ORIGINAL ARTICLE

Variable-fidelity hypervolume-based expected improvement criteria for multi-objective efficient global optimization of expensive functions

Youwei He[1] · Jinju Sun[1,2] · Peng Song[1] · Xuesong Wang[1]

Abstract

Variable-fidelity surrogate-based efficient global optimization (EGO) method with the ability to adaptively select low-fidelity (LF) and high-fidelity (HF) infill point has emerged as an alternative to further save the computational cost of the single-fidelity EGO method. However, in terms of the variable-fidelity surrogate-assisted multi-objective optimization methods, existing methods rely on empirical parameters or are unable to adaptively select LF/HF sample in the optimal search process. In this paper, two variable-fidelity hypervolume-based expected improvement criteria with analytic expressions for variable-fidelity multi-objective EGO method are proposed. The first criterion relies on the concept of variable-fidelity expected improvement matrix (VFEIM) and is obtained by aggregating the VFEIM using a simplified hypervolume-based aggregation scheme. The second criterion termed as VFEMHVI is derived analytically based on a modified hypervolume definition. Both criteria can adaptively select new LF/HF samples in the iterative optimal search process to update the variable-fidelity models towards the HF Pareto front, distinguishing the proposed methods to the rests in the open literature. The constrained versions of the two criteria are also derived for problems with constraints. The effectiveness and efficiency of the proposed methods are verified and validated over analytic problems and demonstrated by two engineering problems including aerodynamic shape optimizations of the NACA0012 and RAE2822 airfoils. The results show that the VFEMHVI combined with the normalization-based strategy to define the reference point is the most efficient one over the compared methods.

Keywords Variable-fidelity surrogate · Multi-objective efficient global optimization · Surrogate · Expected improvement · Hypervolume

1 Introduction

Surrogate-based optimization (SBO) method is a promising approach to reduce the computational cost in expensive simulation-based engineering design optimization problems, as the surrogate model can be used to replace the expensive simulations during the optimization and the cost of building a surrogate model is almost negligible compared with an expensive simulation. Among the available surrogate models, Kriging [1] or Gaussian process regression [2] has gained popularity due to its capability of providing the prediction of a function and the uncertainty of the prediction. In the field of engineering optimization design, the Kriging-based efficient global optimization (EGO) [3] method utilizing the expected improvement (EI) criterion has gotten popularity after its birth, for it enables the searching of global optimum with much less number of expensive function evaluations than the methods such as evolutionary algorithms. For example, EGO has been applied to improve the performance of the axial flow compressor rotor [4] or cascades [5], cryogenic liquid turbine stage [6] and turbine blade [7], etc. The advent of the EGO method greatly inspired the research and development of the SBO [8, 9], including utilizing the variable-fidelity (VF) surrogate to save the computational cost of the optimization task further [10, 11] or extending the EGO for single-objective optimization to solve multi-objective problems [12, 13], etc.

✉ Jinju Sun
   jjsun@mail.xjtu.edu.cn

[1] School of Energy and Power Engineering, Xi'an Jiaotong University, 28 West Xianning Road, Xi'an 710049, China

[2] Collaborative Innovation Center for Advance Aero-Engine (CICAAE) of China, 37 Xueyuan Road, Beijing 100191, China

VF surrogate uses a set of numerical methods with varying degrees of fidelity and computational expense (e.g., potential theory, Euler equations, and Navier–Stokes equations) to approximate the unknown function of interest as a function of the independent variables while reducing the more expensive high-fidelity simulations [14]. Among the different types of VF surrogate, the Co-Kriging and its variants [15, 16], have gained much attention and been adapted to develop VF surrogate-assisted EGO-like method (VFEGO). The augmented expected improvement and the statistic lower bounding criterion were proposed to guide the selection of the location and fidelity level of the next infill sample in [17, 18], respectively. More recently, the extended expected improvement-based on an enhanced Co-Kriging, VF expected improvement (VF-EI) based on Hierarchical Kriging (HK), VF lower confidence bound based on additive scaling VF surrogate, and the expected improvement-based criterion for gradient-enhanced VF surrogate were developed in [19–22], respectively. Those methods can be viewed as a generalized EGO method using different types of VF surrogate models and infill sampling criteria (ISC). Compared with the EI criterion for the single-fidelity EGO method, the distinct feature of the ISCs used in the various VFEGO methods is the ability to adaptively determine the fidelity level of the infill sampling point to be evaluated at.

On extending the EGO to treat multi-objective problems, the multi-objective efficient global optimization (MOEGO) is one of the most frequently studied surrogate model-based optimization algorithms as it enforces a more careful search since it integrates both the prediction and error estimates [23–31]. The MOEGO method shares the same framework of EGO method, consisting of the initialization phase and the iterative search phase. In the initialization phase, the initial samples are generated by the design of experiments (DOE) method and then evaluated using the computational expensive simulation tools. Based on those initial samples, the Kriging models for each objective function are constructed. Then MOEGO proceeds to the iterative phase. At each iteration, the multi-objective EI function is maximized to obtain one infill point which is evaluated with the expensive function. Such a procedure will iterate until the stopping rule is met. It can be noted that the MOEGO approach differs from the EGO method at two aspects: multiple Kriging models are built for each objective and multi-objective EI-based criterion is adopted to guide the selection of the next infill point. The efficiency of the MOEGO method mainly depends on the multi-objective EI infill criterion used [31]. There are different kinds of criteria to be used in MOEGO, the expected angle-penalized length improvement criterion [13], the hypervolume-based EI criterion [24], probability of hypervolume improvement criterion [25], the Euclidean distance-based EI criterion [26], the maximin distance-based EI criterion [27], the expected inverted penalty boundary

intersection improvement [28], the expected improvement matrix-based infill criterion [29], the modified hypervolume-based EI criterion [32], etc. Among those criteria, the last two criteria are of the distinct feature being cheap-to-evaluate and yet efficient to obtain well-distributed solutions compared to other aforementioned criteria, but the performance comparison between those two criteria has never been conducted. Nevertheless, the MOEGO method has been demonstrated to be effective and efficient in engineering applications, such as the airfoil shape optimization [31], bump inlet shape optimization to minimize the total pressure distortions [33], S-shape supersonic axial flow compressor cascade optimization [5], etc.

In terms of using the VF surrogate-based optimization method to solve the expensive multi-objective engineering problems, the existing method can be classified into three categories. The first group of VF surrogate-based algorithms [34–38] finds the Pareto Front (PF) on the LF surrogate first, followed by choosing promising solutions on the PF to be evaluated by HF simulations. With HF samples, the iterative optimization process starts with tuning VF surrogate and followed by determining the PF on the VF surrogate. The third step of the iterative phase is choosing solutions on the obtained PF to be evaluated at the HF level. The termination condition is checked and if not met, the VF surrogate is updated with the newly obtained HF samples. This kind of method has been applied to optimize the antenna [34] and airfoil [37]. It should be noted that how to select a limited number of promising solutions from a large amount of PF solutions is still a challenging task. Moreover, for methods belonging to the first group, the number of LF samples is fixed during the design, while HF samples are repetitively added in the iterative optimization stage. It is, of course, not the most efficient way, since it is always questionable how many LF samples would be sufficient. The other way, and a potentially better way, is that both LF and HF samples are adaptively added. The second group of methods incorporates the VF surrogate into the population-based evolutionary algorithms [39–46]. Such algorithms usually start from the population initialization and then evaluate the individuals using LF and HF simulators, respectively. With the available LF and HF samples, the VF surrogates for each objective are then constructed. The population is evolved iteratively with the fitness of each individual calculated by using the VF surrogate model. After certain evolutionary iterations, individuals in the current population are chosen by sophisticated model manage strategies and to be evaluated by LF/HF simulators. Effectiveness of such VF surrogate-based evolutionary algorithms have been demonstrated by optimizing a torque arm [42] and a stiffened cylindrical shell [43]. Disappointedly, the population-based methods still need thousands of function evaluations which might be unacceptable in some scenarios and the sophisticated model manage

strategy usually relies on empirical parameters. For example, a threshold of the minimum of minimum distance is required for the online variable-fidelity surrogate-assisted multi-objective genetic algorithm [42]. The VF surrogate-based method following the MOEGO framework (VFMOEGO) is classified into the third group, including an output space entropy search method [47] in the Bayesian optimization field. With the initialization stage to prepare the initial data set and build the VF surrogates for each objective, the iterative optimization process starts. In each iteration, an output space entropy-based ISC is utilized to determine the spatial position and fidelity level of the next sample point to be evaluated at. In other words, the LF and HF samples will be adaptively added. It can be noted that such VF surrogate-based method follows the MOEGO framework and is free of the empirical parameter. Like MOEGO, the efficiency of this kind of method heavily relies on the adopted ISC. Unfortunately, the ISC proposed in the output space entropy search method stands on the assumption that the value of lower fidelity of an objective must be smaller than that of the corresponding higher fidelity [47]. This indicates the output space entropy search method is applicable to solve problems with such features, limiting its usefulness in engineering applications. Therefore, proposing a general ISC for VFMOEGO method is required to solve engineering problems.

To fill in this gap, two general cheap-to-evaluate ISCs for VFMOEGO method are proposed, by extending the HK model-based VF-EI criterion for single-objective optimization to the multi-objective scenario. The performance of the VFMOEGO methods utilizing the proposed ISCs is illustrated using analytic examples and two real-world engineering examples. The comparisons between the proposed approaches and existing approaches considering the computational efficiency and robustness are made. The merits of the proposed approach are analyzed and summarized.

