lf}^{(n_{lf})}]^T \in \mathcal{Y}^{n_{lf}}, \\ \mathbf{S} &= [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}]^T \in \mathcal{X}^{n \times m}, \mathbf{y}_S = [y^{(1)}, \dots, y^{(n)}]^T \in \mathcal{Y}^n, \end{aligned} \tag{22}$$

where the subscript ‘‘lf’’ denotes ‘‘low fidelity.’’

#### Kriging model of cheap low-fi function

To build a surrogate model for the hi-fi and expensive function, we first build a surrogate model for the lower fidelity but cheaper function that will be used thereafter to assist the prediction. Assume a random process corresponding to the unknown low-fi function  $y_{lf}(\mathbf{x})$

$$Y_{lf}(\mathbf{x}) = \beta_{0,lf} + Z_{lf}(\mathbf{x}), \tag{23}$$

where  $\beta_{0,lf}$  is an unknown constant and  $Z_{lf}(\mathbf{x})$  is a stationary random process. Then, we can follow (Sacks et al. 1989) to build a kriging based on the sampled data set  $(\mathbf{S}_{lf}, \mathbf{y}_{S,lf})$ . After the kriging is fitted, the prediction of the low-fidelity function at any untried point  $x$  can be written as

$$\hat{y}_{lf}(\mathbf{x}) = \beta_{0,lf} + \mathbf{r}_{lf}^T(\mathbf{x}) \mathbf{R}_{lf}^{-1}(\mathbf{y}_{S,lf} - \beta_{0,lf} \mathbf{1}), \tag{24}$$

where  $\beta_{0,lf} = (\mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \mathbf{R}_{lf}^{-1} \mathbf{y}_{S,lf}$ ;  $\mathbf{R}_{lf} \in \mathcal{Y}^{n_{lf} \times n_{lf}}$  is the correlation matrix representing the correlation between the observed low-fi sample points;  $\mathbf{1} \in \mathcal{Y}^{n_{lf}}$  is a column vector filled with ones;

and  $\mathbf{r}_{\text{lf}} \in \mathcal{D}^{\mathcal{N}^{\text{lf}}}$  is the correlation vector representing the correlation between the untried point and the observed low-fi sample points. The MSE of the kriging prediction at any untried  $\mathbf{x}$ , the uncertainty due to the lack of a low-fi sample, is

$$\text{MSE}[\hat{y}_{\text{lf}}(\mathbf{x})] = s_{\text{lf}}^2(\mathbf{x}) = \sigma_{\text{lf}}^2 \left[ 1.0 - \mathbf{r}_{\text{lf}}^T \mathbf{R}_{\text{lf}}^{-1} \mathbf{r}_{\text{lf}} + (\mathbf{r}_{\text{lf}}^T \mathbf{R}_{\text{lf}}^{-1} \mathbf{1} - 1)^2 / \mathbf{1}^T \mathbf{R}_{\text{lf}}^{-1} \mathbf{1} \right]. \tag{25}$$

The reader is referred to (Simpson et al. 2001; Martin and Simpson 2005; Toal et al. 2008) for more details of building such a kriging.

### Hierarchical kriging for expensive high-fidelity function

By directly taking the low-fi kriging multiplied by a scaling factor  $\beta_0$  as the model trend, the random process for the unknown hi-fi function is assumed as

$$Y(\mathbf{x}) = \beta_0 \hat{y}_{\text{lf}}(\mathbf{x}) + Z(\mathbf{x}), \tag{26}$$

where  $\beta_0$  is a scaling factor representing how the low-fi kriging matches the hi-fi function;  $\hat{y}_{\text{lf}}(\mathbf{x})$  denotes the prediction of low-fi kriging at  $\mathbf{x}$ ; and  $Z(\mathbf{x})$  is a stationary random process with zero mean. In this way, the trend of low-fi function is mapped to the sampled hi-fi data, resulting in a more accurate surrogate model for hi-fi function of interest. Then, a HK model (Han and Görtz 2012) can be built through the hi-fi sample datasets  $(\mathbf{S}, \mathbf{y}_{\text{S}})$ . By minimizing the MSE of the prediction, the HK prediction for the hi-fi and expensive function at any untried  $\mathbf{x}$  can be written as

