



# A bi-fidelity Bayesian optimization method for multi-objective optimization with a novel acquisition function

Kaiqin Xu<sup>1</sup> · Leshi Shu<sup>1</sup> · Linjun Zhong<sup>1</sup> · Ping Jiang<sup>1</sup> · Qi Zhou<sup>2</sup>

Received: 26 May 2022 / Revised: 19 January 2023 / Accepted: 25 January 2023 / Published online: 22 February 2023  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

In engineering design optimization, there are often multiple conflicting optimization objectives. Bayesian optimization (BO) is successfully applied in solving multi-objective optimization problems to reduce computational expense. However, the expensive expense associated with high-fidelity simulations has not been fully addressed. Combining the BO methods with the bi-fidelity surrogate model can further reduce expense by using the information of samples with different fidelities. In this paper, a bi-fidelity BO method for multi-objective optimization based on lower confidence bound function and the hierarchical Kriging model is proposed. In the proposed method, a novel bi-fidelity acquisition function is developed to guide the optimization process, in which a cost coefficient is adopted to balance the sampling cost and the information provided by the new sample. The proposed method quantifies the effect of samples with different fidelities for improving the quality of the Pareto set and fills the blank of the research domain in extending BO based on the lower confidence bound (LCB) function with bi-fidelity surrogate model for multi-objective optimization. Compared with the four state-of-the-art BO methods, the results show that the proposed method is able to obviously reduce the expense while obtaining high-quality Pareto solutions.

**Keywords** Bayesian optimization · Bi-fidelity surrogate model · Lower confidence bounds function · Multi-objective optimization

## List of symbols

$CR$	Cost ratio of samples with different fidelity
$F(x)$	Objective of multi-objective optimization
$f_l(x)$	Predicted mean of low-fidelity surrogate model
$f_i^j$	$j$ Th Pareto solution of the $i$ th objective
$f_{LCB}^l(x)$	LCB function at sample $x$
$f_{LCB}^l(x)$	Low-fidelity LCB function at sample $x$
$f_{LCB}^h(x)$	High-fidelity LCB function at sample $x$
$g_j(x)$	$j$ Th Constrain of multi-objective optimization
$H(p)$	Hypervolume indicator
$I(x)$	Improvement function at sample $x$

$I_{LCB}(x)$	Improvement of LCB function at sample $x$ beyond current Pareto set
$I_{LCB}^l(x)$	Improvement of low-fidelity LCB at sample $x$ beyond current Pareto set
$I_{LCB}^h(x)$	Improvement of high-fidelity LCB at sample $x$ beyond current Pareto set
$I(x, t)$	The novel improvement function with different fidelity
$N_l$	Quantity of the low-fidelity samples
$N_h$	Quantity of the high-fidelity samples
$R$	Covariance matrix of the hierarchical Kriging model
$R_l$	Covariance matrix of the low-fidelity Kriging model
$s_l^2(x)$	Mean-squared error of the low-fidelity Kriging model at unobserved points
$t$	Fidelity level
$X_j$	$j$ Th Pareto solution
$Z(\cdot), Z_l(\cdot)$	Gaussian random process
$\beta_{0,l}, \beta_0$	Scaling factor of trend model
$r$	Correlation vector of the hierarchical Kriging model
$r_l$	Correlation vector of the low-fidelity Kriging model

Responsible Editor: Ramin Bostanabad.

✉ Leshi Shu  
lsshu\_hust@163.com

<sup>1</sup> The State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, Hubei, People's Republic of China

<sup>2</sup> School of Aerospace Engineering, Huazhong University of Science and Technology, Wuhan 430074, Hubei, People's Republic of China

$\theta$	“Roughness” parameter
$\sigma^2$	Process variance

### Subscripts

$LCB$	Function associated with LCB
$l$	Low-fidelity value
$h$	High-fidelity value
$lb$	Lower bound of the value
$ub$	Upper bound of the value

### Superscripts

$h$	High-fidelity value
$l$	Low-fidelity value
$\mu$	Value associated with predicted mean

## 1 Introduction

Simulation models have been widely used in engineering design optimization. However, obtaining the Pareto solutions of multi-objective optimization problems usually requires many evaluations of objective functions. The evaluation processes of objective functions rely on time-consuming simulation models, which increases the burden of the optimization tasks. To solve this problem, surrogate models, such as Kriging models (Jerome et al. 1989; Liu et al. 2018; Williams et al. 2021), radial basis function models (Sóbestér et al. 2005; Zhou et al. 2017), support vector regression models (Shi et al. 2020), are commonly adopted to replace expensive simulations in engineering design optimization.

Bayesian optimization (Shahriari et al. 2016) is a typical surrogate-based optimization method, which is able to balance the exploitation and exploration by utilizing the uncertainty term of the surrogate model with an acquisition function (Jeong et al. 2005). Commonly used acquisition functions include the probability of improvement (PI) (Ruan et al. 2020), expected improvement (EI) (Jones 2001; Zhan et al. 2022), and lower confidence bound (LCB) (Srinivas et al. 2012; Zheng et al. 2016). In recent years, extending these single-objective BO methods to multi-objective optimization has also gained much attention. For example, Knowles (2006) proposed the ParEGO method, which extended the EI function to multi-objective optimization. Emmerich et al. (2011) combined the hypervolume improvement with the EI function, which clarifies the regions where the Pareto frontier can be improved by establishing the monotonicity properties. Zhan et al. (2017) expanded the EI function into a matrix form (EIM), where the elements of the matrix are the EI value corresponding to each objective. Shu et al. (2020) defined an acquisition function for multi-objective optimization problems, considering convergence and diversity of the Pareto frontier.

However, these methods only utilize the high-accuracy samples to build the surrogate model, which still expends

much computational cost (Chen et al. 2016). A potential way to further reduce the expense is adopting the multi-fidelity (MF) surrogate model (Zhang et al. 2018a, b), constructed based on a few costly high-fidelity (HF) samples and plenty of cheap low-fidelity (LF) samples. There are three main MF surrogate modeling methods: the scaling function method (Choi et al. 2009; Li et al. 2017), the space mapping method (Leifsson et al. 2015), and the Co-Kriging method (Kennedy et al. 2000; Perdikaris et al. 2015; Singh et al. 2017). Moreover, several new MF modeling methods have been proposed in the past few years, such as the method proposed by Eweis-Labolle et al. (2022), which utilizes latent-map Gaussian processes and outperforms the space mapping method. With above MF surrogate models gaining much attention, several multi-fidelity BO methods have been developed. For example, Zhang et al. (2018a, b) proposed a multi-fidelity method with an extension of Co-Kriging model, hierarchical Kriging (HK) model (Han et al. 2012), which avoids the construction of cross covariance in Co-Kriging. Jiang et al. (2019) proposed a new method to obtain the new sample and fidelity level, extended BO based on the LCB function with MF surrogate model. However, very limited work is accomplished for multi-objective optimization with BO based on the MF surrogate model. He et al. (2022) introduced a novel method for multi-objective problems, combining the HK model and EIM method. The blank of the research domain in extending BO based on the LCB function with MF surrogate model for multi-objective optimization still needs to fill. The bi-fidelity (BF) surrogate model is the surrogate model with two different fidelities, which is the special case of the MF surrogate model and needs to be considered first.

In this paper, a bi-fidelity BO method for multi-objective optimization is proposed, which utilizes the advantages of the bi-fidelity surrogate model and BO framework to further reduce the expense. In the method, a novel acquisition function is developed, in which a cost coefficient considering a cost ratio between different fidelity levels is introduced to balance the sampling cost and information of the new sample. The proposed method is compared with four state-of-the-art BO methods using several numerical examples and two engineering design optimization problems. The results show that the proposed method presents prominent performance in computational cost, especially for solving high-dimension problems.

The remainder of the paper is organized as follows. In Sec. 2, the technical background of the multi-objective optimization, HK model, and Bayesian optimization are recalled. Sec. 3 introduces the proposed method in detail. In Sec. 4, the results of comparative study between the proposed method and other state-of-the-art methods are presented. Sec. 5 summarizes this paper with conclusions and future work provided.

## 2 Background

### 2.1 Multi-objective optimization

Multi-objective optimization exists multiple conflicting objectives to optimize, and the optimization problem can be formulated as follows:

$$\begin{aligned} \min F(x) &= \{f_1(x), f_2(x), \dots, f_i(x), \dots, f_M(x)\} \\ \text{s.t. } g_j(x) &\leq 0, j = 1, 2, \dots, q \\ x_{lb} &\leq x \leq x_{ub} \end{aligned} \tag{1}$$

where the dimension of objective functions  $F(x)$  is  $M$ ,  $x = (x_1, x_2, \dots, x_N)^T$  denotes  $N$ -dimension design variable vector with lower/upper bounds of  $x_{lb}$  and  $x_{ub}$ ,  $g = (g_1, g_2, \dots, g_q)$  is the constrain vector that can be simple linear or nonlinear.

The multi-objective improvement functions are usually employed as acquisition functions in BO for multi-objective optimization. Euclidean distance improvement function (Keane 2006), maximin distance improvement function (Svenson et al. 2016), and hypervolume (HV) (Yang et al. 2019) improvement function are three state-of-the-art multi-objective improvement functions. The Euclidean distance improvement function can be expressed as follows:

$$I(x) = \min_{j=1}^k \sqrt{\sum_{i=1}^m (f_i^j - y_i(x))^2} \tag{2}$$

where  $f_i^j$  is the  $j$ th Pareto solution of the  $i$ th objective,  $y_i(x)$  is the prediction of the  $i$ th objective.

The maximin distance improvement function is expressed as follows:

$$I(x) = -\max_{j=1}^k \left[ \min_{i=1}^m (y_i(x) - f_i^j) \right] \tag{3}$$

Before the HV improvement function is introduced in detail, firstly the concept of the HV should be clarified. With a reference point dominated by the Pareto set, the HV of the dominated region can be measured as follows:

$$H(p) = \text{Volume}(y \in \mathbb{R}^M | p < y < R) \tag{4}$$

where  $p < y < R$  means the dominated region bounded by current Pareto  $p$  and the reference point,  $\text{Volume}(\cdot)$  is the HV indicator.

The HV improvement is the difference of the HV indicator between the current Pareto set and Pareto set with the next sample. It can be formulated as follows:

$$I(y) = H(p \cup y) - H(p) \tag{5}$$

where  $p \cup y$  represents the Pareto set obtained by the non-dominated-sort method after the new sample  $y$  added.  $I(y)$  is the HV improvement beyond the current Pareto set after

adding a new sample  $y$ . Figure 1 presents the 2D example of the HV improvement. The light gray area is the HV of current Pareto set and the dark gray area illustrates the  $I(y)$ . A higher HV improvement means a better-quality improvement beyond the current Pareto set.

More information about these improvement functions can be consulted in Svenson’s work (2016).

### 2.2 Hierarchical kriging (HK) model

Kriging model (Williams et al. 2006) is an interpolative surrogate model, which originated from the geostatistics and was applied to fit expensive simulation (Jerome et al. 1989).

HK model is a MF surrogate model based on the Kriging model, which takes LF prediction as an overall trend of the MF model with the HF function adopted to be a correction. To build an HF surrogate model for expensive function, first we build a Kriging model with LF samples, which is adopted thereafter to assist the construction of the MF model. The LF Kriging model is expressed as follows:

$$Y_l(x) = \beta_{0,l} + Z_l(x) \tag{6}$$

where  $\beta_{0,l}$  is the mean term and  $Z_l(x)$  is a random process with variance  $\sigma^2$  and zero mean. One of the most commonly used spatial correlation functions between the error term of two points  $x$  and  $x'$ , named the Gaussian correlation function, is given as follows:

$$\text{cov}(Z(x), Z(x')) = \sigma^2 \exp \left\{ -\sum_{i=1}^d \theta_i |x_i - x'_i|^2 \right\} \tag{7}$$

where  $\sigma^2$  is the variance of the  $Z(x)$ ,  $d$  is the dimension of design variables, and  $\theta = \{\theta_1, \theta_2, \dots, \theta_d\}$  is a “roughness” parameter associated with dimension  $i$ .