The remainder of this paper is organized as follows. Sect. 2 recalls the theoretical background. Then, in Sect. 3, the proposed ISCs are introduced in detail with the implementation and illustration. The compared methods, analytic, and real-world test problems for the comparative study, and the results of the comparative study are given in Sect. 4. Conclusions and future works are summarized in Sect. 5.

2 Background and terminology

This study aims to solve the following multi-objective optimization (MOO) problem based on HF expensive numerical simulation with the assistance of LF but cheaper numerical analysis:

$$\min_{\mathbf{x}\in\Omega} F(\mathbf{x}) = \left\{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_i(\mathbf{x}), \dots, f_M(\mathbf{x})\right\}$$

with assistance of $F_{lf}(\mathbf{x}) = \left\{f_{1,lf}(\mathbf{x}), f_{2,lf}(\mathbf{x}), \dots, f_{i,lf}(\mathbf{x}), \dots, f_{M,lf}(\mathbf{x})\right\}$, (1)

where $F(\mathbf{x})$ and $F_{lf}(\mathbf{x})$ denote the objective functions evaluated by HF and LF numerical simulators, respectively, $\mathbf{x}$ is the vector of design parameters with length being d, M represents the number of objectives, $\Omega \in \mathbb{R}^d$ is the design space. A point $\mathbf{x}$ is said to Pareto-dominate another point $\mathbf{x}'$ if $f_i(\mathbf{x}) \leq f_i(\mathbf{x}')\forall i$ and there exists some $j \in \{1, 2, \dots, M\}$ such that $f_j(\mathbf{x}) < f_j(\mathbf{x}')$. The optimal solutions of a MOO problem is a set of points $\mathcal{X}^* \subset \Omega$ such that no point $\mathbf{x}' \in \Omega\backslash\mathcal{X}^*$ Pareto-dominates a point $\mathbf{x} \in \mathcal{X}^*$. The solution set $\mathcal{X}^*$ is called the optimal *Pareto set* and the corresponding set of function values F^* is called the optimal *Pareto front* (PF). The PF with k solutions is denoted as:

$$\text{PF} = \left\{f_1^j, f_2^j, \dots f_M^j\right\}, j = 1, 2, \dots, k. \tag{2}$$

Our goal is to approximate PF by minimizing the computational cost through the VFMOEGO method or obtain a better approximation of PF with a budgeted cost. In the proposed VFMOEGO methods, each objective is to be approximated by an HK model. Meanwhile, the HK model is built based on Kriging. And the ISCs used in the VFMOEGO methods are developed by extending the VF-EI function for single-objective VFEGO method to VFMOEGO. Thus, the basics of the Kriging, HK, and VF-EI are briefed as follows.

2.1 Kriging

Kriging is an interpolation technique that originated from the geostatistical community and then applied to approximate expensive black-box functions [1]. It hypothesizes that a random process exits in each sampling site, and consists of two parts: the trend function and the random process:

$$Y(\mathbf{x}) = \beta_0 + Z(\mathbf{x}). \tag{3}$$

For ordinary Kriging, the trend function β_0 is an unknown constant and $Z(\mathbf{x})$ is a stationary random process with zero mean and process variance σ^2. The covariance of $Z(\mathbf{x})$ can be expressed as:

$$\text{CoV}\left(Z(\mathbf{x}), Z(\mathbf{x}')\right) = \sigma^2 R(\mathbf{x}, \mathbf{x}'), \tag{4}$$

where $R(\mathbf{x}, \mathbf{x}')$ is the spatial correlation function depending on the Euclidean distance between two sites $\mathbf{x}$ and $\mathbf{x}'$. The common-used correlation functions are the squared exponential function, the Matérn function, and the exponential function. In the field of design and analysis of computer experiment (DACE), the widely used squared exponential function (aka Gaussian correlation) is expressed as:

$$R(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_{i=1}^{d} \theta_i |x_i - x_i'|^2\right), \tag{5}$$

where $\theta_i, i = 1, 2, \ldots d$ are parameters measuring the activity of the variable and being determined in the model fitting process.

By minimizing the mean square error (MSE) of the random process subject to the unbiasedness constraint, the Kriging prediction $\hat{y}(\mathbf{x})$ along with the MSE $s^2(\mathbf{x})$ of the prediction at any unvisited point $\mathbf{x}$ can be derived as:

$$\hat{y}(\mathbf{x}) = \beta_0 + \mathbf{r}^T \mathbf{R}^{-1}(\mathbf{y}_s - \beta_0 \mathbf{1})$$
$$s^2(\mathbf{x}) = \sigma^2 \left\{ 1.0 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \left[\mathbf{r}^T \mathbf{R}^{-1}\mathbf{1} - 1\right]^2 \Big/ \left(\mathbf{1}^T \mathbf{R}^{-1}\mathbf{1}\right) \right\}, \tag{6}$$

where $\mathbf{r}$ is a correlation vector between the unobserved point and the sample points, $\mathbf{R}$ is the correlation matrix whose elements are correlation functions between sample points, $\mathbf{y}_s$ is the column vector that contains the responses at the sample points, $\mathbf{1}$ is a column vector filled with ones. For more details about Kriging refer to [1].

2.2 Hierarchical Kriging for variable-fidelity surrogate modeling

HK [48] is distinct from other VF model by taking the low-fidelity (LF) function as the model trend for the HF model to avoid calculating the covariance matrix between LF and HF samples. HK has been proved to be more computational efficient without losing accuracy compared with Co-Kriging [49–51]. To begin with, the Kriging model of LF function is built based on the LF samples. The prediction and MSE of the LF Kriging are denoted as $\hat{y}_{lf}(\mathbf{x})$ and $s_{lf}^2(\mathbf{x})$ respectively. Then the LF Kriging is used as the model trend of the Kriging for the HF function of the form:

$$Y(\mathbf{x}) = \beta_{lf}\hat{y}_{lf}(\mathbf{x}) + Z(\mathbf{x}), \tag{7}$$

where β_{lf} is a scaling factor representing how the LF Kriging $\hat{y}_{lf}(\mathbf{x})$ correlates the HF function, $Z(\mathbf{x})$ represents the random process having zero mean and variance with the same form shown in (4). Via minimizing the MSE of the random process subject to the unbiasedness constraint, the HK predictor $\hat{y}(\mathbf{x})$ and its MSE $s^2(\mathbf{x})$ are formulated as:

$$\hat{y}(\mathbf{x}) = \beta_{lf}\hat{y}_{lf}(\mathbf{x}) + \mathbf{r}^T \mathbf{R}^{-1}(\mathbf{y}_s - \beta_{lf}\mathbf{F})$$
$$s^2(\mathbf{x}) = \sigma^2 \left\{ 1 - \mathbf{r}^T \mathbf{R}^{-1}\mathbf{r} + \left[\mathbf{r}^T \mathbf{R}^{-1}\mathbf{F} - \hat{y}_{lf}(\mathbf{x})\right]^2 \Big/ \left(\mathbf{F}^T \mathbf{R}^{-1}\mathbf{F}\right) \right\}, \tag{8}$$

where $\beta_{lf} = (\mathbf{F}^T \mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T \mathbf{R}^{-1}\mathbf{y}_s$, $\mathbf{r}$ is a correlation vector between the unobserved point and the HF sample points, $\mathbf{R}$ is the correlation matrix whose elements are correlation functions between HF sample points, $\mathbf{y}_s$ is the column vector

that contains the true responses at the HF sample points, $\mathbf{F} = \left[\hat{y}_{lf}(\mathbf{x}^{(1)}), \ldots, \hat{y}_{lf}(\mathbf{x}^{(n)})\right]^T$ represents the n predictions from the LF Kriging at the HF sample sites $\left[\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}\right]^T$. Details about the HK model can be found in [48].

2.3 HK-based variable-fidelity EI

In the VFEGO method proposed in [19], a new infill sample point is obtained by the VF-EI criterion at each iteration of the iterative optimization phase to update the HK model to reach the global optimum of the HF function efficiently. The VF-EI is derived based on the assumption that the prediction from the HK model for the objective function at any unvisited point $\mathbf{x}$ obeys the following normal distribution:

$$\hat{Y}(\mathbf{x}, l) \sim N\left[\hat{y}(\mathbf{x}), s^2(\mathbf{x}, l)\right], \, l = 1, 2, \tag{9}$$

with

$$s^2(\mathbf{x}, l) = \begin{cases} \beta_{lf}^2 s_{lf}^2(\mathbf{x}), & l = 1 \text{ for LF level} \\ s^2(\mathbf{x}), & l = 2 \text{ for HF level} \end{cases}. \tag{10}$$

With this, the improvement w.r.t. the best-observed HF objective value so far $f_{\min}$ is formulated as:

$$I_{vf}(\mathbf{x}, l) = \max\left(f_{\min} - \hat{Y}(\mathbf{x}, l), 0\right), \tag{11}$$

which is a function of both the spatial location $\mathbf{x}$ and fidelity level l. Then the VF-EI relating to the above improvement function can be derived as:

$$\mathrm{EI}_{vf}(\mathbf{x}, l) = (f_{\min} - \hat{y}(\mathbf{x}))\Phi\left(\frac{f_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x}, l)}\right) + s(\mathbf{x}, l)\phi\left(\frac{f_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x}, l)}\right), \tag{12}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and probability density function of the standard normal distribution. It can be noted that the above VF-EI only relates to the improvement of the HF function, which means the infilling of both LF and HF samples are HF-optimum oriented. It should be noted that, in the above VF-EI function, the cost ratio between LF and HF simulations are not taken into consideration. The disadvantage of such a feature is it might not be as efficient as other methods including a cost-control strategy in some applications [21]. But for problems with varying cost over the design space or the cost ratio between LF and HF simulations are not know beforehand, its performance will be more robust [52]. The new sample point $\mathbf{x}$ and its fidelity l at each iteration are determined by maximizing the VF-EI, that is:

$$\mathbf{x}, l = \underset{\mathbf{x} \in \Omega, l=1,2}{\arg\max} \mathrm{EI}_{vf}(\mathbf{x}, l). \tag{13}$$

If the new sample is found to be an LF one ($l = 1$), which means that adding an LF sample will improve the HF objective function most, it will be evaluated by the cheaper

analysis; otherwise, the new sample will be evaluated by the expensive analysis ($l = 2$).