$$\hat{y}(\mathbf{x}) = \beta_0 \hat{y}_{\text{lf}}(\mathbf{x}) + \mathbf{r}^T(\mathbf{x}) \mathbf{R}^{-1} (\mathbf{y}_{\text{S}} - \beta_0 \mathbf{F}), \tag{27}$$

where  $\mathbf{F} \in \mathcal{D}^{\mathcal{N}^{\text{lf}}}$  is a column vector filled with predictions of the low-fi kriging at the sites of hi-fi samples;  $\mathbf{R} \in \mathcal{D}^{\mathcal{N}^{\text{lf}} \times \mathcal{N}^{\text{lf}}}$  is the correlation matrix representing the correlation between the observed hi-fi sample points; and  $\mathbf{r} \in \mathcal{D}^{\mathcal{N}^{\text{lf}}}$  is the correlation vector representing the correlation between the untried point and the observed hi-fi sample points

$$\begin{aligned} \beta_0 &= (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_{\text{S}} \\ \mathbf{F} &= [\hat{y}_{\text{lf}}(\mathbf{x}^{(1)}), \dots, \hat{y}_{\text{lf}}(\mathbf{x}^{(n)})]^T \in \mathcal{D}^{\mathcal{N}^{\text{lf}}}, \\ \mathbf{R} &= \begin{bmatrix} R(\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) & \dots & R(\mathbf{x}^{(1)}, \mathbf{x}^{(n)}) \\ \vdots & \ddots & \vdots \\ R(\mathbf{x}^{(n)}, \mathbf{x}^{(1)}) & \dots & R(\mathbf{x}^{(n)}, \mathbf{x}^{(n)}) \end{bmatrix} \in \mathcal{D}^{\mathcal{N}^{\text{lf}} \times \mathcal{N}^{\text{lf}}}, \\ \mathbf{r} &= [R(\mathbf{x}, \mathbf{x}^{(1)}), \dots, R(\mathbf{x}, \mathbf{x}^{(n)})]^T \in \mathcal{D}^{\mathcal{N}^{\text{lf}}}, \end{aligned} \tag{28}$$

where  $R(\mathbf{x}, \mathbf{x}')$  is the spatial correlation function which only depends on the Euclidean distance between the two sites  $\mathbf{x}, \mathbf{x}'$ . Generally, the  $R(\mathbf{x}, \mathbf{x}')$  can be a Gaussian exponential function or a cubic spline function, see (Han and Görtz 2012).

The MSE of the HK prediction at any untried  $\mathbf{x}$ , the uncertainty due to lack of a hi-fi sample, is

$$\begin{aligned} \text{MSE}\{\hat{y}(\mathbf{x})\} &= s^2(\mathbf{x}) \\ &= \sigma^2 \left\{ 1.0 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \left[ \mathbf{r}^T \mathbf{R}^{-1} \mathbf{F} - \hat{y}_{\text{lf}}(\mathbf{x}) \right]^2 / (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \right\}. \end{aligned} \tag{29}$$

In comparison to a cokriging model, the HK model does not need to calculate the cross covariance between low- and hi-fi samples. As a result, the correlation matrix of a HK model is relatively smaller. In addition, the HK model can provide a more reasonable MSE estimation than any of the existing kriging and cokriging models (Han and Görtz 2012), which is very beneficial for infill-sampling based on the method such as EI.

### A.2 Standard expected improvement method

For using the expected improvement (EI) method proposed by Jones et al. (1998), we can assume that the prediction of the HK model at any untried site  $\mathbf{x}$  obeys a normal distribution  $\hat{Y}(\mathbf{x}) \sim \mathcal{N}[\hat{y}(\mathbf{x}), s^2(\mathbf{x})]$ , with the mean being the surrogate prediction  $\hat{y}(\mathbf{x})$  and the standard deviation  $s(\mathbf{x})$  being its root mean-squared error (RMSE). Then the statistical improvement at any untried location w.r.t. the best hi-fi objective function observed so far  $y_{\text{min}}$  is defined as:

$$I(\mathbf{x}) = \max(y_{\text{min}} - \hat{Y}(\mathbf{x}), 0). \tag{30}$$

Then, the EI function can be written as

$$EI(\mathbf{x}) = \begin{cases} (y_{\text{min}} - \hat{y}(\mathbf{x})) \Phi\left(\frac{y_{\text{min}} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) + s(\mathbf{x}) \phi\left(\frac{y_{\text{min}} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right), & \text{if } s(\mathbf{x}) > 0, \\ 0, & \text{if } s(\mathbf{x}) = 0 \end{cases} \tag{31}$$

where  $\Phi$  and  $\phi$  are the cumulative distribution function and probability density function of standard normal distribution, respectively.