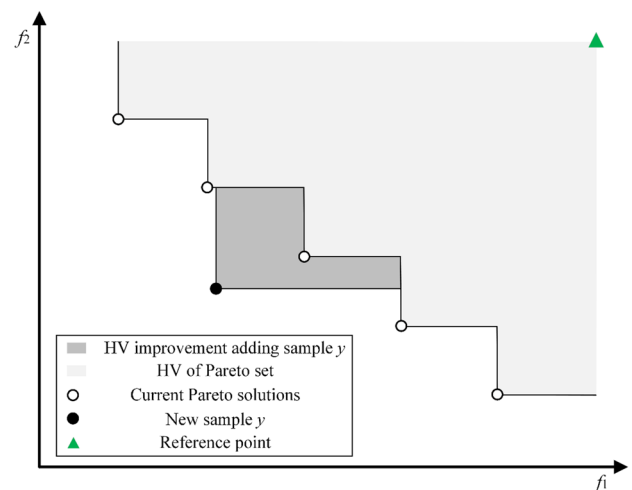


Fig.1 2-dimension example of the HV improvement

The prediction of mean and mean-squared error (MSE) from LF model at any unobserved points can be formulated as follows:

$$\hat{y}_l(x) = \beta_{0,l} + r_l^T(x)R_l^{-1}(f_l(x) - 1\beta_{0,l}) \tag{8}$$

$$s_l^2(x) = \sigma^2(1 - r_l^T(x)R_l^{-1}r_l(x) + \frac{(1 - 1^T R_l^{-1} r_l(x))^2}{1^T R_l^{-1} 1}) \tag{9}$$

respectively, where  $\beta_{0,l} = (1^T R_l^{-1} 1)^{-1} 1^T R_l^{-1} f_l(x)$ ,  $f_l(x)$  is the LF response of samples  $x$ ,  $1$  is a vector filled with 1,  $R_l$  is the covariance matrix with  $R_{l,mm} = \text{cov}(Z(x^m), Z(x^n))$  as elements, and  $r_l$  is a vector of correlation between the unobserved point and samples. Above hyperparameters  $\beta_{0,l}, \sigma^2, \theta_i$  are calculated by Maximum Likelihood Estimation (MLE).

Taking the predicted value of the LF model to construct the prior mean of the HF Kriging model, it will be built as follows:

$$Y(x) = \beta_0 \hat{y}_l(x) + Z(x) \tag{10}$$

where the  $\hat{y}_l(x)$  is the prediction of the LF model scaled by a constant factor  $\beta_0$ ,  $Z(x)$  is a random process with zero mean and a covariance of  $\text{cov}(Z(x), Z(x'))$ .

The prediction of mean value and mean-squared error (MSE) of HK model at the unobserved point can be formulated as follows:

$$\hat{y}(x) = \beta_0 \hat{y}_l(x) + r^T(x)R^{-1}(f_h(x) - \beta_0 f_l(x)) \tag{11}$$

$$MSE\{\hat{y}(x)\} = \sigma^2 \{1 - r^T(x)R^{-1}r(x) + [r^T(x)R^{-1}f_l(x) - \hat{y}_l(x)][f_l(x)^T R^{-1} f_l(x)]^{-1} [r^T(x)R^{-1}f_l(x) - \hat{y}_l(x)]^T\} \tag{12}$$

respectively, where  $\beta_0 = [f_l(x)^T R^{-1} f_l(x)]^{-1} f_l(x)^T R^{-1} f_h(x)$  is a scaling factor, and  $f_h(x)$  is the high-fidelity response of sample points  $x$ .

More details about hierarchical Kriging model are introduced in Han’s work (2012).

### 2.3 Bayesian optimization based on the LCB function

Bayesian optimization is a type of the surrogate model-based method for expensive optimization problems. The BO method utilizes the uncertainty and prediction information from surrogate model to quantify the improvement or probability of improvement between new sample and current optimum, which can balance the exploitation and exploration. EI and LCB functions are commonly used acquisition functions (Tran et al. 2019). With the number of objectives increased, due to the complexity of the non-dominated region, a piecewise integral approach is often

used to calculate the EI value, where the integral region is commonly resolved into some regular cells (Zhan et al. 2017). Hence, such improvement-based acquisition function as above multi-objective EI criteria becomes so complicated that the EI value is hard to be calculated. On the contrary, an outstanding advantage using LCB function is that it is no need to calculate such tedious piecewise integral like EI and PI function. This property economizes many computational resources from the calculation of high-dimension integrals.

Given the samples  $\{x_1, x_2, \dots, x_N\}$  and their associated response  $\{f_1, f_2, \dots, f_N\}$ , LCB function is formulated as follows:

$$f_{LCB}(x) = \hat{f}(x) - ks(x) \tag{13}$$

where  $\hat{f}(x)$  represents the prediction value of surrogate model,  $s(x)$  is the square root of uncertainty,  $k$  is a parameter balanced the exploitation and exploration. In multi-objective optimization, the form of LCB function will be a vector (Shu et al. 2020):

$$f_{LCB}(x) = [f_{LCB,1}(x), f_{LCB,2}(x), \dots, f_{LCB,M}(x)] \tag{14}$$

where  $f_{LCB,i}(x) = \hat{f}_i(x) - k_i s_i(x)$

where  $M$  is the number of objectives, each objective is fitted by the surrogate model respectively, and  $s_i(x)$  is the square root of the posterior predictive variance. A larger value of the coefficient  $k_i$  represents that the algorithm is encouraged to search the solution in the uncertain region, conversely a small value encourages exploitation in the current optimum region.

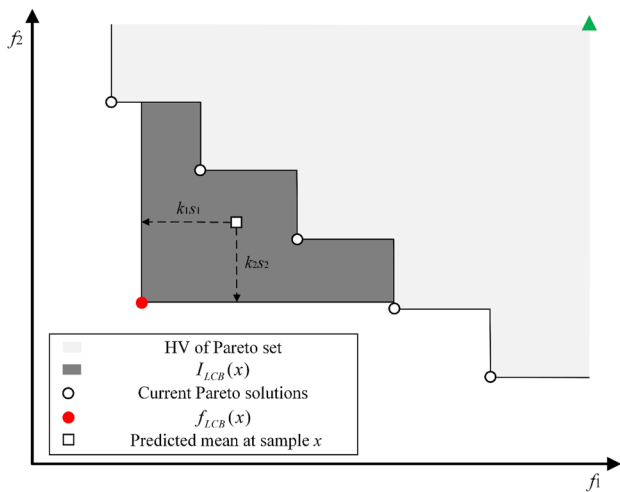
More information about EI, PI, and other acquisition functions can be consulted in following reference papers (Jones et al. 1998) (Ruan et al. 2020; Shahriari et al. 2016).

## 3 Proposed method

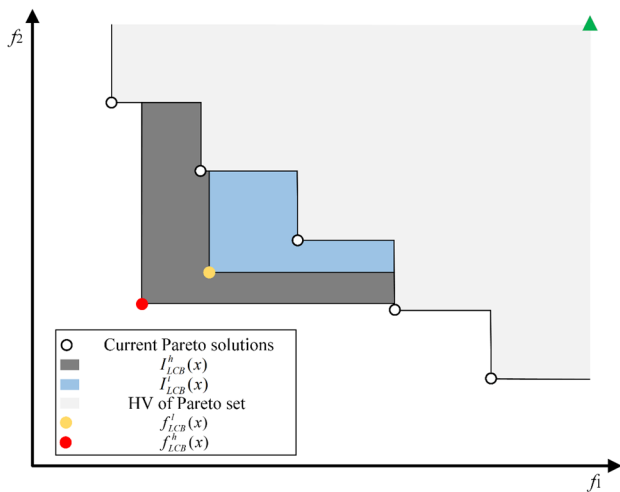
In this section, a bi-fidelity BO method for multi-objective optimization is proposed. Firstly, the cost and HV improvement after adding a new sample with different fidelity are analyzed, then the new acquisition function is introduced in detail. Finally, the terminal condition and procedure of the method are presented.

### 3.1 Predicted HV improvement after adding a new sample with different fidelities

The LCB function utilizes the predicted mean and uncertainty to measure the worth of each point in the unobserved area. The HV improvement function is employed to quantify the predicted quality improvement after adding a new



**Fig.2** 2-dimension example of Pareto set and predicted HV improvement



**Fig.3** the predicted HV improvement after adding HF and LF samples

sample in the proposed method, as presented in Eqs. (4) and (5). The HV improvement function can be formulated as follows:

$$I_{LCB}(x) = H(p \cup f_{LCB}(x)) - H(p) \tag{15}$$

where the  $f_{LCB}(x)$  is the LCB function defined in Eq. (14). Figure 2 presents the Pareto set and predicted HV improvement after adding a new sample ( $I_{LCB}(x)$ ) in 2D objective space. The red point in Fig. 2 represents the  $f_{LCB}(x)$  at sample  $x$ .

Figure 3 presents the HV improvement while the surrogate model has two fidelity levels and the HV improvement functions are formulated, respectively, as follows:

$$I_{LCB}^l(x) = H(p \cup f_{LCB}^l(x)) - H(p) \tag{16}$$

$$I_{LCB}^h(x) = H(p \cup f_{LCB}^h(x)) - H(p) \tag{17}$$

where the  $I_{LCB}^l(x)$  and  $I_{LCB}^h(x)$  denote the HV improvement after adding HF and LF samples, respectively,  $f_{LCB}^l(x)$  and  $f_{LCB}^h(x)$  are expressed as follows:

$$f_{LCB}^l(x) = [f_{LCB,1}^l(x), f_{LCB,2}^l(x), \dots, f_{LCB,M}^l(x)] \tag{18}$$

$$f_{LCB}^h(x) = [f_{LCB,1}^h(x), f_{LCB,2}^h(x), \dots, f_{LCB,M}^h(x)] \tag{19}$$

respectively, where

$$f_{LCB,i}^l(x) = \hat{f}_i(x) - k_i \beta_0 s_i^l(x) \tag{20}$$

$$f_{LCB,i}^h(x) = \hat{f}_i(x) - k_i s_i^h(x) \tag{21}$$

$s_i^h(x)$  is the error term arising from the lack of an HF sample in HK model,  $\beta_0 s_i^l(x)$  is the error of the LF model prediction because of the lack of a LF sample (Zhang et al. 2018a, b).

The light gray area in Fig. 3 is the HV of the current Pareto frontier. The blue and dark gray areas represent the HV improvement after adding the LF sample ( $I_{LCB}^l(x)$ ) and the HF sample ( $I_{LCB}^h(x)$ ) beyond current Pareto set, respectively. It should be noted that the  $I_{LCB}^h(x)$  includes the blue area, which denotes the HV indicator of the region between  $f_{LCB}^h(x)$  and the Pareto set. In most circumstances, the  $I_{LCB}^h(x)$  is larger than the corresponding  $I_{LCB}^l(x)$  due to the discrepant uncertainty term so that there is a conflict between computational cost and HV improvement of the sample with different fidelity.