3 Proposed criteria for VFMOEGO

This work aims to develop the ISCs for the VFMOEGO method. The overall steps of the VFMOEGO are described first. VFMOEGO follows the framework of EGO/MOEGO and consists of the initialization and iterative optimization phases. The initialization phase includes the generation of the LF and HF initial samples by the DOE method individually and the evaluation of the LF and HF samples using the simulation tool at the corresponding fidelity level. The iterative optimization stage proceeds with the construction of the HK models for each objective function based on the initial data set. Then the spatial position and fidelity level to be evaluated at the next sample point are obtained by maximizing the ISC. Third, the simulator is invoked to get the LF/HF response of the sample point. The data set is augmented before the check of the termination condition. The iterative optimization phase will continue iteratively until the stopping rule is met.

In VFMOEGO, each objective is approximated by an HK model. The objective vector of the studying point **x** then can be considered as the following M-dimensional independent Gaussian random variables:

$$\hat{Y}_i(\mathbf{x}, l) \sim N\big[\hat{y}_i(\mathbf{x}), s_i^2(\mathbf{x}, l)\big], \ l = 1, 2, i = 1, \ldots, M, \tag{14}$$

where $\hat{y}_i(\mathbf{x})$ is the prediction of the studying point **x** offered by the ith HK model, $s_i(\mathbf{x}, l)$ with $l = 1$ or $l = 2$ represents the variance at LF or HF level offered by the ith HK model. With this, the proposed two criteria are derived as follows.

3.1 Variable-fidelity expected improvement matrix-based criterion

When dealing with the MOO problem, the current best-observed HF objective value $f_{\min}$ in single-objective VFEGO expands in two directions: (1) the number of points in the current best solution increases from one to k and (2) the dimension of each point changes from one to M. Hence the current best solution for MOO is a 2-D matrix:

$$\begin{bmatrix} f_1^1 & f_1^2 & \cdots & f_1^M \\ f_2^1 & f_2^2 & \cdots & f_2^M \\ \vdots & \vdots & \ddots & \vdots \\ f_k^1 & f_k^2 & \cdots & f_k^M \end{bmatrix}. \tag{15}$$

Inspired by this, the scalar function VF-EI for single objective VFEGO can also be expanded into a 2-D matrix for MOO. Correspondingly, the variable-fidelity expected improvement matrix (VFEIM) is formulated as:

$$\begin{bmatrix} \mathrm{EI}_{vf}^{1,1}(\mathbf{x}, l) & \mathrm{EI}_{vf}^{2,1}(\mathbf{x}, l) & \cdots & \mathrm{EI}_{vf}^{M,1}(\mathbf{x}, l) \\ \mathrm{EI}_{vf}^{1,2}(\mathbf{x}, l) & \mathrm{EI}_{vf}^{2,2}(\mathbf{x}, l) & \cdots & \mathrm{EI}_{vf}^{M,2}(\mathbf{x}, l) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{EI}_{vf}^{1,k}(\mathbf{x}, l) & \mathrm{EI}_{vf}^{2,k}(\mathbf{x}, l) & \cdots & \mathrm{EI}_{vf}^{M,k}(\mathbf{x}, l) \end{bmatrix}, \tag{16}$$

with

$$EI_{\mathrm{vf}}^{i,j}(\mathbf{x}, l) = \left(f_j^i - \hat{y}_i(\mathbf{x})\right)\Phi\left(\frac{f_j^i - \hat{y}_i(\mathbf{x})}{s_i(\mathbf{x}, l)}\right) + s_i(\mathbf{x}, l)\phi\left(\frac{f_j^i - \hat{y}_i(\mathbf{x})}{s_i(\mathbf{x}, l)}\right), \tag{17}$$

$$s_i(\mathbf{x}, l) = \begin{cases} \beta_0^2 s_{lf,i}^2(\mathbf{x}), & l = 1 \text{ for LF level} \\ s_i^2(\mathbf{x}), & l = 2 \text{ for HF level} \end{cases}, \tag{18}$$

where $F_j = \left\{f_j^i\right\}, i = 1, 2, \ldots, M, j = 1, 2, \ldots, k$ are the k non-dominated front points in the current HF sample set. Figure 1a, b illustrate the VFEIM with $l = 1$ (LF level) and $l = 2$ (HF level), respectively. It can be noted that the ith row in the matrix represents the VF-EIs beyond the ith non-dominated point in all objectives and the jth column in the matrix represents the VF-EI in jth objective beyond all the non-dominated points. To sum up, the VFEIM contains all the information about the VF-EIs of the studying point **x** beyond all the non-dominated front points in each direction.

It can be noted that the VFEIM is a simple and natural extension. However, the VFEIM gives no comprehensive measurement of how much the studying point can improve the PF approximation. Thus, a strategy to aggregate VFEIM into a scalar value to judge the overall improvement of the studying point compared against the current PF approximation should be utilized. The aggregation scheme based on hypervolume measurement in [29] is adopted.

Hypervolume is an indicator measuring the volume (or hypervolume in high dimensional space) of the region dominated by the PF approximation **P** and bounded by a reference point $\mathbf{r} \in \mathbb{R}^M$:

$$H(\mathbf{P}) = V\big(\mathbf{y} \in \mathbb{R}^M | \mathbf{P} \prec \mathbf{y} \prec \mathbf{r}\big), \tag{19}$$

where $\mathbf{P} \prec \mathbf{y}$ means **y** Pareto-dominates **P**, $V(\cdot)$ denotes the hypervolume of a bounded space. Figure 2 presents the hypervolume in 2-D space. The reference point **r** is specified by the user with a relatively large value to ensure that it is dominated by any point in the PF. A higher value of the hypervolume indicator means a better distribution of a PF approximation. However, the hypervolume indicator is expensive to compute when the number of objectives is larger than two.

For a given point $\mathbf{y} \in \mathbb{R}^M$, the hypervolume improvement related to add it into the current PF approximation is defined as the growth of $H(\mathbf{P})$. It is illustrated in Fig. 3 and formulated as:
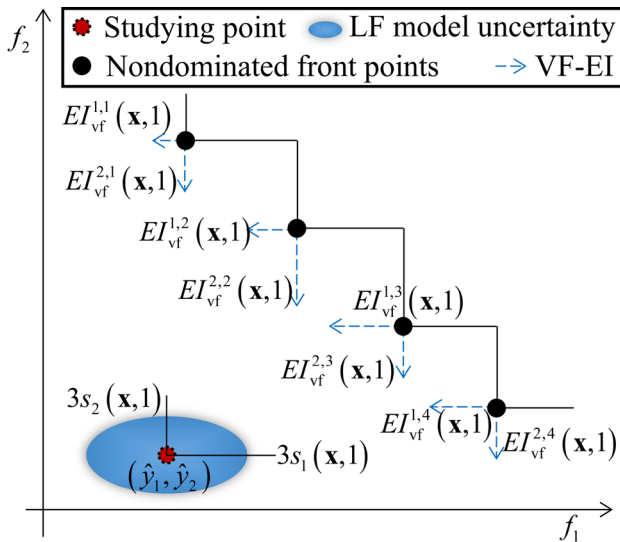
(a) VFEIM with $l = 1$
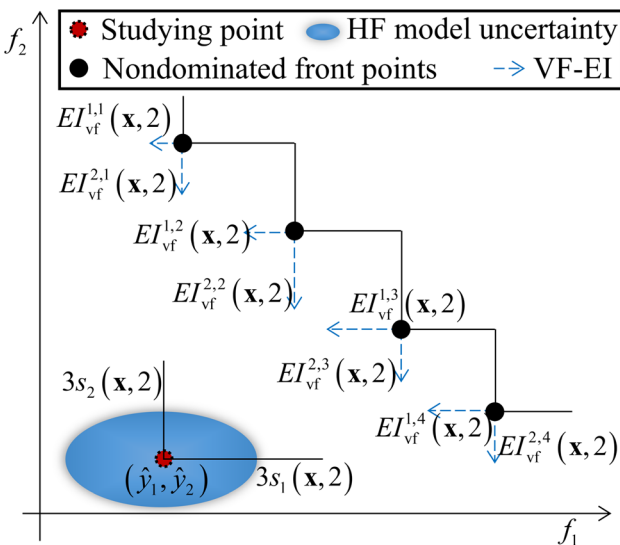
(b) VFEIM with $l = 2$

Fig. 1 2-D illustrations of the VF-EIM

Fig. 2 2-D illustration of the hypervolume

Fig. 3 The 2-D illustration of the hypervolume improvement

$$I_{\text{hv}}(\mathbf{y}) = H(\mathbf{P} \cup \mathbf{y}) - H(\mathbf{P}). \tag{20}$$