For a constrained optimization, the HK model for the constraint function,  $\hat{g}(\mathbf{x})$ , is also built. We can also assume the prediction at any untried site obeys a normal distributed,  $\hat{G}(\mathbf{x}) \sim \mathcal{N}[\hat{g}(\mathbf{x}), s_g^2(\mathbf{x})]$ , with the mean value being the predictor  $\hat{g}(\mathbf{x})$  and the standard deviation being its RMSE  $s_g(\mathbf{x})$ . Therefore, the probability of satisfying the constraint is

$$P[\hat{G}(\mathbf{x}) \leq 0] = \Phi\left(\frac{-\hat{g}(\mathbf{x})}{s_g(\mathbf{x})}\right). \tag{32}$$

And the constrained EI function can be given by

$$CEI(\mathbf{x}) = E[I(\mathbf{x}) \cap \{G(\mathbf{x}) \leq 0\}] = EI(\mathbf{x}) \cdot P[G(\mathbf{x}) \leq 0]. \quad (33)$$

If there are  $N_C$  constraints, we should build  $N_C$  HK models for every constraint function, and the resulting constrained EI function is

$$CEI(\mathbf{x}) = EI(\mathbf{x}) \cdot \prod_{i=1}^{N_C} P[G_i(\mathbf{x}) \leq 0]. \quad (34)$$

Then the following unconstrained sub-optimization problem is formulated as

$$\mathbf{x} = \underset{\mathbf{x}_{low} \leq \mathbf{x} \leq \mathbf{x}_{up}}{\operatorname{argmax}} CEI(\mathbf{x}). \quad (35)$$

A hybrid method of combining GA, Hooke and Jeeves pattern search, and BFGS gradient-based method is used to solve the above sub-optimization problem to suggest the new sample point (Han 2016a, b), which is to be evaluate by hi-fi numerical analysis again. Please note only the hi-fi sample point is obtained here, since the above EI function is defined based on the uncertainty due to lack of a hi-fi sample. Therefore, if we need to adaptively select low-fi sample points as well, the EI function based on the uncertainty coming from the low-fi kriging model should be formulated, which is almost a blank area so far and will be done in Section 3.