### 3.2 The proposed acquisition function

As illustrated in the following Fig. 4, the predicted HV improvement can be partitioned into two parts, one from the predicted mean and the other from the uncertainty region.

The green region in Fig. 4 means the HV improvement of the predicted mean after adding a new sample ( $I_{LCB}^\mu(x)$ ) and it is expressed as follows:

$$I_{LCB}^\mu(x) = H(p \cup \hat{f}(x)) - H(p) \tag{22}$$

$$\hat{f}(x) = [\hat{f}_1(x), \hat{f}_2(x), \dots, \hat{f}_i(x), \dots, \hat{f}_M(x)] \tag{23}$$

where  $\hat{f}_i(x)$  means the prediction mean of each objective. The blue and dark gray areas represent the HV improvement



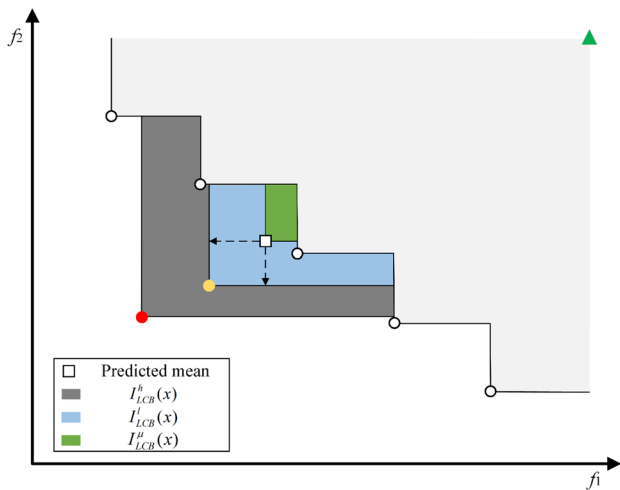


Fig.4 the predicted HV improvement of the predicted mean after adding new samples

as the same as shown in Fig. 3, and they can be expressed as  $I_{LCB}^l(x)$  and  $I_{LCB}^h(x)$  calculated by Eqs. (16) and (17).

To balance the cost and improvement of a new sample with different fidelity, here a novel improvement function with different fidelity is introduced, which can be expressed as follows:

$$I(x, t) = \begin{cases} (I_{LCB}^l(x) - I_{LCB}^\mu(x)) \times CR(t) + I_{LCB}^\mu(x) & t = 1 \\ (I_{LCB}^h(x) - I_{LCB}^\mu(x)) \times CR(t) + I_{LCB}^\mu(x) & t = 2 \end{cases} \quad (24)$$

where  $t = 1$  and  $t = 2$  represent the improvement function with low fidelity and high fidelity, respectively,  $CR(t)$  is the cost coefficient that is formulated as follows:

$$CR(t) = \begin{cases} c & t = 1 \\ 1 & t = 2 \end{cases} \quad (25)$$

where  $c \geq 1$ , which is the ratio of cost between HF and LF simulations. With the cost coefficient applied, the HV improvement with LF sample is expanded proportionally by controlling the computational cost to the same level for different fidelities. Thus, the HV improvement with different fidelities can be reasonably compared and choose the most appropriate one. When the cost ratio  $c$  is larger than one, it is invidious to compare  $I_{LCB}^l(x)$  with corresponding  $I_{LCB}^h(x)$  after multiplying the cost coefficient  $CR(t)$  because the predicted mean with all fidelity levels has no discrepancy between each other in BF surrogate model. Hence, the effect of the cost coefficient on the HV improvement of the predicted mean should be eliminated.

As the cost coefficient  $CR(t)$  with HF level is equal to 1, the term where  $t = 2$  in Eq. (24) can be regarded as follows:

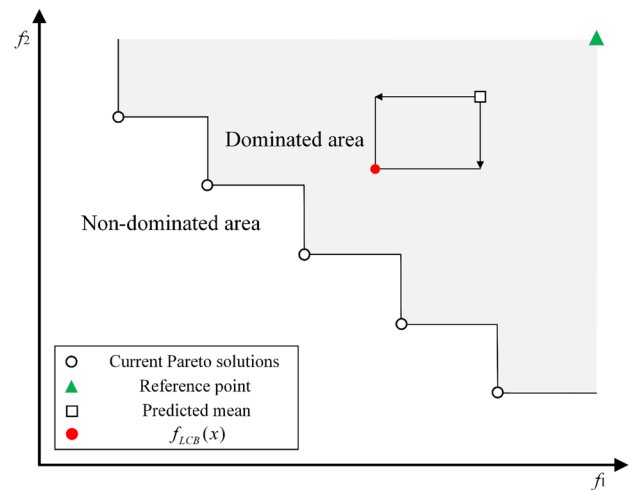


Fig.5 the LCB function located in dominated region

$$(I_{LCB}^h(x) - I_{LCB}^\mu(x)) \times 1 + I_{LCB}^\mu(x) = I_{LCB}^h(x) \quad (26)$$

and the  $I(x, t)$  can be simplified as follows:

$$I(x, t) = \begin{cases} (I_{LCB}^l(x) - I_{LCB}^\mu(x)) \times CR(t) + I_{LCB}^\mu(x) & t = 1 \\ I_{LCB}^h(x) & t = 2 \end{cases} \quad (27)$$

When the  $f_{LCB}(x)$  locates in dominated region, all of terms in  $I(x, t)$  is equal to zero and it cannot estimate which fidelity level is more appropriate. On the other hand, the algorithm of optimizing the acquisition function, such as genetic algorithm (GA) (Jin 2011), adopts random sampling to obtain the initial population, whereas Pareto solutions usually only occupy a narrow region relative to the whole design space. Therefore, GA may not be able to acquire a solution located in the area of Pareto set during the searching process, which leads to the decline of search efficiency in optimization process. Figure 5 illustrates the situation that the LCB function locates in dominated region of Pareto frontier.

To address this issue, we refer to our previous work (Shu et al. 2020), a novel acquisition function is defined as follows:

$$a(x, t) = \begin{cases} I(x, t), & \text{where } f_{LCB} \text{ is non-dominated} \\ - \min_{i=1,2,\dots,q} (\|f_{LCB}(x) - f(X_i)\|_2), & \text{where } f_{LCB} \text{ is dominated} \end{cases} \quad (28)$$

where  $I(x, t)$  is Eq. (27),  $X_i$  ( $i = 1, 2, \dots, q$ ) denotes the current Pareto solution,  $\|f_{LCB}(x) - f(X_i)\|_2$  means the Euclidean distance between  $f_{LCB}(x)$  and the Pareto solution  $X_i$ . The acquisition function makes an encouragement to help GA find the solution close to the Pareto frontier so that the optimization efficiency can be improved.

### 3.3 Procedure and termination condition

The proposed method, based on BF model and LCB function for multi-objective optimization (BF-MOLCB), follows a process as illustrated in Fig. 6. In the study, the Latin hypercube sampling (LHS) (Wang 2003) policy is applied to generate two initial sample sets for two fidelity levels. With the corresponding responses of initial samples evaluated, the current Pareto frontier can be obtained by the non-dominated sorting method (Deb et al. 2002).

The procedure of the BF-MOLCB is summarized in the flowchart as shown in Fig. 6, includes

**Step 1** Generate two sample sets for LF and HF model by LHS.

**Step 2** Evaluate the true responses of the corresponding samples.

**Step 3** Construct the LF Kriging models for each objective utilizing the DACE toolbox (Lophaven et al. 2002).

**Step 4** Construct the hierarchical Kriging model based on LF Kriging model and HF samples. The initial value of the hyperparameter  $\theta$  is set to be 1. The region of  $\theta$  is set to be  $[10^{-6}, 10^3]$ .

**Step 5** Obtain the current Pareto frontier by the non-dominated sorting method.

**Step 6** Judge the termination criterion, if it is not satisfied, go to step 7. Otherwise, the method is stopped and output the solution.

**Step 7** Maximize the proposed acquisition function to obtain a new sample, turn back to step 2.

A stop criterion is adopted to terminate the optimization process after the following two conditions are satisfied:

- (1) After the algorithm finds at least 20 solutions, thus, the Pareto set will provide enough choice to the decision maker; Or the computational expense gets the maximum budget, defined as 200 in this paper.
- (2) The discrepancy of the quality metrics in two adjacent iterations represents the variation of the current Pareto frontier. Hence, when the ratio of the difference and the current HV value is less than a specific value (e.g., 0.1% in this paper), the algorithm will be terminated, which is defined as follows:

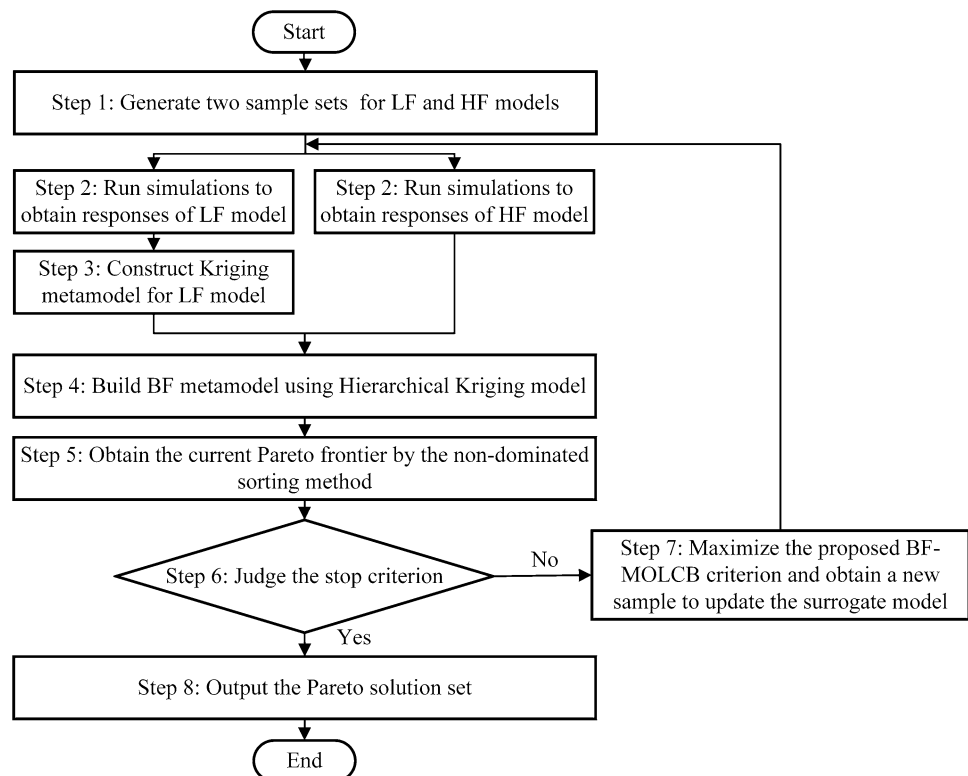
$$HV_T = \frac{HV_i - HV_{i-1}}{HV_i} \quad (29)$$

where the  $HV_i$  means the HV value of the  $i$ th iteration (Shu et al. 2020).