As the $I_{\text{hv}}(\mathbf{y})$ being the difference between two hypervolumes, it inherits the computational burden of hypervolume when the number of objectives is larger than two. Therefore, a simplified hypervolume improvement function is proposed in [29] and defined to be the minimum of the k hypervolume improvements of $\mathbf{y}$ with respect to each point on the PF approximation, which is expressed as:
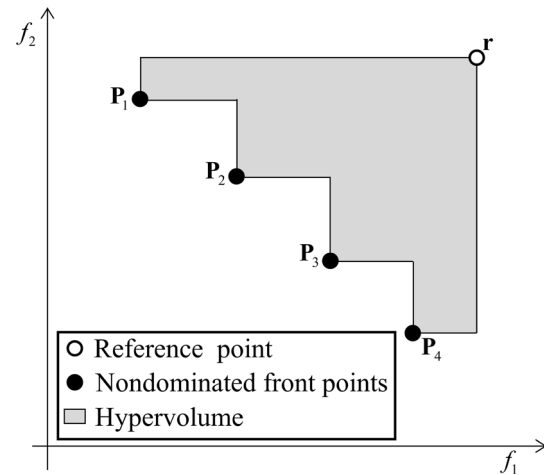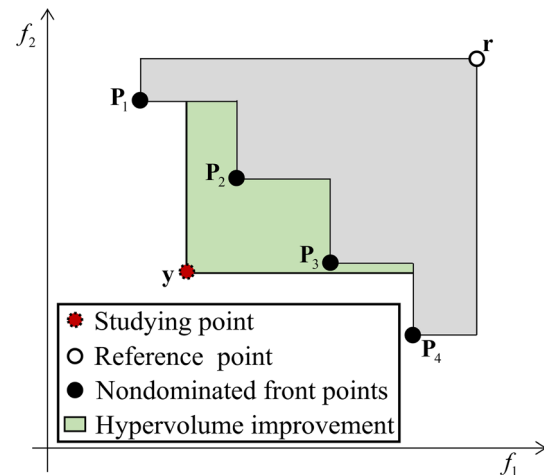
$$I_{\text{hvs}}(\mathbf{x}) = \min_{j=1}^{k} \left[H(\mathbf{P}_j \cup \mathbf{y}) - H(\mathbf{P}_j) \right], \tag{21}$$

where $\mathbf{P}_j = \left(f_1^j, f_2^j, \ldots, f_i^j, \ldots, f_M^j \right)$. The above formula is cheap-to-evaluate as the hypervolume improvement of a point over the other single point can be calculated analytically and formula (21) can be expressed as:

$$I_{\text{hvs}}(\mathbf{x}) = \min_{j=1}^{k} \left[\prod_{i=1}^{M} \left(r_i - y_i(\mathbf{x}) \right) - \prod_{i=1}^{M} \left(r_i - f_i^j \right) \right]. \tag{22}$$

Notably, in Eq. (22), $y_i(\mathbf{x}) = f_i^j - \left(f_i^j - y_i(\mathbf{x}) \right)$. Then taking the form of the simplified hypervolume improvement and replacing the improvement $f_i^j - y_i(\mathbf{x})$ by the

corresponding expected improvement, i.e. the (i,j) element of the VFEIM $\mathrm{EI}_{\mathrm{vf}}^{i,j}(\mathbf{x},l)$, the VFEIM can be aggregated as:

$$\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l) = \min_{j=1}^{k} \left[\prod_{i=1}^{M} \left[r_i - \left(f_i^j - \mathrm{EI}_{\mathrm{vf}}^{i,j}(\mathbf{x},l) \right) \right] - \prod_{i=1}^{M} \left(r_i - f_i^j \right) \right]. \tag{23}$$

It is clear to note that the $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l)$ is cheap-to-evaluate as it is expressed in the closed-form. Though the $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l)$ is obtained by simply taking the form of the simplified hypervolume improvement function, it is still interpretable. The following example with two objectives is adopted to illustrate this. Assuming an arbitrary point $\mathbf{P}_j$ on the current PF approximation set as shown in Fig. 4, then the expected improvements over $\mathbf{P}_j$ with respect to a studying point $\mathbf{y}(\mathbf{x}) \in \mathbb{R}^2$ are $\left(\mathrm{EI}_{\mathrm{vf}}^{1,j}(\mathbf{x},l), EI_{\mathrm{vf}}^{2,j}(\mathbf{x},l) \right)$ and illustrated as the dashed arrows in the figure. With this, the term $\left(f_1^j - \mathrm{EI}_{\mathrm{vf}}^{1,j}(\mathbf{x},l), f_2^j - \mathrm{EI}_{\mathrm{vf}}^{2,j}(\mathbf{x},l) \right)$ can be interpreted as the expected position of the concerned non-dominated front point $\mathbf{P}_j$, denoted as $\mathbf{P}'_j$ in the figure. The hypervolume improvement of $\mathbf{P}'_j$ over $\mathbf{P}_j$ will be calculated exactly as $\prod_{i=1}^{2} \left[r_i - \left(f_i^j - EI_{\mathrm{vf}}^{i,j}(\mathbf{x},l) \right) \right] - \prod_{i=1}^{2} \left(r_i - f_i^j \right)$, which is termed as the simplified expected hypervolume improvement and indicated by the region filled with the darker color.

It is worth to note that aggregated VFEIM function $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l)$ only relates to the improvement of the HF approximation of PF. The aggregated VFEIM function for the HF level, $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l=2)$, indicates the expected improvement of the HF approximation of PF if an HF sample point is added. But aggregated VFEIM function for the LF level, $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l=1)$, is the expected improvement of the HF approximation of PF when an LF sample is added. In a word, the infillings of both LF and HF samples are toward

to improve the HF approximation of PF, which is a unique property of the VFEIM-based criteria.

As the $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l)$ function takes the minimal of the k simplified expected hypervolume improvements, then the PF approximation will get the integral improvement if $\mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l)$ is maximized. In other words, in VFMOEGO with the VFEIM-based criterion, the new sample point $\mathbf{x}$ and its fidelity level l can be determined by solving the following maximization problem:

$$\mathbf{x}, l = \underset{\mathbf{x} \in \Omega, l=1,2}{\arg\max} \ \mathrm{VFEIM}_{\mathrm{h}}(\mathbf{x},l). \tag{24}$$

If the fidelity level $l = 1$, which means that adding an LF sample will improve the PF approximation most, the true responses of the sample point $\mathbf{x}$ will be obtained by calling the cheaper analysis tool; otherwise, the new sample will be evaluated at the HF level (i.e. $l = 2$). It should be mentioned that the responses of the sample point $\mathbf{x}$ for each objective can also be evaluated at different fidelity levels by solving the following sub-optimization problem:

$$\mathbf{x}, \mathbf{l} = \underset{\mathbf{x} \in \Omega, \mathbf{l} \in \Omega_l}{\arg\max} \ VFEIM_{\mathrm{h}}(\mathbf{x}, \mathbf{l}), \tag{25}$$

where $\mathbf{l}$ is now a vector containing the fidelity levels for each objective to be evaluated at; Ω_l represents the design space of the fidelity level. For a 2-objective problem with two fidelity levels, the possible value for $\mathbf{l}$ includes $[(1,1),(1,2),(2,1),(2,2)]$. The criterion (25) will lead to a more flexible VFMOEGO method and it is possible to utilizing such setting in the above ISC. Empirically to say, it depends on the problem to be solved in choosing (24) or (25). If the objectives are all obtained by using a single simulator or the evaluation costs for the different objectives are identical, the criterion (24) should be adopted as it will reduce the implementation effort and the difficulty in solving the mixed-integer maximization problem. Otherwise, the criterion (25) is recommended. For example, for an aircraft wing optimization problem, the drag coefficient C_{D} is to be minimized and the lift coefficient C_{L} is to be maximized. As both the drag and lift coefficient are calculated through the computational fluid dynamic (CFD) simulator, it would be more convenient to use criterion (24) as the CFD simulator will be invoked once in evaluating the sample point $\mathbf{x}$ at the identical fidelity level. While, if the drag coefficient and the structural stress of the wing are to be minimized, two simulators are involved to obtain the responses no matter what the fidelity level vector is. Hence, the more flexible criterion (25) is recommended. As the two real-world problems to be solved later are both based on CFD simulations, the criterion (24) is to be used hereafter.
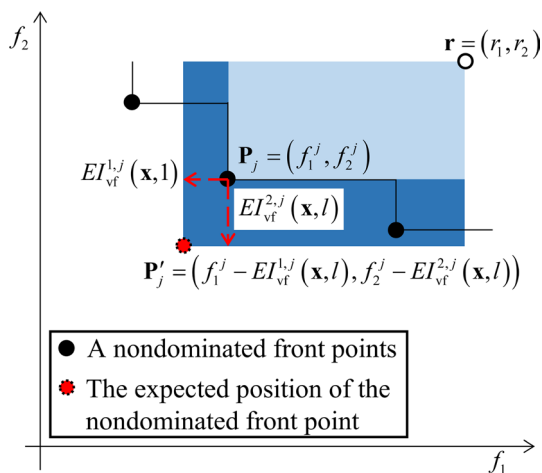