## References

- Ackley DH (1987) A connectionist machine for genetic hill-climbing. Kluwer, Boston
- Alexandrov N, Dennis JE, Lewis RM, Torczon V (1998) A trust-region framework for managing the use of approximation models in optimization. *Struct Optim* 15(1):16–23
- Alexandrov N, Lewis RM, Gumbert CR, Green LL, Newman PA (2001) Approximation and model management in aerodynamic optimization with variable-fidelity models. *J Aircr* 38(6):1093–1101
- Bakr MH, Bandler JW, Madsen K, SØndergaard J (2001) An introduction to the space mapping technique. *Optim Eng* 2(4):369–384
- Benamara T, Breikopt P, Lepot I, Sainvitu C (2016) Adaptive infill sampling criterion for multi-fidelity optimization based on Gappy-POD. *Struct Multidisc Optim* 54(4):843–855
- Cai X, Qiu H, Gao L, Wei L, Shao X (2017) Adaptive radial-basis-function-based multifidelity metamodeling for expensive black-box problems. *AIAA J* 55(7):2424–2436
- Chang KJ, Haftka RT, Giles GL, Kao PJ (1993) Sensitivity-based scaling for approximation structural response. *J Aircr* 30(2):283–288
- Choi S, Alonso JJ, Kroo IM, Wintzer M (2004) Multi-fidelity design optimization of low-boom supersonic business jets. In: 10th AIAA/ISSMO Multidiscip Anal Optim Conf, AIAA paper 2004–4371, Albany, NY, US, 30 August–1 September
- Choi S, Alonso JJ, Kim S, Kroo IM (2009) Two-level multifidelity design optimization studies for supersonic jets. *J Aircr* 46(3):776–790
- Courrier N, Boucard PA, Soulier B (2016) Variable-fidelity modeling of structural analysis of assemblies. *J Glob Optim* 64(3):577–613
- Forrester AIJ, Keane AJ (2009) Recent advances in surrogate-based optimization. *Prog Aerosp Sci* 45(1–3):50–79
- Forrester AIJ, Sóbester A, Keane AJ (2007) Multi-fidelity optimization via surrogate modelling. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 463(2088):3251–3269
- Forrester AIJ, Sóbester A, Keane AJ (2008) *Engineering design via surrogate modelling— a practical guide*. Wiley, New York
- Gano SE, Renaud JE, Sanders B (2005) Hybrid variable fidelity optimization by using a kriging-based scaling function. *AIAA J* 43(11):2422–2430
- Giunta AA, Wojtkiewicz SF, Eldred MS (2003) Overview of modern design of experiments methods for computational simulations. In: 41st Aerosp Sci Meet Exhib, AIAA paper 2003–649, Reno, Nevada, 6–9 January
- Ha H, Oh S, Yee K (2014) Feasibility study of hierarchical kriging model in the design optimization process. *J Korean Soc Aeronaut Space Sci* 42(2):108–118
- Haftka RT (1991) Combining global and local approximations. *AIAA J* 29(9):1523–1525
- Han Z-H (2016a) SurroOpt: a generic surrogate-based optimization code for aerodynamic and multidisciplinary design. In: 30th Cong. Int. Counc. Aeronaut. Sci. ICAS, paper no. 2016–0281, Daejeon, Korea, 25–30 September
- Han Z-H (2016b) Kriging surrogate model and its application to design optimization: A review of recent progress. *Chin J Aeronaut* 37(11):3197–3225
- Han Z-H, Görtz S (2012) Hierarchical kriging model for variable-fidelity surrogate modeling. *AIAA J* 50(5):1285–1296
- Han Z-H, Zhang K-S (2012) Surrogate-based optimization. In: Roeva O (ed) *Real-World Applications of Genetic Algorithms*, InTech, pp. 343–362
- Han Z-H, He F, Song W-P, Qiao Z-D (2007) A preconditioned multigrid method for efficient simulation of three-dimensional compressible and incompressible flows. *Chin J Aeronaut* 20(4):289–296
- Han Z-H, Zimmermann R, Görtz S (2012) An alternative cokriging model for variable-fidelity surrogate modeling. *AIAA J* 50(5):1205–1210
- Han Z-H, Görtz S, Zimmermann R (2013) Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aerosp Sci Technol* 25:177–189
- Han Z-H, Zhang Y, Song C-X, Zhang K-S (2017) Weighted gradient-enhanced kriging for high-dimensional surrogate modeling and design optimization. *AIAA J* 55(12):4330–4346. <https://doi.org/10.2514/1.J055842>
- Hartman JK (1973) Some experiments in global optimization. *Nav Res Logist Q* 20:569–576
- Holland JH (1975) *Adaptation in natural and artificial systems*. Control & Artificial Intelligence University of Michigan Press, 6(2):126–137
- Huang D, Allen TT, Notz WI, Miller RA (2006) Sequential kriging optimization using multi-fidelity evaluations. *Struct Multidiscip Optim* 32(5):369–382
- Jo Y, Yi S, Choi S, Lee DJ, Choi DZ (2016) Adaptive variable-fidelity analysis and design using dynamic fidelity indicators. *AIAA J* 54(11):3564–3579
- Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. *J Glob Optim* 21(4):345–383
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *J Glob Optim* 13(4):455–492
- Journel AG, Huijbregts CJ (1978) *Mining geostatistics*. Academic Press, New York
- Kennedy J, Eberhart R (1995) *Particle swarm optimization*. In: *Proceedings of IEEE International Conference on Neural Networks*. 4:1942–1948
- Kennedy MC, O’Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87(1):1–13