## 4 Examples and results analysis

In this section, four numerical examples and an engineering design problem are adopted to test the efficiency and applicability of the proposed method. The formulations and true

**Fig. 6** the flowchart of the proposed method



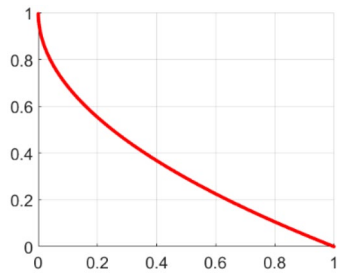
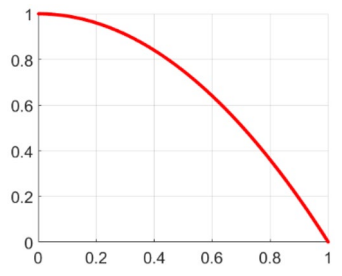
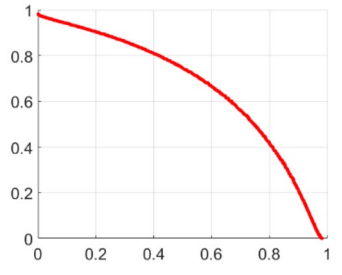
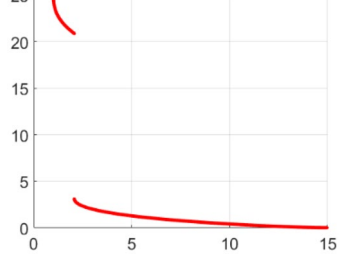
Pareto solutions of the four numerical examples (Shu et al. 2019) are summarized in Table 1. There are two single-fidelity BO methods (Sun et al. 2021; Zhan et al. 2017) only utilizing HF simulations and other two bi-fidelity BO methods (He et al. 2022) to compare with the proposed BF-MOLCB method: (1) Euclidean distance-based EIM method (EIMe), (2) Euclidean distance-based LCBM method (LCBMe), (3) Hypervolume-based VFEIM method (VFEIMh), and (4) VFEMHVI method.

### 4.1 Numerical examples

ZDT1, ZDT2, FON, and POL examples are solved by above methods and the results are compared in this part. The cost of the initial sample is set as  $10 \cdot n$  ( $n$  is the dimension of design variable), where the cost of the bi-fidelity method is formulated by an equivalent calculation as follows:

$$COST = \frac{N_l}{CR} + N_h \tag{30}$$

**Table 1** The formulation of the four numerical examples

Test function	Formulation	True Pareto
ZDT1	minimize $f_1(x) = x_1$ HF : $f_2(x) = h(x) \times g(x)$ LF : $f_2(x) = (0.8g(x) - 0.2) \times (1.2h(x) + 0.2)$ where $g(x) = 1 + \frac{9}{n-1} \sum_{i=2}^n x_i$ $h(x) = 1 - \sqrt{f_1(x)/g(x)}$ $n = 3$ $0 \leq x_i \leq 1, i = 1, \dots, n$	
ZDT2	minimize $f_1(x) = x_1$ HF : $f_2(x) = h(x) \times g(x)$ LF : $f_2(x) = (1.1h(x) - 0.1) \times (0.9g(x) + 1.1)$ where $g(x) = 1 + \frac{9}{n-1} \sum_{i=2}^n x_i$ $h(x) = 1 - (f_1(x)/g(x))^2$ $n = 3, 6, 8, 10$ $0 \leq x_i \leq 1, i = 1, \dots, n$	
FON	minimize $f_1(x) = 1 - \exp(-\sum_{i=1}^3 (x_i - \frac{1}{\sqrt{3}})^2)$ HF : $f_2(x) = 1 - \exp(-\sum_{i=1}^3 (x_i + \frac{1}{\sqrt{3}})^2)$ LF : $f_2(x) = (1 - \exp(-(x_1 + 0.5)^2 - (x_2 + 0.55)^2 - (x_3 + 0.6)^2))(1.1 + 0.25 \sin x_1)$ $-4 \leq x_i \leq 4, i = 1, 2, 3$	
POL	minimize HF : $f_1(x) = [1 + (A_1 - B_1)^2 + (A_2 - B_2)^2]$ LF : $f_1(x) = [1 + (0.9A_1 - 1.2B_1)^2 + 0.9(1.2A_2 - 0.9B_2)^2]$ $f_2(x) = (x_1 + 3)^2 + (x_2 + 1)^2$ where $A_1 = \sin 2 - 2 \cos 1 + 0.5 \sin 1 - 1.5 \cos 2$ $A_2 = 2 \sin 2 - \cos 1 + 1.5 \sin 1 - 0.5 \cos 2$ $B_1 = 0.5 \sin x_1 + \sin x_2 - 2 \cos x_1 - 1.5 \cos x_2$ $B_2 = 1.5 \sin x_1 + 2 \sin x_2 - \cos x_1 - 0.5 \cos x_2$ $-\pi \leq x_i \leq \pi, i = 1, 2$	



**Table 2** Reference points to calculate HV indicator in each test function

Test function	Reference point
ZDT1	[11,11]
ZDT2	[11,11]
FON	[1.4,1.4]
POL	[70,70]

where  $N_l/N_h$  denotes the number of LF/HF samples. For initial sampling,  $N_l$  and  $N_h$  are set as  $5*n*CR$  and  $5*n$ , respectively. The  $CR$  is the ratio of cost between LF and HF samples, which is set to be 4 in the following numerical test problems. To obtain the HV indicator, a reference point is set in each test function, as the shown in Table 2.

Firstly, a detailed comparison between these methods is demonstrated using the example ZDT1 (dimension  $n = 3$ ). Each method is run 10 times, where the statistics of the HV indicator and computational expense (COST) are recorded. The comparison results of HV and COST using different methods are summarized in Table 3, including the mean, median value, and STDs of HV and COST from 10 runs. The best indicators are shown in bold in the Table. Figure 7

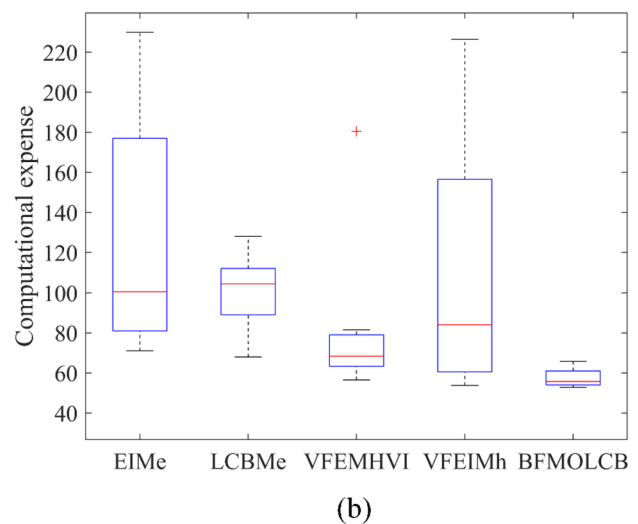
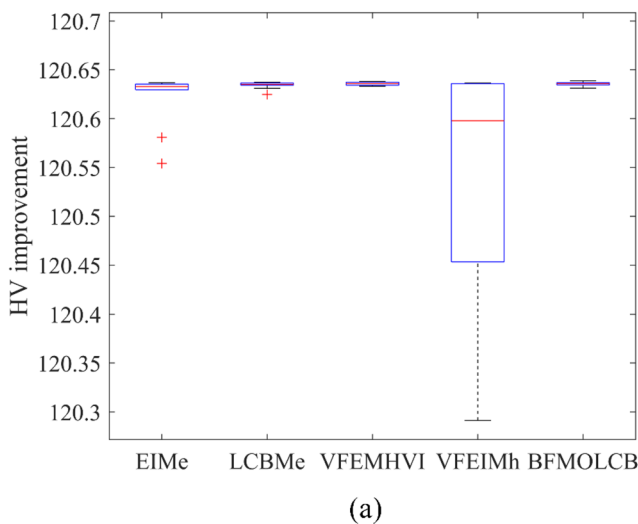
illustrates the box charts of the HV and COST obtained from different methods in ZDT1 problem.

As the HV indicators presented in Table 3, the LCBMe method, the VFEMHVI method, and proposed method have the most ideal Pareto frontier approximation quality according to the HV indicator. Meanwhile, the minimum mean and STD value of the COST demonstrate that the proposed method enables to have an ideal performance with lowest computational expense simultaneously, in which the COST means of the proposed method have 55.67%, 43.74%, 46.03%, and 28.13% decline over EIMe, LCBMe, VFEIMh, and VFEMHVI, respectively. Figure 7 illustrates the comparison over different methods intuitively. It can be seen that COST indicator of the proposed method has the most inconspicuous abnormal value in this box chart.

Table 4 presents the average computational expense required to obtain different number of Pareto solutions in ZDT1, further comparing the efficiency of finding Pareto solutions over these methods. As shown in Table 4, the proposed method always has the lowest value of the average computational expense. It is important to find enough Pareto solutions to provide adequate options for designers.

**Table 3** The compared results of different methods for ZDT1 problem

Test	Method	HV			COST		
		Median	Mean	Std	Median	Mean	Std
ZDT1 ( $n=3$ )	EIMe	120.6327	120.6204	0.0287	100.5000	128.8000	61.8310
	LCBMe	120.6352	120.6342	0.0037	104.5000	101.5000	18.2041
	VFEIMh	120.5979	120.5301	0.1370	84.0000	105.8000	59.1117
	VFEMHVI	<b>120.6359</b>	<b>120.6357</b>	<b>0.0017</b>	68.3750	79.4500	36.2950
	BF-MOLCB	120.6358	120.6355	0.0021	<b>55.7500</b>	<b>57.1500</b>	<b>4.1820</b>



**Fig.7** the box chart of the HV and COST. **a** HV, **b** COST

**Table 4** The average computational expense required to obtain different number of Pareto solutions for ZDT1 problem

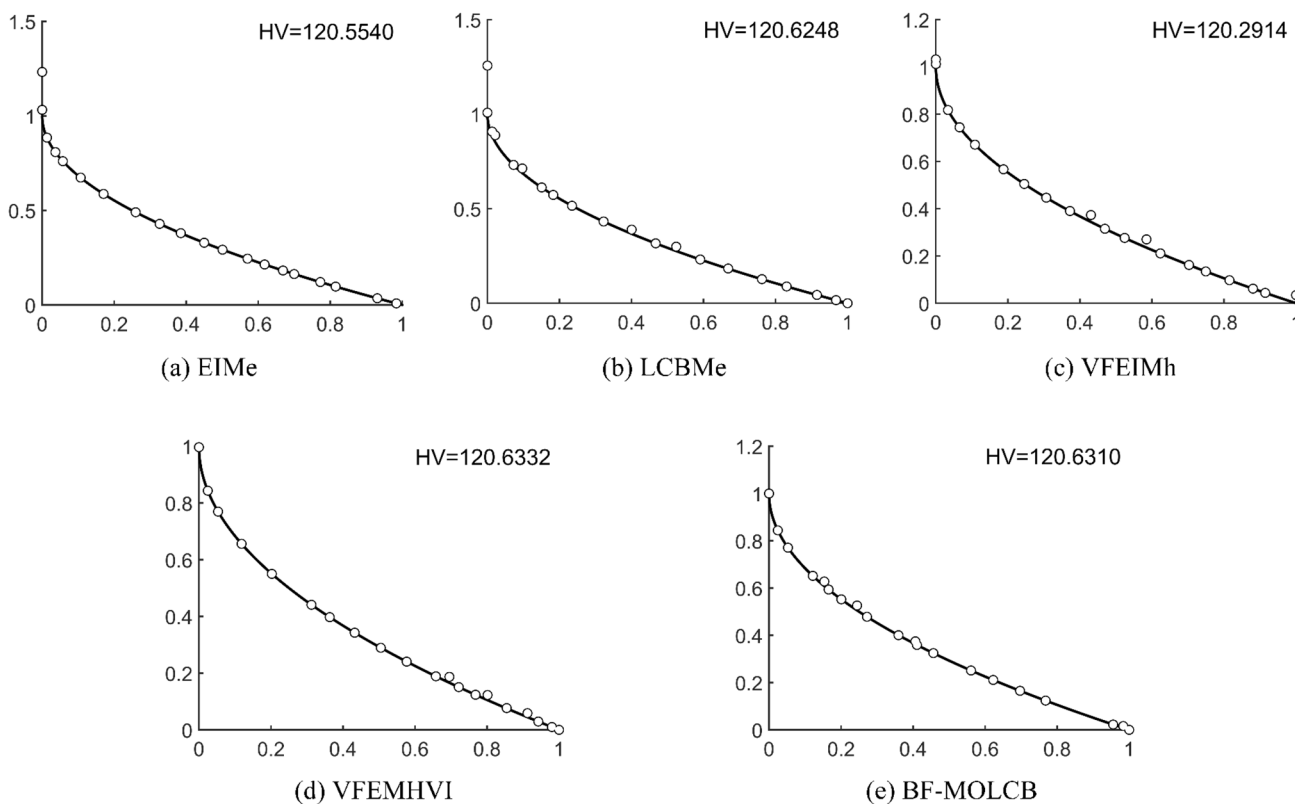
Pareto solutions number	EIMe	LCBMe	VFEIMh	VFEMHVI	BF-MOLCB
10	45.00	47.80	49.25	42.83	<b>42.33</b>
15	61.50	72.60	67.40	53.35	<b>49.95</b>
20	128.80	101.50	105.70	79.45	<b>57.15</b>