Fig. 4 2-D illustration of the simplified hypervolume improvement

3.2 Variable-fidelity expected modified hypervolume improvement-based criterion

The second criterion is developed based on a modified hypervolume improvement function proposed in [32]. First, the hypervolume improvement function in (20) is reorganized into the following form by introducing the local upper bounds **U** (LUBs) [53]:

$$I_{\mathrm{hv}}(\mathbf{x}) = \sum_{j=1}^{m} V\big([\mathbf{y}, \mathbf{U}_j]\big) - V_{\mathrm{overlap}}(\mathbf{y}), \quad (26)$$

where m indicates the number of LUBs, $\mathbf{U}_j \in \mathbb{R}^M, j = 1, 2, \dots m$ represents the LUBs related to the current PF approximation. The LUBs denote a set of points that cannot be dominated by one another that $\forall \mathbf{U}_i \in \mathbf{U}, \forall \mathbf{U}_j \in \mathbf{U} \Rightarrow \mathbf{U}_i \nprec \mathbf{U}_j$ (The procedure and code to generate LUBs with a reference point and an initial PF approximation can be found in [32]). With LUBs, the nondominated region can be defined as $\Theta = \big\{ \mathbf{a} \in \mathbb{R}^M | \exists \mathbf{U}_i \in \mathbf{U} : \mathbf{a} \ll \mathbf{U}_i \big\}$, ($\ll$ means strong Pareto-dominance, i.e. $\mathbf{a} \ll \mathbf{b} \Leftrightarrow \forall 1 \le i \le M : a_i < b_i$). The relationship between the LUBs **U** and current PF approximation is exemplified in Fig. 5. Detail properties of the LUBs can be found in [53]. In (26), $V\big([\mathbf{y}, \mathbf{U}_i]\big)$ represents the hypervolume of the hyper-rectangle with the two points **y** and $\mathbf{U}_j$ as the diagonal points, the first term $\sum_{j=1}^{m} V\big([\mathbf{y}, \mathbf{U}_j]\big)$ denotes the volume sum of those regions dominated by **y**, $V_{\mathrm{overlap}}(\mathbf{y})$ is the volume of the overlapped part in the first term. As shown in Fig. 5, the areas with different shaded lines represent the regions for different $[\mathbf{y}, \mathbf{U}_j]$ and the colored area relates to the value of V_{overlap}.

As the $V_{\mathrm{overlap}}(\mathbf{y})$ cannot be calculated analytically when the number of objectives is larger than two. Hence a
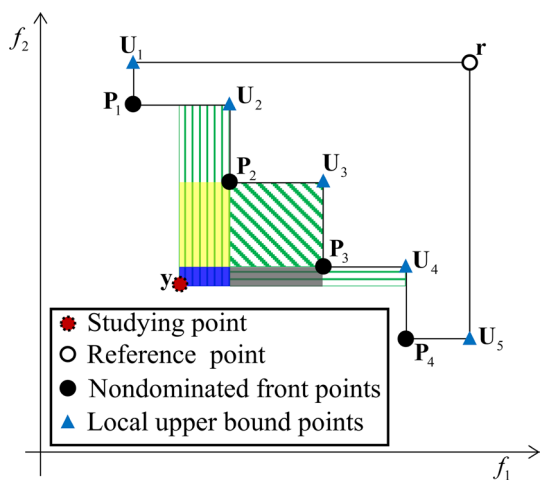
modified hypervolume improvement function is defined by neglecting the second term of (26) and expressed in detail as:

$$I_{\mathrm{hvm}}(\mathbf{x}) = \sum_{j=1}^{m} V\big([\mathbf{y}, \mathbf{U}_j]\big) = \sum_{j=1}^{m} \left[\prod_{i=1}^{M} \big(U_j^i - y_i(\mathbf{x}) \big) \right]. \quad (27)$$

It has been proved that $I_{\mathrm{hvm}}(\mathbf{x}) \ge 0$ and $I_{\mathrm{hvm}}(\mathbf{x}) > 0$ when and only when $\mathbf{y} \in \Theta$ [32]. Moreover, it has been demonstrated that the above modified hypervolume improvement shares the identical properties with the original hypervolume improvement function in MOO, e.g. the improvement function is monotonically decreasing with increasing $y_i(i = 1, 2, \cdots M)$. While the modified hypervolume improvement can be calculated analytically compared with the original one. Recall that, in VFMOEGO, the objective vector $y_i(\mathbf{x}), i = 1, 2, \dots, M$ of the studying point **x** follows M-dimensional independent Gaussian distribution as shown in (14), the expectation of the modified hypervolume improvement (VFEMHVI) function can be derived as:

$$\mathrm{VFEMHVI}(\mathbf{x}, l) = E\big[I_{\mathrm{hvm}}(\mathbf{x})\big] = E\left[\sum_{j=1}^{m} \left[\prod_{i=1}^{M} \big(U_j^i - y_i(\mathbf{x}) \big) \right] \right],$$

$$= \sum_{j=1}^{m} \left[\prod_{i=1}^{M} E\big(U_j^i - y_i(\mathbf{x}) \big) \right] = \sum_{j=1}^{m} \left[\prod_{i=1}^{M} EI_{\mathrm{vf}}^{i,j}(\mathbf{x}, l, U_j^i) \right] \quad (28)$$

where

$$\mathrm{EI}_{vf}^{i,j}(\mathbf{x}, l, U_j^i) = \big(U_j^i - \hat{y}_i(\mathbf{x}) \big) \Phi\left(\frac{U_j^i - \hat{y}_i(\mathbf{x})}{s_i(\mathbf{x}, l)} \right)$$
$$+ s_i(\mathbf{x}, l) \phi\left(\frac{U_j^i - \hat{y}_i(\mathbf{x})}{s_i(\mathbf{x}, l)} \right). \quad (29)$$

It can be also noted that the infillings of both LF and HF samples based on VFEMHVI are toward improving the HF approximation of PF. Moreover, Eq. (29) takes the identical form of Eq. (17) but the values of LUBs are used in (29) as the reference value instead of the current PF approximation. Thus Eq. (28) can be regarded as the aggregation result of the variable-fidelity LUBs-based expected improvement matrix which is expressed as:

$$\begin{bmatrix} \mathrm{EI}_{vf}^{1,1}(\mathbf{x}, l, U_1^1) & EI_{vf}^{2,1}(\mathbf{x}, l, U_1^2) & \cdots & \mathrm{EI}_{vf}^{M,1}(\mathbf{x}, l, U_1^M) \\ \mathrm{EI}_{vf}^{1,2}(\mathbf{x}, l, U_2^1) & EI_{vf}^{2,2}(\mathbf{x}, l, U_2^2) & \cdots & EI_{vf}^{M,2}(\mathbf{x}, l, U_2^M) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{EI}_{vf}^{1,m}(\mathbf{x}, l, U_m^1) & EI_{vf}^{2,m}(\mathbf{x}, l, U_m^2) & \cdots & \mathrm{EI}_{vf}^{M,m}(\mathbf{x}, l, U_m^M) \end{bmatrix}. \quad (30)$$

In other words, the VFEMHVI function can also be regarded as based on the VFEIM concept. But the reference values used in each element of the VFEIM and the aggregation scheme are different from that of the VFEIM$_h$. For

Fig. 5 LUBs and the modified hypervolume improvement

VFMOEGO with the VFEMHVI-based criterion, the new sample point $\mathbf{x}$ and its fidelity level l can be determined by solving the following maximization problem:

$$\mathbf{x}, l = \arg\max_{\mathbf{x}\in\Omega, l=1,2} \text{VFEMHVI}(\mathbf{x}, l). \tag{31}$$

The VFEMHVI criterion is derived analytically hence to be computational efficient when solving (31). It should be mentioned that, in criterion (31), treating l as a vector as that of the VFEIM-based criterion in (25) is also possible. The reason for proposing a criterion with l as a scalar is identical to that explained in the last subsection.

3.3 Constraints handling

For a problem with M_C computationally expensive inequality constraints $g_i(\mathbf{x}) \le 0, i = 1, \dots, M_C$, the constraints are also modeled by HK. The prediction of the HK model for the constraints at a point can be considered as the following normal distribution,

$$\hat{G}_i(\mathbf{x}, l) \sim N\left[\hat{g}_i(\mathbf{x}), s_{g,i}^2(\mathbf{x}, l)\right], \ l = 1, 2, i = 1, \dots, M_C, \tag{32}$$

where $\hat{g}_i(\mathbf{x})$ is the prediction of the ith HK model for constraint, $s_{g,i}^2(\mathbf{x}, l), l = 1, 2$ represents the uncertainty of the ith model due to the lack of LF and HF samples. Then, the probability of satisfying the ith constraint is

$$P[G_i(\mathbf{x}, l) \le 0] = \Phi\left(\frac{0 - \hat{g}_i(\mathbf{x})}{s_{g,i}(\mathbf{x}, l)}\right). \tag{33}$$

Eventually, the criteria for the constrained problem can be formulated as:

$$\text{CVFEIM}_h(\mathbf{x}, l) = \text{VFEIM}_h(\mathbf{x}, l) \cdot \prod_{i=1}^{M_C} P\left[G_i(\mathbf{x}, l) \le 0\right], \tag{34}$$

$$\text{CVFEMHVI}(\mathbf{x}, l) = \text{VFEMHVI}(\mathbf{x}, l) \cdot \prod_{i=1}^{M_C} P\left[G_i(\mathbf{x}, l) \le 0\right]. \tag{35}$$

By maximizing the above criteria, the spatial location of a new sample point and its fidelity level to be evaluated at can then be determined.

3.4 Implementation of the VFMOEGO method

The algorithm 1 is given as the implementation of the VFMOEGO with the VFEMHVI criterion and the corresponding framework is sketched in Fig. 6a. It is noted that a reference point is required from the user for both criteria. In this work, except for imputing the reference point from the user, introducing a normalization procedure to release such a burden for the user is also available. In detail, all the LF and HF objective responses are scaled using the minima and maxima of the HF objective responses of the current HF samples, then the reference point is set to $r_j = 1.1$ for $j = 1, 2, \dots, M$. Figure 6b shows the VFMOEGO method using such a normalization procedure. The influence of the two strategies to specify the reference point will be discussed within the section of numerical experiments. For VFMOEGO with VFEIM$_h$ criterion, the procedure is identical thus not repeated.