- Kim Y, Lee S, Yee K, Rhee D (2017) High-to-low initial sample ratio of hierarchical kriging for film hole array optimization. *J Propuls Power*. <https://doi.org/10.2514/1.B36556>
- Koch PN, Simpson TW, Allen JK, Mistree F (1999) Statistical approximations for multidisciplinary design optimization: The problem of the size. *J Aircr* 36(1):275–286
- Koziel S, Leifsson L (2013) Surrogate-based aerodynamic shape optimization by variable-resolution models. *AIAA J* 51(1):94–106
- Koziel S, Leifsson L, Yang XS (2013) Surrogate-based optimization. In: Koziel S, Yang XS, Zhang QJ (eds) *Simulation-driven design optimization and modeling for microwave engineering*. Imperial College Press, London
- Krige DG (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. *J South Afr Inst Min Metall* 52(6):119–139
- Kulfan BM (2008) Universal parametric geometry representation method. *J Aircr* 45(1):142–158
- Leifsson L, Koziel S, Tesfahunegn YA (2016) Multiobjective aerodynamic optimization by variable-fidelity models and response surface surrogates. *AIAA J* 54(2):531–541
- Liu J, Han Z-H., Song W-P (2012) Comparison of infill sampling criteria in kriging-based aerodynamic optimization. In: 28th Cong Int Counc Aeronaut Sci ICAS, Brisbane, Australia, 23–38 September
- Liu J, Song W-P, Han Z-H, Zhang Y (2017) Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models. *Struct Multidiscip Optim* 55(3):925–943
- Martin JD, Simpson TW (2005) Use of kriging models to approximate deterministic computer models. *AIAA J* 43(4):853–863
- Matheron GM (1963) Principles of geostatistics. *Econ Geol* 58(8):1246–1266
- McDaniel WR, Ankenman BE (2000) A response surface test bed. *Qual Reliab Eng Int* 16:363–372
- Palar PS, Shimoyama K (2017) Multi-fidelity uncertainty analysis in CFD using hierarchical kriging. In: 35th AIAA Appl Aerodyn Conf, AIAA paper 2017–3261, Denver, Colorado, US, 5–9 June
- Park C, Haftka RT, Kim NH (2017) Remarks on multi-fidelity surrogates. *Struct Multidisc Optim* 55(3):1029–1050
- Queipo NV, Haftka RT, Shyy W, Goela T, Vaidyanathan R, Tucker PK (2005) Surrogate-based analysis and optimization. *Prog Aerosp Sci* 45(1):1–28
- Reisenthel PH, Allen TT (2014) Application of multifidelity expected improvement algorithms to aeroelastic design optimization. In: 10th AIAA Multidisc Des Optim Spec Conf, AIAA paper 2016–1542, San Diego, US, 7–10 January
- Robinson TD, Eldred MS, Willcox KE, Haimes R (2008) Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping. *AIAA J* 46(11):2814–2822
- Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. *Stat Sci* 4(4):409–423
- Shan S, Wang GG (2010) Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. *Struct Multidiscip Optim* 41(2): 219–241
- Simpson TW, Mauery TM, Korte JJ, Mistree F (2001) Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA J* 39(12):2233–2241
- Toal DJJ, Bressloff NW, Kean AJ (2008) Kriging hyperparameter tuning strategies. *AIAA J* 46(5):1240–1252
- Viana FAC, Simpson TW, Balabanov V, Toropov V (2014) Metamodeling in multidisciplinary design optimization: how far have we really come? *AIAA J* 52(4):670–690
- Wang GG, Shan S (2007) Review of metamodeling techniques in support of engineering design optimization. *J Mech Des* 129(4):370–380
- Xie F, Song W-P, Han Z-H (2008) Numerical study of high-resolution scheme based on preconditioning method. *J Aircr* 46(2):520–525
- Zhang Y, Han Z-H., Liu J, Song W-P (2015) Efficient variable-fidelity optimization applied to benchmark transonic airfoil design. In: 7th Asia-Pac int Symp Aerosp Technol, Cairns, Australia, 25–27 November
- Zhang Y, Han Z-H, Shi L-X, Song W-P (2016) Multi-round surrogate-based optimization for benchmark aerodynamic design problems. In: 54th AIAA Aerosp Sci Meet, AIAA paper 2016–1545, San Diego, US, 7–10 January