The quality of Pareto frontier is also a considerable metric in multi-objective optimization problems, which directly determines the information provided for designers. Hence, it is necessary to compare the quality of Pareto frontier from different methods. There are several figures illustrating the Pareto frontier with the lowest HV indicator over 10 runs, as shown in Fig. 8. In the ZDT1 problem, the most ideal Pareto frontier are still obtained by the VFEMHVI method and the proposed method, but the advantage is not evident enough, in which the difference of the HV means between the proposed method and VFEMHVI is lower than 0.01. Other methods also have nonnegligible competitiveness in this case.

The comparisons of remaining examples are summarized in Tables 5 and 6. All these methods present satisfactory

performance in obtaining the Pareto solutions on FON and ZDT2 problems with  $n = 3$  and the proposed method still has the lowest computational cost and ideal HV indicators. A very diverse phenomenon appears in POL problem. Even though the discrepancy of the cost among different methods is not significant enough, the cost of the proposed method is slightly higher than other single-fidelity methods and VFEIMh method. However, the proposed method has the highest HV indicator, which means it enables to obtain a Pareto set with better quality than other methods. Similar to Fig. 8, Fig. 9 illustrates the Pareto frontier with the lowest HV indicator in 10 runs on POL problem. It can be seen that the Pareto solutions obtained by the proposed method uniformly distribute on the true Pareto frontier, which is obviously better than VFEIMh and other single-fidelity methods. The VFEMHVI method has the same situation but its expense is 3.43% higher than the proposed method.

Additionally, there are several high-dimension test problems (ZDT2 problem with dimension  $n = 68, 10$ ) for measuring the ability to solve the optimization problems under more complex cases. The proposed method reveals a prominent performance in these problems, as shown in Table 6. The “/” symbol denotes that the termination criterion cannot be satisfied when the computational expense reaches the maximum budget 200. For



**Fig.8** Pareto frontier with the lowest HV indicator in 10 runs on ZDT1 problem

**Table 5** The compared results of different methods for ZDT2 ( $n=3$ ), POL and FON problems

Test	Method	HV			COST		
		Median	Mean	Std	Median	Mean	Std
ZDT2 ( $n=3$ )	EIMe	120.3054	120.3057	0.0012	66.0000	66.6000	5.7388
	LCBMe	120.3064	120.3062	0.0012	67.0000	68.4000	6.5013
	VFEIMh	120.3067	120.3064	0.0023	62.2500	64.5000	8.2563
	VFEMHVI	<b>120.3084</b>	<b>120.3085</b>	<b>0.0009</b>	64.1250	65.4000	8.4502
	BF-MOLCB	120.3078	120.3078	0.0012	<b>59.2500</b>	<b>59.9500</b>	<b>3.5645</b>
POL	EIMe	4789.3157	4788.6301	3.2840	47.5000	<b>47.2000</b>	2.1499
	LCBMe	4788.6244	4788.7998	2.4318	47.0000	47.5000	1.9003
	VFEIMh	4789.5654	4790.4894	2.0784	<b>46.7500</b>	47.3250	2.2362
	VFEMHVI	4793.8204	4793.8453	<b>0.1653</b>	53.5000	53.5250	2.1488
	BF-MOLCB	<b>4793.9483</b>	<b>4793.8788</b>	0.2598	51.7500	51.7500	<b>1.2802</b>
FON	EIMe	1.2609	1.2596	0.0073	69.0000	70.0000	7.6012
	LCBMe	1.2629	1.2614	0.0102	68.0000	68.3000	<b>3.4335</b>
	VFEIMh	1.2646	1.2634	0.0073	69.2500	70.9000	7.3817
	VFEMHVI	1.2707	1.2711	0.0023	71.2500	70.4500	5.0329
	BF-MOLCB	<b>1.2728</b>	<b>1.2732</b>	<b>0.0011</b>	<b>66.6250</b>	<b>66.9750</b>	4.1356

**Table 6** The compared results of different methods for high-dimension ZDT2 problems

Test	Method	HV			COST		
		Median	Mean	Std	Median	Mean	Std
ZDT2 ( $n=6$ )	EIMe	120.2893	120.2895	0.0065	210.5000	203.4000	30.6311
	LCBMe	120.2841	120.2847	0.0065	260.0000	258.1000	6.0083
	VFEIMh	120.2845	120.2481	0.1161	260.0000	241.0250	31.7191
	VFEMHVI	<b>120.3073</b>	<b>120.3073</b>	<b>0.0014</b>	134.8750	147.4500	27.2832
	BF-MOLCB	120.2990	120.2966	0.0088	<b>91.2500</b>	<b>95.8250</b>	11.5205
ZDT2 ( $n=8$ )	EIMe	/	/	/	/	/	/
	LCBMe	/	/	/	/	/	/
	VFEIMh	120.2775	120.2334	0.1221	280.0000	277.0000	11.1586
	VFEMHVI	<b>120.3015</b>	<b>120.3004</b>	<b>0.0047</b>	256.0000	248.2000	37.1296
	BF-MOLCB	120.2962	120.2949	0.0060	<b>121.8750</b>	<b>147.5250</b>	47.0816
ZDT2 ( $n=10$ )	Others	/					
	BF-MOLCB	<b>120.2922</b>	<b>120.2878</b>	<b>0.0136</b>	<b>212.8750</b>	<b>224.8500</b>	<b>71.6823</b>

these high-dimension problems, the proposed method always has the lowest computational expense and ideal quality of Pareto frontiers. In the ZDT2 problem with  $n = 6$ , single-fidelity methods sometimes fail to obtain 20 Pareto solutions before the maximum budget reached. The performance of the VFEMHVI and the proposed BF-MOLCB method has notable improvement, in which the COST means of the proposed method have 52.89%, 62.87%, 60.24%, and 34.70% decline over EIMe, LCBMe, VFEIMh, and VFEMHVI, respectively. The VFEMHVI method can obtain the largest HV indicator. The proposed method can also obtain an ideal Pareto frontier (difference of HV between the VFEMHVI and the proposed method is less than 0.01% of the whole HV), meanwhile the computational expense of the proposed method is much lower

than the VFEMHVI. In the test problem with  $n = 8, 10$ , all single-fidelity methods cannot meet the termination criterion until the maximum budget 200 reached. The cost of the VFEIMh and VFEMHVI method is also unsatisfactory. Especially in ZDT2 with  $n = 10$ , only the proposed method can still obtain a satisfactory Pareto frontier with controllable computational expense.

Comparing these test problems in terms of the dimensions, low-dimension problems ( $n = 2, 3$ ) have a smaller difference in the computational cost between these methods, whereas the corresponding discrepancy in high-dimension problems is much more noteworthy. Therefore, the bi-fidelity surrogate models have better performance than single-fidelity models in approximating the complex problems, which confirms the research significance of utilizing bi-fidelity

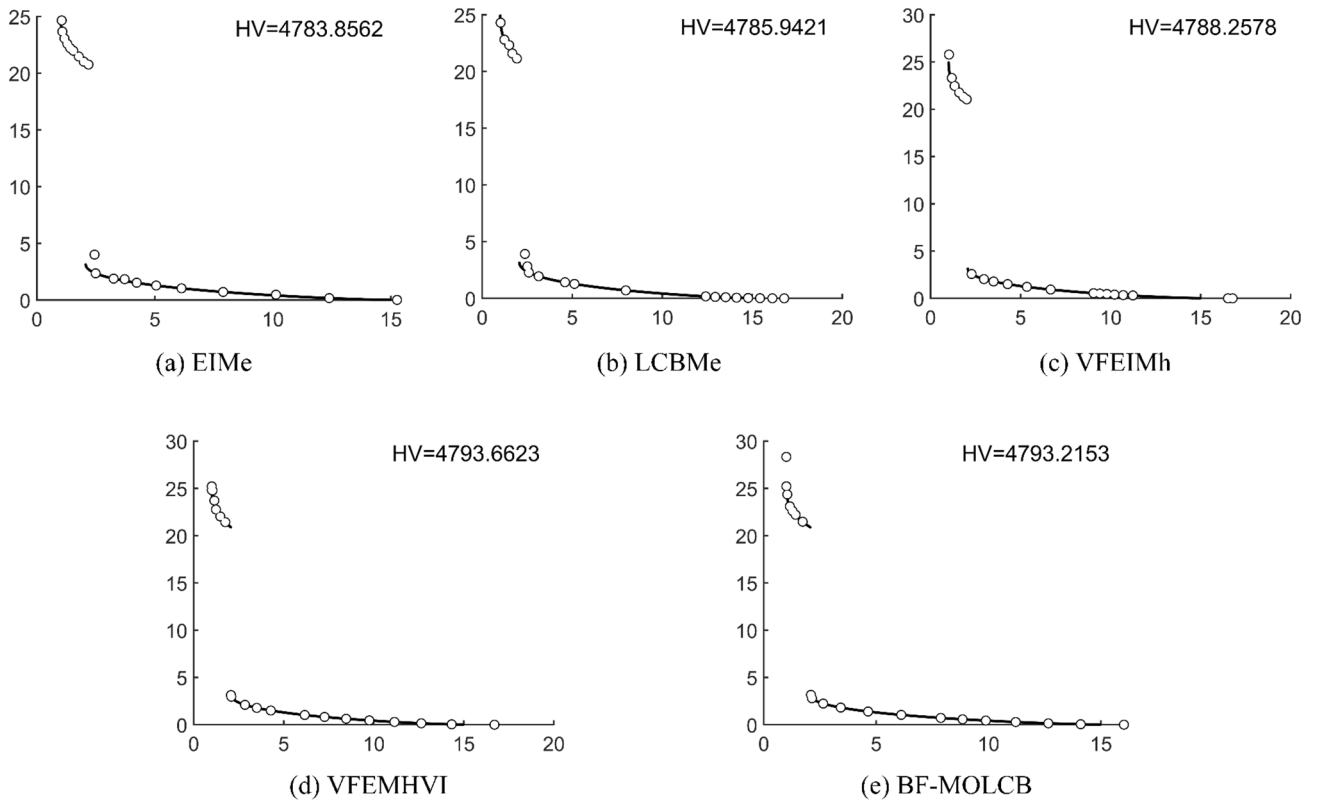


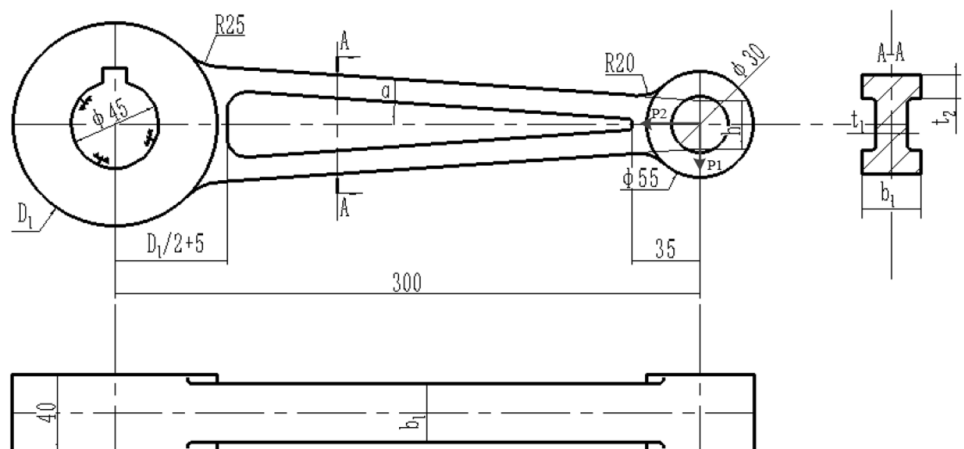
Fig.9 Pareto frontier with the lowest HV indicator in the 10 runs on POL problem

models in BO to solve the multi-objective optimization problems.