Fig. 6 Flowchart of the VFMOEGO method with **a** reference point provided by user and **b** introducing normalization procedure
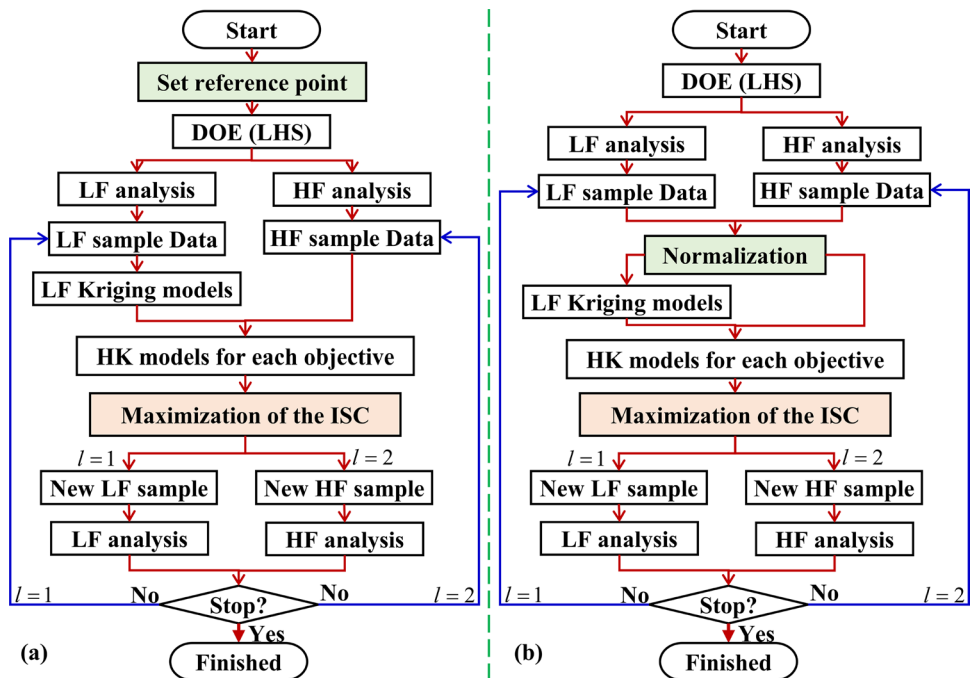
Algorithm 1: Procedure of VFMOEGO method with VFEMHVI

1:	set initial LF and HF sample points, $J = 0$
2:	evaluate the samples by LF and HF analyses, respectively
3:	set reference point
4:	**while** $J < J_{max}$
5:	construct the LF Kriging models for each objective
6:	construct the HK models for each objective
7:	find $\mathbf{x}_{LF}^* = \arg\max \text{VFEMHVI}(\mathbf{x}, l = 1)$
8:	find $\mathbf{x}_{HF}^* = \arg\max \text{VFEMHVI}(\mathbf{x}, l = 2)$
9:	**If** $\text{VFEMHVI}(\mathbf{x}_{LF}^*, l = 1) > \text{VFEMHVI}(\mathbf{x}_{HF}^*, l = 2)$ **then**
	$\mathbf{x}_{new} = \mathbf{x}_{LF}^*, l = 1$
10:	**else**
	$\mathbf{x}_{new} = \mathbf{x}_{LF}^*, l = 1$
	end if
11:	**if** $(l = 1)$ **then**
	evaluate $\mathbf{x}_{new}$ using LF analysis, go to 5
12:	**else**
	evaluate $\mathbf{x}_{new}$ using LF analysis, go to 6
	end if
13:	**end while**

3.5 Illustration of the proposed criteria and method

In this subsection, a two-objective ZDT2 test problem [42] is employed to illustrate the proposed criteria and methods. The formulation of the ZDT2 problem is given as:

$$\min_{\mathbf{x} \in \Omega} F(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x})\}$$

with assistance of $F_{lf}(\mathbf{x}) = \{f_{1,lf}(\mathbf{x}), f_{2,lf}(\mathbf{x})\}$

$$f_1(\mathbf{x}) = f_{1,lf}(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x}) \times h(\mathbf{x})$$

$$f_{2,lf}(\mathbf{x}) = (0.9 \times g(\mathbf{x}) + 1.1) \times (1.1 \times h(\mathbf{x}) - 0.1) \qquad (36)$$

where $g(\mathbf{x}) = 1 + \dfrac{9}{n-1} \sum_{i=2}^{n} x_i$

$$h(\mathbf{x}) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}$$

$$0 \leq x_i \leq 1, i = 1, \ldots, n.$$

First, the VFEIM$_h$ criterion employing the normalization procedure, the VFEMHVI criterion with the user specifying the reference point, and the VFEMHVI utilizing the normalization procedure (denoted as VFEMHVIN for short) are compared. For the reference point of the ZDT2 problem, $\mathbf{r} = [11, 11]$ is adopted here. Identical initial LF/HF samples are used to calculate the 2-objective criteria and the contours in the 2-D design space are shown in Fig. 7, respectively. In these figures, the dots present the sampled points, the color of the contours indicates the value of the criteria. The darker the color the higher the value. First, it can be noted that, by comparing Fig. 7a, b, the contours of the VFEIM$_h$ criteria at LF level (i.e. $l = 2$) have very similar trends to its corresponding HF criteria (i.e. $l = 2$) but differ in magnitude. This observation holds for VFEMHVI criteria as shown in Fig. 7c–f. Through comparing the VFEIM$_h$ and VFEMH-VIN criterion at HF level by observing the Fig. 7b, f, it is expected to see that the trend and magnitude of the contours

Fig. 7 Plots of the criteria in the design space of ZDT2 problem. **a** VFEIM$_h$ with $l = 1$. **b** VFEIM$_h$ with $l = 2$. **c** VFEM-HVI with $l = 1$. **d** VFEMHVI with $l = 2$. **e** VFEMHVIN with $l = 1$. **f** VFEMHVIN with $l = 2$
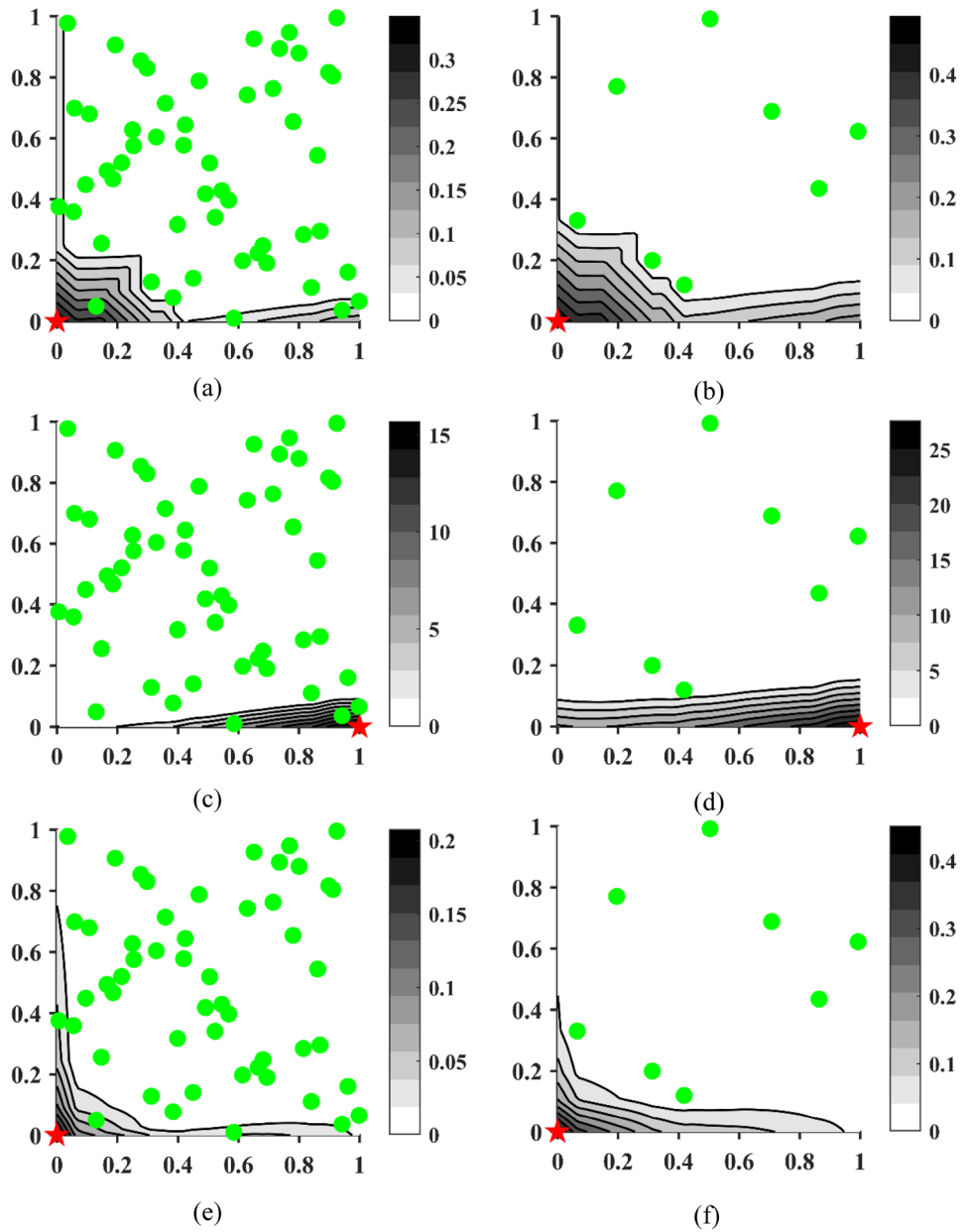

of those two criteria are different as the two criteria differ at the reference values used and the aggregation scheme. The dissimilarity of the criteria at the LF level can also be noticed through Fig. 7a, e. Though the trend and magnitude of the VFEIM$_h$ and VFEMHVIN criterion are different, they suggest the identical next sample point shown as the pentagrams in those figures. The third observation that can be noticed is the influence of the strategy on specifying the reference point. With Fig. 7d, f shown the VFEMHVI and VFEMHVIN criterion at the HF level, not only the trend but also the next sample point suggested are disparate. This indicates the performance of the method utilizing the

normalization procedure might deviate from that by receiving the reference point from the user.