### 4.2 Engineering case

In this section, two engineering examples, torque arm and honeycomb structure vibration isolator design optimization, are adopted to verify the performance of the proposed method. The torque arm is fixed at the left end. The load on the torque arm,  $P_1 = 8.0kN$  and  $P_2 = 4.0kN$ , is exerted at the right end (Fig. 10).

Fig.10 Parameterization of the torque arm (Shu et al. 2020)



Six design variables are applied for the torque in this example, including  $\alpha, b_1, D_1, h, t_1, t_2$ , and  $t_2$ . The optimization problem is expressed as follows:

$$\begin{aligned} \min f_1 &= V(\alpha, b_1, D_1, h, t_1, t_2) \\ \min f_2 &= \max\_d(\alpha, b_1, D_1, h, t_1, t_2) \end{aligned} \tag{31}$$

where  $3 \text{ deg} \leq \alpha \leq 4.5 \text{ deg}; 25\text{mm} \leq b_1 \leq 35\text{mm}$   
 $90\text{mm} \leq D_1 \leq 120\text{mm}; 20\text{mm} \leq h \leq 30\text{mm}$   
 $12\text{mm} \leq t_1 \leq 22\text{mm}; 8\text{mm} \leq t_2 \leq 12\text{mm}$

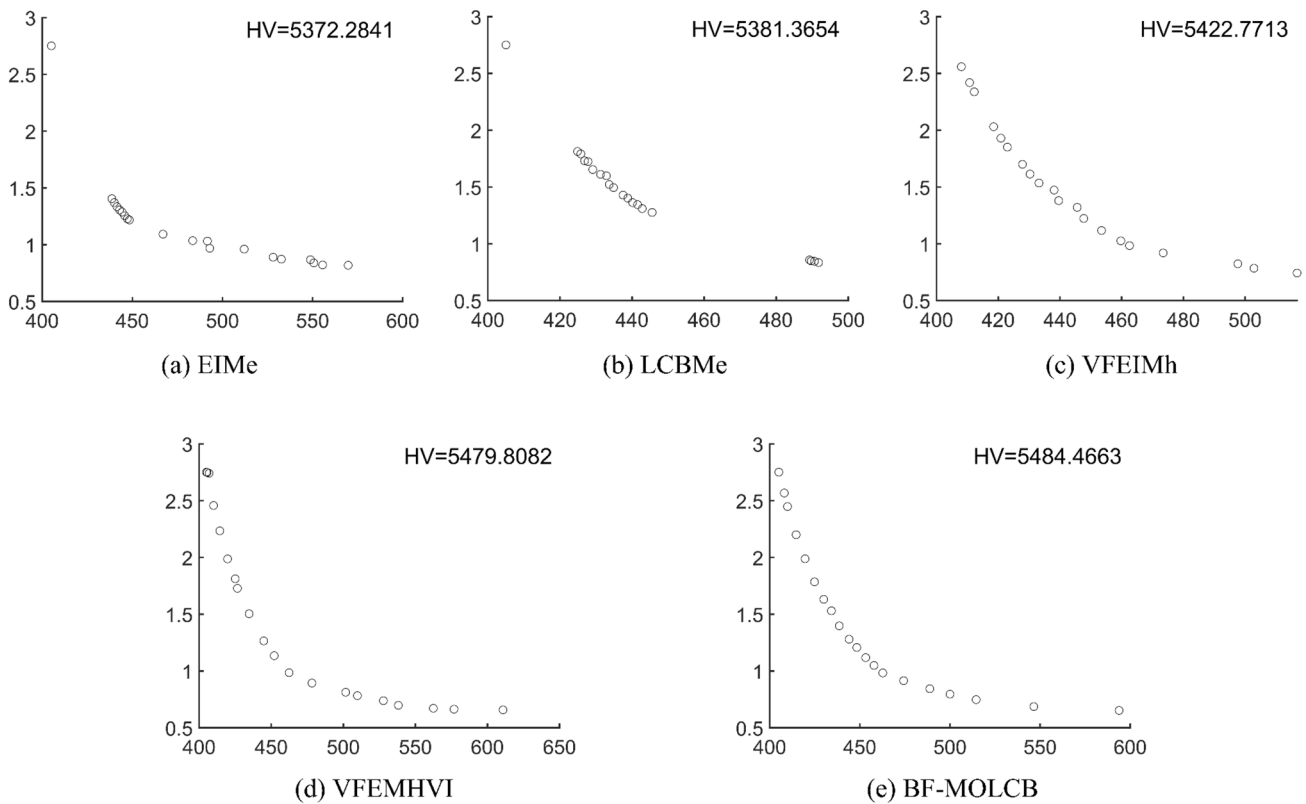
where  $V$  is the total volume,  $\max\_d$  is the maximum displacement of the torque arm, and the Young's modulus is 200 GPa and Poisson's ratio is 0.3. In this example, the

reference point is [1000;10] and the cost ratio between high- and low-fidelity samples is 4.

Each method is run 10 times, where the average of the HV indicator and computational expense is recorded, as the shown in Table 7. All of these methods can meet the termination condition; thereinto, the proposed method has the best performance in the quality of Pareto solutions due to the highest HV indicator. However, the cost of the proposed method is higher than other methods. The reason why this phenomenon appears is that the non-dominated solutions obtained by the single-fidelity methods and the VFEIMh method early on have satisfied the termination criterion, which leads the method cannot obtain an ideal Pareto set. There are some figures in Fig. 11 used to show the Pareto frontier with the lowest HV indicator in 10 runs

**Table 7** The comparison with different methods for the torque arm problem

Test	Method	HV			COST		
		Median	Mean	Std	Median	Mean	Std
Torque arm	EIMe	5413.7313	5415.6613	28.2935	<b>79.0000</b>	<b>78.6000</b>	6.8346
	LCBMe	5421.6972	5417.3965	23.0653	82.0000	81.3000	4.7621
	VFEIMh	5473.2851	5469.4519	18.4871	86.7500	87.1250	<b>2.8874</b>
	VFEMHVI	5484.5960	5484.4753	2.6605	90.5000	91.2000	5.4439
	BF-MOLCB	<b>5488.8273</b>	<b>5488.2470</b>	<b>1.6061</b>	92.7500	95.9250	7.9696



**Fig.11** Pareto frontier with the lowest HV indicator in 10 runs on the torque arm problem

on this torque arm problem. Obviously, the Pareto solutions obtained from proposed method distribute uniformly on the true Pareto frontier, distinctly better than other single-fidelity methods and VFEIMh.

To compare the proposed method with other methods more fairly, the same budget is adopted to compare the quality of the Pareto solution set, which is set 100 in this study. Each method works 10 times repeatedly, terminated when the cost of optimization process reaches 100, and summarizes the median, mean, and std value of the HV indicator. The results are shown in Table 8. It can be seen that the median and mean value of the HV indicator obtained by the proposed method are still better than other methods;

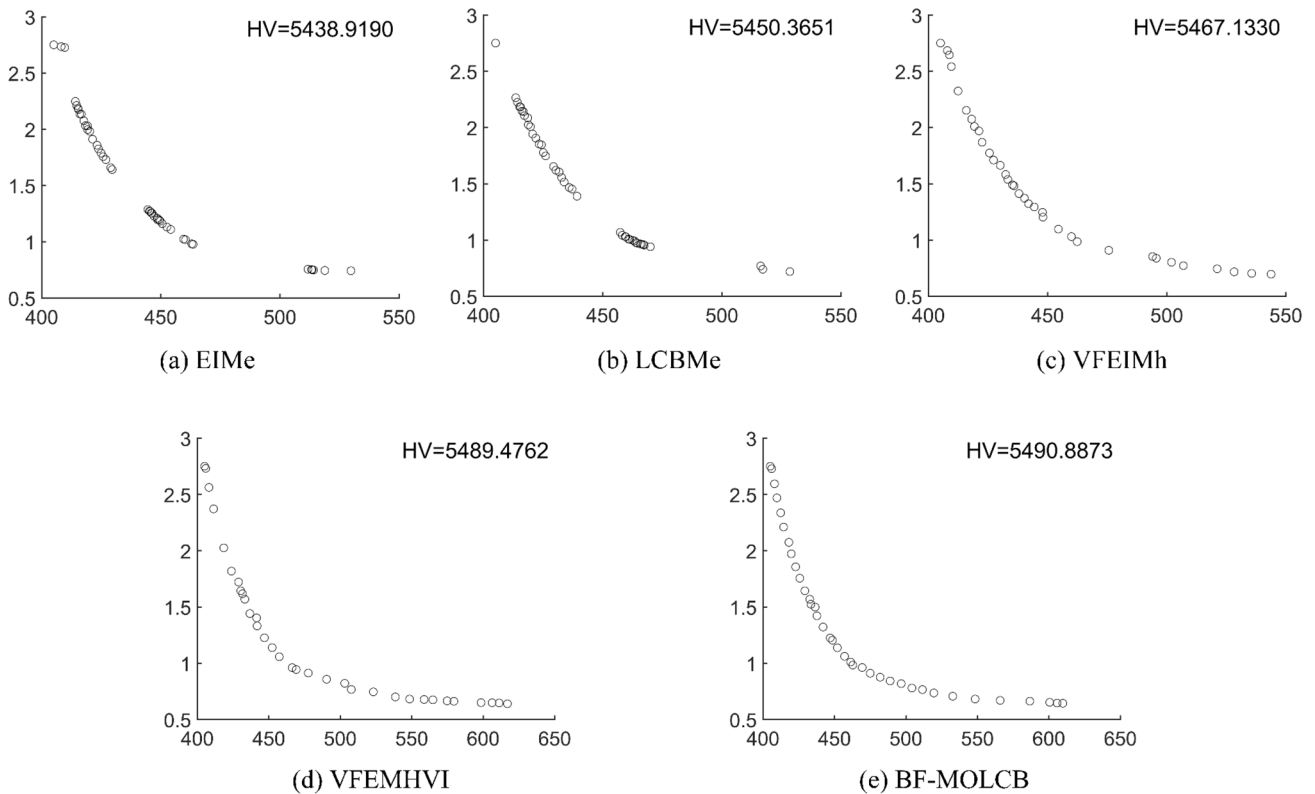
**Table 8** The comparison with the same budget for the torque arm problem

Test	Method	HV			COST
		Median	Mean	Std	Mean
Torque arm	EIMe	5469.6338	5466.1356	13.6531	160
	LCBMe	5473.9124	5469.9749	11.0629	160
	VFEIMh	5487.7229	5483.0587	10.2516	160
	VFEMHVI	5492.3657	5491.8646	1.2179	160
	BF-MOLCB	<b>5492.5094</b>	<b>5492.6020</b>	<b>1.0337</b>	160

meanwhile, the least std value reveals the stability of the proposed method. Even the HV indicators obtained by the single-fidelity methods and the VFEIMh method after 100 iterations are still less than the corresponding value when the proposed method has obtained 20 Pareto solutions (5488.2470 on average), whereas the proposed method only costs 35.93 on average to obtain 20 Pareto solutions in optimization process. The VFEMHVI method also has a good performance, which has slight discrepancy with the proposed method.