Second, the first five refine steps of the VFMOEGO method with different criteria in solving the ZDT2 problem of six design variables using identical initial LF and HF samples are sketched in Fig. 8 and Tables 1, 2, and 3. The initial HF non-dominated solutions (HFNDS), representing the initial HF approximation of PF, are represented by the filled squares. The added LF/HF samples in the iterative optimization procedure are shown by filled triangles/circles. The true PF of the ZDT2 problem is also presented in those figures. It can be noted that all methods select LF/HF samples adaptively in the iterative optimization process, which is the

Fig. 8 Refinement process of the VFMOEGO method on the ZDT2 problem with different criterion: **a** VFEIM$_h$, **b** VFEM-HVI, **c** VFEMHVIN

(a)

(b)

(c)
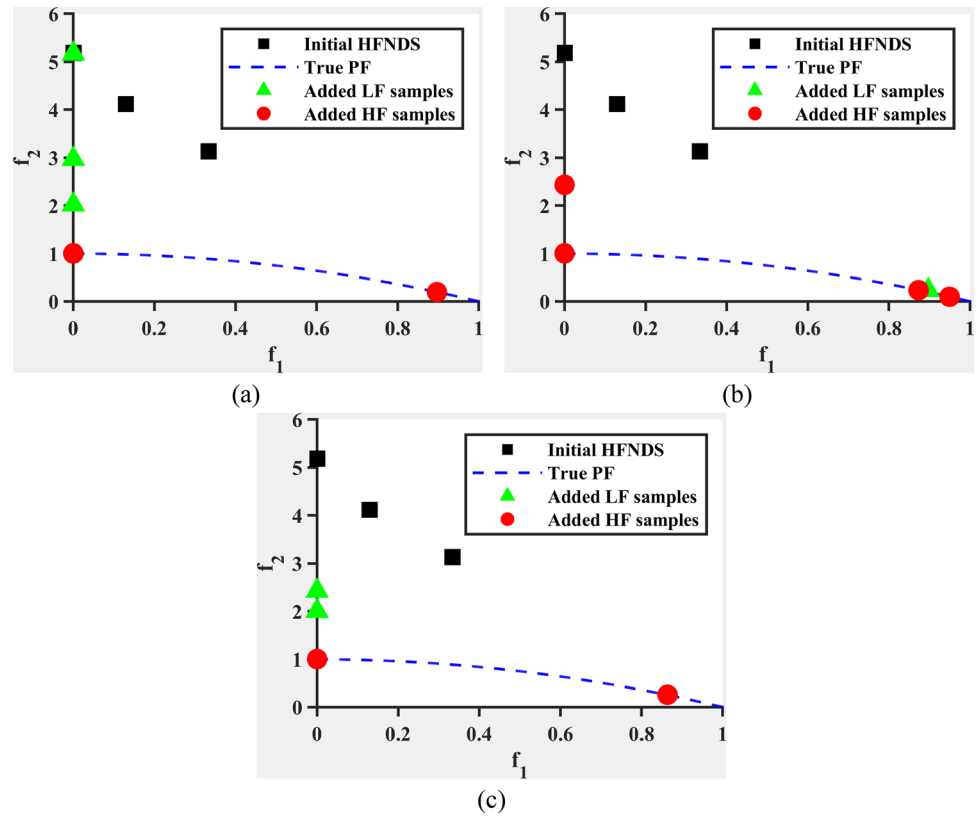
Table 1 Refinement process of the VFEIM$_h$-based optimization of the ZDT2 problem

Iteration	Added sample	Fidelity level	NHV
1	$[f_1, f_2]$ =[6.2e−5, 2.0165]	1 (LF)	0
2	$[f_1, f_2]$ =[0, 1.0001]	2 (HF)	0.0908
3	$[f_1, f_2]$ =[0, 5.1639]	1 (LF)	0.0908
4	$[f_1, f_2]$ = [0.8970, 0.1957]	2 (HF)	0.2257
5	$[f_1, f_2]$ = [0, 2.9710]	1 (LF)	0.2257

Table 2 Refinement process of the VFEMHVI-based optimization of the ZDT2 problem

Iteration	Added sample	Fidelity level	NHV
1	$[f_1, f_2]$ =[0.8969, 0.2339]	1 (LF)	0
2	$[f_1, f_2]$ =[0.9496, 0.0987]	2 (HF)	0.1245
3	$[f_1, f_2]$ =[2.9e-5, 1.0036]	2 (HF)	0.2001
4	$[f_1, f_2]$ =[0.8728, 0.2385]	2 (HF)	0.2487
5	$[f_1, f_2]$ =[0, 2.4314]	2 (HF)	0.2487

Table 3 Refinement process of the VFEMHVIN-based optimization of the ZDT2 problem

Iteration	Added sample	Fidelity level	NHV
1	$[f_1, f_2]$ =[3.7e−5, 2.0077]	1 (LF)	0
2	$[f_1, f_2]$ =[1.8e−5, 2.0016]	1 (LF)	0
3	$[f_1, f_2]$ =[0, 1.003]	2 (HF)	0.0906
4	$[f_1, f_2]$ =[0, 2.4295]	1 (LF)	0.0906
5	$[f_1, f_2]$ =[0.8637, 0.2584]	2 (HF)	0.2355

unique property of the proposed criteria. For example, the five sample points selected by VFEIM$_h$ are evaluated at the LF–HF–LF–HF–LF level. While the VFEMHVI determined the first five points to be evaluated at LF–HF–HF–HF–HF level. In terms of the spatial position the samples added in

the first five iterations (see Tables 1, 2, and 3), the three criteria have diverse favors. For instance, the first point chosen by the three criteria is all evaluated at the LF level but with different spatial locations, i.e. [6.2e−5, 2.0165], [0.8969, 0.2339], and [3.7e−5, 2.0077] respectively. Even though, the performance metric, i.e. the normalized hypervolume (NHV), is increased continuously indicating that the HF approximation of PF is improved remarkably by all methods (The NHV will be detailed in the section of the numerical experiment). The above illustrations demonstrate the effectiveness of the proposed criteria on improving the HF approximation of PF and the ability to adaptively select LF and HF samples in the iterative phase of the VFMOEGO method. More numerical examples are included in the next section to support this observation.
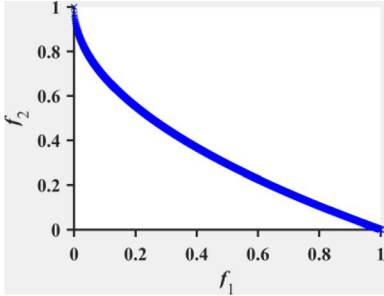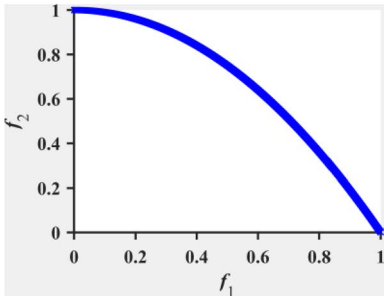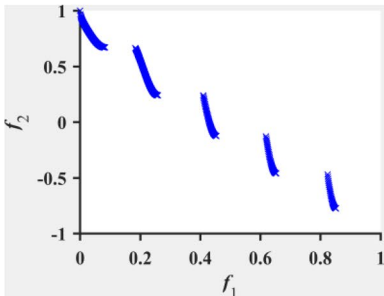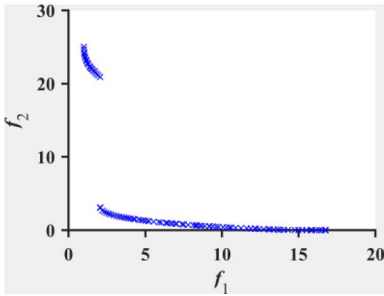
4 Examples and results

4.1 Analytic test cases

In this subsection, analytic numerical examples are adopted to demonstrate the applicability and efficiency of the proposed criteria. These examples are also solved by the corresponding single-fidelity methods employing the criteria of HF level only, i.e. setting $l = 2$ in (24) or (31). The three single-fidelity methods utilizing HF simulations are denoted as EIM_h, EMHVI, EMHVIN, and the three variable-fidelity approaches are termed as $VFEIM_h$, VFEMHVI, VFEMHVIN, respectively.