Figure 12 intuitively presents the Pareto frontier of the lowest HV indicator with the same budget. The Pareto frontiers obtained by the single-fidelity methods are much better than before but still distributed unevenly. The diversity of the Pareto solutions is also worsor than the proposed method obtained, in which a fixed area many non-dominated solutions intensively distributed. Compared with other two bi-fidelity method, the proposed method still has the best performance in the torque arm optimization problem.

In the vibration isolator optimization problem, six design variables are applied, including  $\theta, D, R_f, L, t_h,$  and  $t_l$  as shown in Fig. 13. The ratio of the natural frequency  $e$  and the Equivalent strain  $\epsilon_{max}$  are set as the goal of the multi-objective optimization problem to minimize. The nonlinear coefficient  $\xi$ , the transverse-longitudinal stiffness ratio  $R_k$ ,



**Fig.12** Pareto frontier of the lowest HV indicator with the same budget



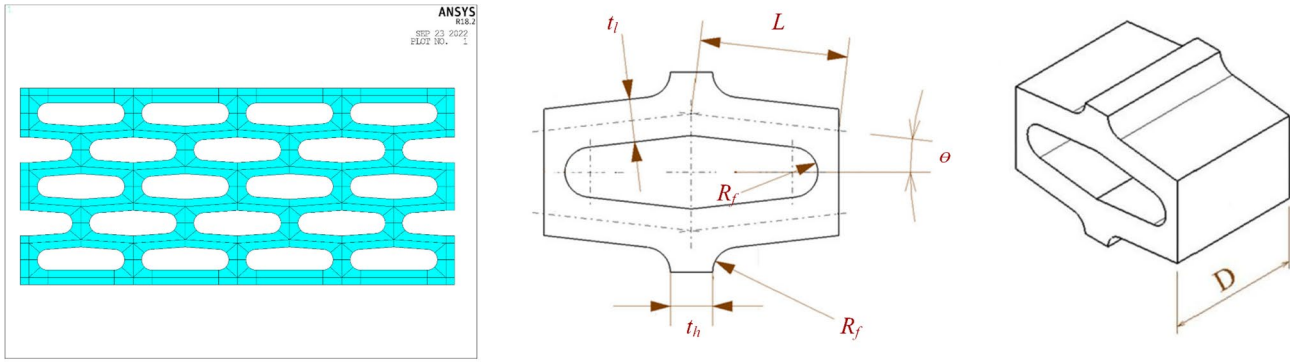


Fig.13 The vibration isolator and the geometry of a honeycomb cell (Qian et al. 2021)

Table 9 Fixed parameters

Fixed parameters	Value
Number of Rows $m$	2
Number of columns $n$	4
Density of polyurethane $\rho$	1166 kg/m <sup>3</sup>
Poisson's ratio $\mu$	0.475
Elastic modulus $E$	33.24 MPa
Rated Load $F$	50 kgf

Table 10 Physical meaning and ranges of design variables

Design variables	Description	Ranges
$\theta$	Oblique angle	4.0–15.0°
$D$	Depth	35.0–100.0 mm
$R_f$	Radius of rounding	3.0–5.0 mm
$L$	Length of oblique bar	15.0–20.0 mm
$t_h$	Thickness of vertical bar	5.0–10.0 mm
$t_l$	Thickness of oblique bar	3.5–10.0 mm

the total height  $H$ , and the dimensional coordination conditions  $T_s$  are set as the constraints, in which  $\xi$  and  $R_k$  should be evaluated by computationally expensive simulations while the  $H$  and  $T_s$  can be obtained by mathematical formulas. The optimization problem is expressed as follows:

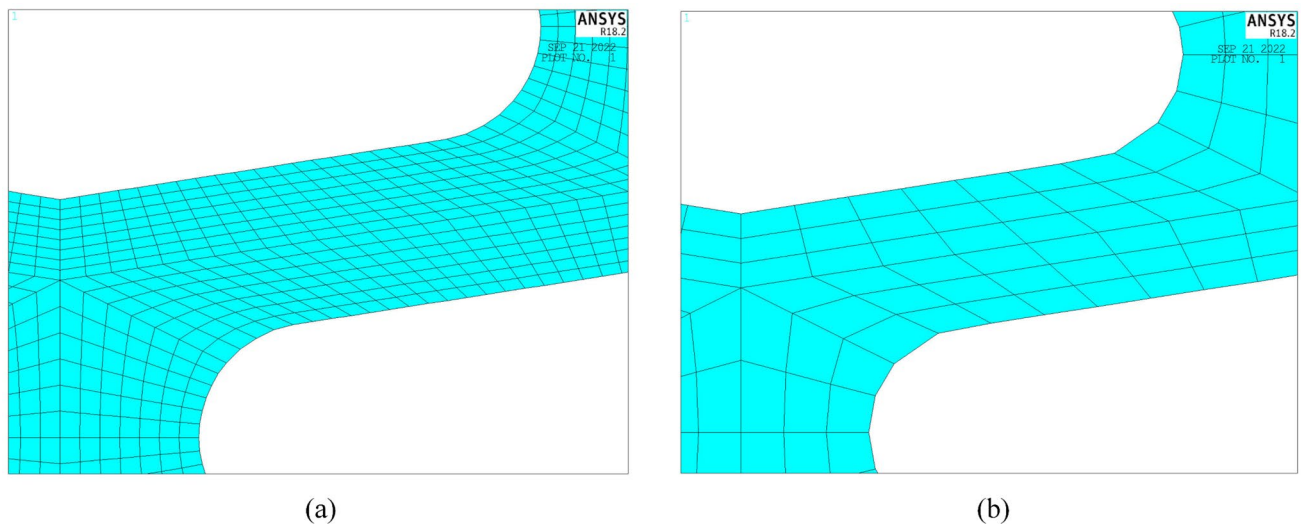
$$\begin{aligned} \min e(\theta, D, R_f, L, t_h, t_l) &= \left| 1 - \frac{f}{f_0} \right| \\ \epsilon_{\max}(\theta, D, R_f, L, t_h, t_l) & \\ \text{s.t. } g_1 = 1 - \frac{\xi}{0.75} \leq 0, g_2 = 1 - \frac{R_k}{1.2} \leq 0, & \\ g_3 = \frac{H}{70} - 1 \leq 0, g_4 = T_s \leq 0 & \end{aligned} \tag{32}$$

where the parameters are shown in Tables 9, 10, and 11.

The vibration isolator test problem is modeled and simulated using ANSYS 18.2, in which high- and low-fidelity levels are set via adjusting the dimension of the mesh division. The number of mesh division of the high-fidelity analysis model in the direction of inclined beam thickness  $t_l$  is set as 6, and the corresponding number in

Table 11 Objective and constraint functions and their boundary conditions

Objectives	Description	Boundary Condition	Whether computationally expensive or not
$e = \left  1 - \frac{f}{f_0} \right $	The ratio of the natural frequency	$f_0 = 6$	Yes
$\epsilon_{\max}$	Equivalent strain	minimize	Yes
<b>Constraints</b>			
$g_1 : \xi$	Minimum value of nonlinear coefficient	$\xi \geq 0.75$	Yes
$g_2 : R_k$	Transverse-longitudinal stiffness ratio	$R_k \geq 1.2$	Yes
$g_3 : H$	Total height of vibration isolator	$H \leq 70$	No
$g_4 : T_s$	Dimensional coordination conditions	$T_s \leq 0$	No



**Fig. 14** The mesh division for high- and low-fidelity finite element. **a** Fine mesh. **b** Coarse mesh (Qian et al. 2021)

the low-fidelity analysis model is set as 2, as shown in Fig. 14. The cost coefficient is set as 3. The penalty function is used to handle the constraints which are obtained by mathematical formulas. For constraints which need time-consuming simulations to obtain, this paper addresses them based on the method proposed by Schonlau (1998), which was extended to solve multi-fidelity optimization problems in our previous work (Shu et al. 2021). The reference point is set as [10;0.2].

Each method is run 10 times, where the average of the HV indicator and computational expense is recorded, as the shown in Table 12. All of these methods can meet the termination condition; thereinto, the proposed method has the best performance in the quality of Pareto solutions due to the highest HV indicator. With regard to the computational expense, there is only a small variation between these methods, in which the EIME has the lowest cost and the VFEMHVI is the highest one. Other three methods are almost the same level. However, the proposed method has notable advantage in term of the HV indicator, especially comparing with the EIME, the LCBMe, and the VFEIMh method. Figure 15 illustrates the Pareto frontier with the lowest HV indicator in 10 runs. It can be seen that the

Pareto frontier obtained from proposed method is better than other methods obtained.

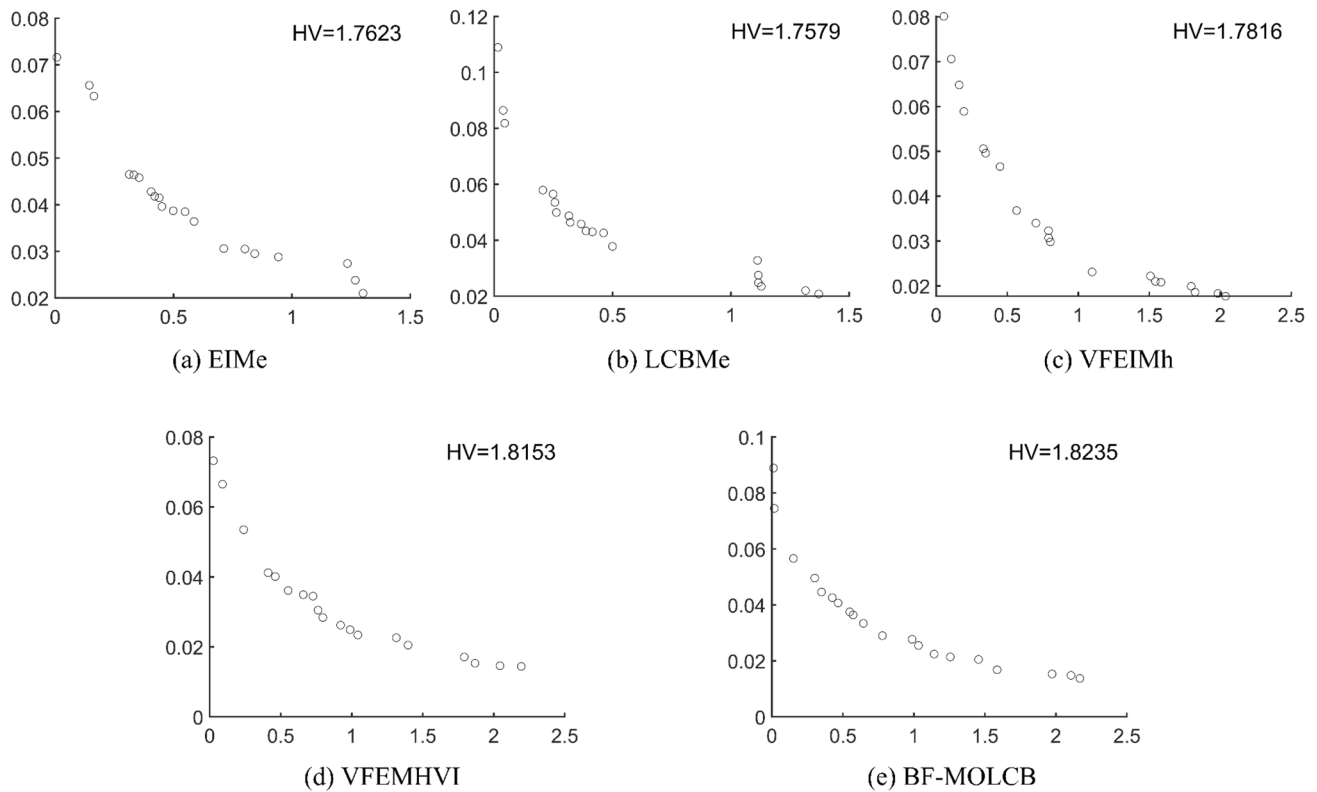
## 5 Conclusion

A bi-fidelity Bayesian optimization method for multi-objective optimization is proposed. In the proposed method, the LCB functions are adopted to utilize the uncertainty of surrogate model. On this basis, a novel acquisition function is defined, in which a cost coefficient balances the computational expense and accuracy of new sample points with different fidelity levels. Then, the new sample can be determined by maximizing the proposed acquisition function.

Four numerical test problems with different dimensions and two real-world engineering design optimization problems are applied to investigate the feasibility and efficiency of the proposed method. Some conclusions are drawn as follows:

**Table 12** The comparison with different methods for the engineering problem

Test	Method	HV			COST		
		Median	Mean	Std	Median	Mean	Std
Vibration isolator	EIME	1.7913	1.7861	0.0176	<b>124.0000</b>	<b>126.5000</b>	13.6728
	LCBMe	1.7976	1.7940	0.0185	127.0000	130.8000	22.0091
	VFEIMh	1.8087	1.8061	0.0122	131.6667	127.1000	13.1600
	VFEMHVI	1.8224	1.8222	0.0052	133.0000	132.6333	10.9448
	BF-MOLCB	<b>1.8259</b>	<b>1.8260</b>	<b>0.0021</b>	128.1667	130.2333	<b>10.7279</b>



**Fig.15** Pareto frontier with the lowest HV indicator in 10 runs on the engineering problem

- 1) The effect of samples with different fidelities for improving the quality of the current Pareto set is quantified in the proposed method;
- 2) The proposed method balances the cost and effect of samples with different fidelities, which can utilize the LF samples to potentially further reduce the computational expense;
- 3) Compared with other two state-of-the-art single-fidelity methods and two bi-fidelity methods, the proposed method shows prominent performance in computational cost and improving of Pareto solutions, especially for solving high-dimension problems.

In the future, several potential directions are worth exploring. First, the acquisition function of choosing new samples from multiple fidelities ( $> 2$ ) still remains to be researched and tested. Second, combined with the research from Foumani et al. (2022), other acquisition functions like EI and PI can be utilized in multi-objective optimization with multi-fidelity model. Last but not the least, the proposed method can be extended to solve multi-objective robust optimization problems.

**Funding** This research has been supported by the National Natural Science Foundation of China (NSFC) under Grant No. 52105256 and No. 52188102.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Replication of results** The main step of the proposed method is introduced in Sect. 3. The Matlab codes can be available from the website: <https://pan.baidu.com/s/1B0vf4QaWByIoNUvoa9vFUA> by using the password: 0ubx.

## References

- Chen S, Jiang Z, Yang S, Chen W (2016) Multimodel fusion based sequential optimization. *AIAA J* 55(1):241–254
- Choi S, Alonso JJ, Kroo IM (2009) Two-level multifidelity design optimization studies for supersonic jets. *J Aircr* 46(3):776–790
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6(2):182–197
- Emmerich MTM, Deutz AH, Klinkenberg JW (2011) Hypervolume-based expected improvement: Monotonicity properties and exact computation. Paper presented at the 2011 IEEE congress of evolutionary computation (CEC)

- Eweis-Labolle JT, Oune N, Bostanabad R (2022) Data fusion with latent map gaussian processes. *J Mech Des* 144(9):091703
- Han ZH, Görtz S (2012) Hierarchical kriging model for variable-fidelity surrogate modeling. *AIAA J* 50:1885–1896
- He Y, Sun J, Song P, Wang X (2022) Variable-fidelity hypervolume-based expected improvement criteria for multi-objective efficient global optimization of expensive functions. *Eng Comput* 38(4):3663–3689
- Jeong S, Obayashi S (2005) Efficient global optimization (EGO) for multi-objective problem and data mining. Paper presented at the 2005 IEEE congress on evolutionary computation
- Jerome S, William JW, Toby JM, Henry PW (1989) Design and analysis of computer experiments. *Stat Sci* 4(4):409–423
- Jiang P, Cheng J, Zhou Q, Shu L, Hu J (2019) Variable-fidelity lower confidence bounding approach for engineering optimization problems with expensive simulations. *AIAA J* 57(12):5416–5430
- Jin Y (2011) Surrogate-assisted evolutionary computation: recent advances and future challenges. *Swarm Evol Comput* 1(2):61–70
- Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. *J Global Optim* 21(4):345–383
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *J Global Optim* 13(4):455–492
- Keane AJ (2006) Statistical improvement criteria for use in multiobjective design optimization. *AIAA J* 44(4):879–891
- Kennedy MC, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87(1):1–13
- Knowles J (2006) ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans Evol Comput* 10(1):50–66
- Leifsson L, Koziel S, Tesfahunegn YA (2015) Multiobjective aerodynamic optimization by variable-fidelity models and response surface surrogates. *AIAA J* 54(2):531–541
- Li X, Qiu H, Jiang Z, Gao L, Shao X (2017) A VF-SLP framework using least squares hybrid scaling for RBDO. *Struct Multidiscip Optim* 55(5):1629–1640
- Liu H, Ong YS, Cai J, Wang Y (2018) Cope with diverse data structures in multi-fidelity modeling: a Gaussian process method. *Eng Appl Artif Intell* 67:211–225
- Lophaven SN, Nielsen HB, Søndergaard J (2002) DACE: a Matlab kriging toolbox (Vol. 2): Citeseer
- Perdikaris P, Venturi D, Roysset JO, Karniadakis GE (2015) Multi-fidelity modelling via recursive co-kriging and Gaussian–Markov random fields. *Proc R Soc a: Math, Phys Eng Sci* 471:20150018
- Qian J, Cheng Y, Zhang A, Zhou Q, Zhang J (2021) Optimization design of metamaterial vibration isolator with honeycomb structure based on multi-fidelity surrogate model. *Struct Multidiscip Optim* 64(1):423–439
- Ruan X, Jiang P, Zhou Q, Hu J, Shu L (2020) Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems. *Struct Multidiscip Optim* 62(6):3021–3052
- Schonlau M, Welch W, Jones D (1998) Global versus local search in constrained optimization of computer models. *Inst Math Stat Lect Notes Monograph Ser* 34:11–25
- Shahriari B, Swersky K, Wang Z, Adams RP, Freitas Nd (2016) Taking the human out of the loop: a review of bayesian optimization. *Proc IEEE* 104(1):148–175
- Shi M, Lv L, Sun W, Song X (2020) A multi-fidelity surrogate model based on support vector regression. *Struct Multidiscip Optim* 61(6):2363–2375
- Shu L, Jiang P, Zhou Q, Xie T (2019) An online variable-fidelity optimization approach for multi-objective design optimization. *Struct Multidiscip Optim* 60(3):1059–1077
- Shu L, Jiang P, Shao X, Wang Y (2020) A new multi-objective bayesian optimization formulation with the acquisition function for convergence and diversity. *J Mech Design*. <https://doi.org/10.1115/14046508>
- Shu L, Jiang P, Wang Y (2021) A multi-fidelity Bayesian optimization approach based on the expected further improvement. *Struct Multidiscip Optim* 63(4):1709–1719
- Singh P, Couckuyt I, Elsayed K, Deschrijver D, Dhaene T (2017) Multi-objective geometry optimization of a gas cyclone using triple-fidelity co-kriging surrogate models. *J Optim Theory Appl* 175(1):172–193
- Sóbester A, Leary SJ, Keane AJ (2005) On the design of optimization strategies based on global response surface approximation models. *J Global Optim* 33(1):31–59
- Srinivas N, Krause A, Kakade SM, Seeger MW (2012) Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans Inf Theory* 58(5):3250–3265
- Sun G, Li L, Fang J, Li Q (2021) On lower confidence bound improvement matrix-based approaches for multiobjective Bayesian optimization and its applications to thin-walled structures. *Thin Walled Struct* 161:107248
- Svenson J, Santner T (2016) Multiobjective optimization of expensive-to-evaluate deterministic computer simulator models. *Comput Stat Data Anal* 94:250–264
- Tran A, Tran M, Wang Y (2019) Constrained mixed-integer Gaussian mixture Bayesian optimization and its applications in designing fractal and auxetic metamaterials. *Struct Multidiscip Optim* 59(6):2131–2154
- Wang GG (2003) Adaptive response surface method using inherited Latin hypercube design points. *J Mech Des* 125(2):210–220
- Williams B, Cremaschi S (2021) Selection of surrogate modeling techniques for surface approximation and surrogate-based optimization. *Chem Eng Res Des* 170:76–89
- Williams CK, Rasmussen CE (2006) Gaussian processes for machine learning. MIT press Cambridge, MA
- Yang K, Emmerich M, Deutz A, Bäck T (2019) Multi-objective Bayesian global optimization using expected hypervolume improvement gradient. *Swarm Evol Comput* 44:945–956
- Zanjani Foumani Z, Shishehbor M, Yousefpour A, Bostanabad R (2022). Multi-Fidelity Cost-Aware Bayesian Optimization. [arXiv:2211.02732](https://arxiv.org/abs/2211.02732)
- Zhan D, Cheng Y, Liu J (2017) Expected improvement matrix-based infill criteria for expensive multiobjective optimization. *IEEE Trans Evol Comput* 21(6):956–975
- Zhan D, Meng Y, Xing H (2022) A fast multi-point expected improvement for parallel expensive optimization. *IEEE Trans Evol Comput*. <https://doi.org/10.1109/TEVC.2022.3232776>
- Zhang Y, Han ZH, Zhang KS (2018) Variable-fidelity expected improvement method for efficient global optimization of expensive functions. *Struct Multidiscip Optim* 58(4):1431–1451
- Zhang Y, Kim NH, Park C, Haftka RT (2018) Multifidelity surrogate based on single linear regression. *AIAA J* 56(12):4944–4952
- Zheng J, Li Z, Gao L, Jiang G (2016) A parameterized lower confidence bounding scheme for adaptive metamodel-based design optimization. *Eng Comput* 33(7):2165–2184
- Zhou Q, Jiang P, Shao X, Hu J, Cao L, Wan L (2017) A variable fidelity information fusion method based on radial basis function. *Adv Eng Inform* 32:26–39

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.