Table 4 Formulations and true PF for numerical examples

Problem	Formula	True PF
ZDT1	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times h(\mathbf{x})$ $f_{2,\mathrm{lf}}(\mathbf{x}) = (0.8g(\mathbf{x}) - 0.2) \times (1.2h(\mathbf{x}) + 0.2)$ where $g(\mathbf{x}) = 1 + \dfrac{9}{n-1}\sum_{i=2}^{n} x_i$ $h(\mathbf{x}) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}$ $n = 6$ $0 \le x_i \le 1, i = 1,\ldots,n$	
ZDT2	$f_1(\mathbf{x}) = f_{1,\mathrm{lf}}(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times h(\mathbf{x})$ $f_{2,\mathrm{lf}}(x) = (0.9g(\mathbf{x}) + 1.1) \times (1.1h(\mathbf{x}) - 0.1)$ $g(\mathbf{x}) = 1 + \dfrac{9}{n-1}\sum_{i=2}^{n} x_i$ $h(\mathbf{x}) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}$ $n = 6$ $0 \le x_i \le 1, i = 1,\ldots,n$	
ZDT3	$f_1(\mathbf{x}) = f_{1,\mathrm{lf}}(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times h(\mathbf{x})$ $f_{2,\mathrm{lf}}(\mathbf{x}) = (0.75g(\mathbf{x}) - 0.25) \times (1.25h(\mathbf{x}) + 0.25)$ $g(\mathbf{x}) = 1 + \dfrac{9}{n-1}\sum_{i=2}^{n} x_i$ $h(\mathbf{x}) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})} - (f_1(\mathbf{x})/g(\mathbf{x}))\sin(10\pi f_1)$ $n = 6$ $0 \le x_i \le 1, i = 1,\ldots,n$	
POL	$f_1(\mathbf{x}) = \left[1 + (A_1 - B_1)^2 + (A_2 - B_2)^2\right]$ $f_{1,\mathrm{lf}}(\mathbf{x}) = \left[1 + (0.9A_1 - 1.2B_1)^2 + 0.9(1.2A_2 - 0.9B_2)^2\right]$ $f_2(\mathbf{x}) = f_{2,\mathrm{lf}}(\mathbf{x}) = (x_1 + 3)^2 + (x_2 + 1)^2$ where $A_1 = 0.5\sin 1 - 2\cos 1 + \sin 2 - 1.5\cos 2$ $A_2 = 1.5\sin 1 - \cos 1 + 2\sin 2 - 0.5\cos 2$ $B_1 = 0.5\sin x_1 - 2\cos x_1 + \sin x_2 - 1.5\cos x_2$ $B_2 = 1.5\sin x_1 - \cos x_1 + 2\sin x_2 - 0.5\cos x_2$ $-\pi \le x_i \le \pi, i = 1,2$	

For clarity, the procedure of single-fidelity MOEGO employing the EIM_h criteria is introduced briefly. In the initialization phase, HF initial samples are generated by DOE method and then HF simulator is invoked to obtain the responses of those HF samples. The iterative optimization stage proceeds with the construction of the Kriging models for each objective function based on the initial data set. Then the spatial position of the next sample point is obtained by maximizing the EIM_h function (i.e. $EIM_h(\mathbf{x}) = VFEIM_h(\mathbf{x}, l = 2)$). Third, the simulator is called to get the HF response of the sample point. The HF data set is augmented before the check of the termination condition. The iterative optimization phase will continue iteratively until the stopping rule is met.

The formulations of analytic numerical examples from [41, 42] are summarized in Table 4. The computational cost ratio (CR) between an HF evaluation and an LF one for all the test cases is set to be 4. On 6-D ZDT problems, the initial HF sample size for single-fidelity MOEGO method is set as $11d$-1 following [29], where d represents the dimension of the design variable; for variable-fidelity methods, the initial HF and LF sample sizes are set as $5d$ and $(6d$-1)*CR to ensure the cost on the evaluation of the initial samples are identical to that of the single-fidelity methods. On the 2-D POL problem, 21 HF initial samples are used in single-fidelity method, 16 HF and 20 LF samples are adopted for the variable-fidelity methods. For the maximization of the VFEIM or VFEMHVI-based criterion, which is a mix-integer problem of one integer variable with two levels, it is maximized at LF (i.e. $l = 1$) and HF (i.e. $l = 2$) level respectively by the genetic algorithm (GA) as depicted in

Algorithm 1. The GA is used with the following settings: the maximum generation being 100, population size being 100, probability of mutation and crossover being 0.2 and 0.8, respectively. Such settings are also employed to maximize single-fidelity criteria. A maximal equivalent iteration number for the iterative phase is adopted as the termination condition of each algorithm. For the equivalent iteration number is expressed as:

$$\text{EquivIter} = \frac{1}{\text{CR}}\text{iter}_{LF} + \text{iter}_{HF}, \tag{37}$$

where iter_{LF} and iter_{HF} represent the number of iterations choosing an LF and HF simulation calls. Notably, the number of equivalent iterations reduces to the iteration number for the single-fidelity method. The maximal allowed equivalent iteration for all methods on the 6-D ZDT problems and the 2-D POL problem is set as 100 and 70, respectively. Considering the randomness of the LHS method and GA, all problems are solved by each method with 11 times. The reference point needed for EMHVI and VFEMHVI is set as (11, 11) for three ZDT problems and (80, 80) for the POL problem.

The numerical experiments over those analytic problems are conducted with the aid of PlatEMO, an open-source platform for comparison of multi-objective optimization method [54]. In the platform, several metrics are available to measure the quality of solutions obtained. In this work, two metrics are employed to compare the performance of different methods on solving analytic problems, that is, normalized hypervolume (NHV) [55] and inverted generational distance (IGD) [56] that measure the proximity to the true

Table 5 Statistics of the NHV values of the final PF approximation obtained by the compared methods

Problem	ZDT1	ZDT2	ZDT3	POL	+/–/=
EIM_h					
Mean	6.6025e–1	1.5710e–1	6.0475e–1	8.7684e–1	0/3/1
STD	2.34e–2 (–)	7.75e–2 (–)	6.60e–2 (=)	6.50e–4 (–)	
EMHVI					
Mean	6.9726e–1	3.2743e–1	5.4310e–1	8.7745e–1	0/2/2
STD	6.62e–3 (=)	8.58e–2 (–)	2.88e–2 (–)	1.85e–4 (=)	
EMHVIN					
Mean	6.9987e–1	3.0096e–1	5.8567e–1	8.7749e–1	0/1/3
STD	4.19e–3 (=)	8.04e–2 (–)	5.34e–2 (=)	1.30e–4 (=)	
$VFEIM_h$					
Mean	6.9720e–1	4.1127e–1	5.8321e–1	8.7545e–1	0/2/2
STD	4.29e–3 (=)	9.40e–3 (–)	1.15e–2 (=)	1.54e–3 (–)	
VFEMHVI					
Mean	7.0220e–1	4.2088e–1	5.7471e–1	8.7738e–1	0/0/4
STD	4.14e–3 (=)	3.28e–3 (=)	1.98e–3 (=)	2.17e–4 (=)	
VFEMHVIN					
Mean	7.0106e–1	4.1928e–1	5.7546e–1	8.7751e–1	
STD	5.07e–3	3.93e–3	2.43e–3	1.31e–4	

Table 6 Statistics of the IGD values of the final PF approximation obtained by the compared methods

	ZDT1	ZDT2	ZDT3	POL	+/–/=
EIM$_h$					
Mean	4.7183e−2	4.2143e−1	7.0382e−2	2.1338e−1	0/4/0
STD	1.39e−2 (−)	1.62e−1 (−)	2.50e−2 (−)	5.87e−2 (−)	
EMHVI					
Mean	2.3495e−2	1.7151e−1	6.5344e−2	1.4722e−1	0/3/1
STD	4.85e−3 (−)	1.53e−1 (−)	2.30e−2 (−)	1.83e−2 (=)	
EMHVIN					
Mean	2.1767e−2	1.8908e−1	8.8467e−2	1.3975e−1	0/3/1
STD	2.82e−3 (−)	1.46e−1 (−)	3.54e−2 (−)	8.25e−3 (=)	
VFEIM$_h$					
Mean	2.3948e−2	3.5003e−2	2.6038e−2	1.7746e−1	0/2/2
STD	3.69e−3 (−)	7.48e−3 (=)	3.95e−3 (=)	3.50e−2 (−)	
VFEMHVI					
Mean	1.8361e−2	2.8651e−2	2.8579e−2	1.4750e−1	0/0/4
STD	3.16e−3 (=)	2.89e−3 (=)	5.41e−3 (=)	2.61e−2 (=)	
VFEMHVIN					
Mean	1.9028e−2	2.9102e−2	2.6137e−2	1.3581e−1	
STD	3.75e−3	4.05e−3	4.08e−3	1.29e−2	

PF and the diversity of solutions simultaneously. NHV and IGD have widely been adopted for fair assessing the performance of multi-objective algorithms [57, 58]. A larger value of NHV or a smaller value of IGD indicates better quality of the HF approximation of the PF. The definition of reference points for NHV calculation and true PF for IGD calculation of those analytic problems can be found in [54]. The number of HF solutions to approximate the PF is also counted to show the differences among all compared methods. The Wilcoxon rank-sum test with significance level 0.05 is applied to identify whether the means of the results obtained by one criterion is significantly different from the results obtained by the VFEMHVIN criterion. The result is labeled by the signs '+', '−' and '≈', which indicate that the result is significantly better, significantly worst, and comparable statistically.

Tables 5 and 6 present the statistic results of NHV and IGD over the analytic test instances. The Wilcoxon rank-sum test results between one method and VFEMHVIN are presented in the parentheses in those tables. Moreover, the number of one method performs significantly better, worst, and comparable statistically to VFEMHVIN is counted and listed in Tables 5 and 6. From the statistic results of NHV metric shown in Table 5, it can be noted that the VFEMHVIN performs better or comparative to all the rest five methods over all test problems. In detail, compared with the three single-fidelity methods, VFEMHVIN is better on 3, 2, and 1 of the four problems and comparable on 1, 2, and 3 of the four problems. This demonstrates the optimization gains can be improved by utilizing VF surrogate-based method. The VFEIM$_h$ performs statistically worse than VFEMHVIN on

ZDT2 and POL problems, and comparable to VFEMHVIN on ZDT1 and ZDT3 problems. Recall the two differences between VFEIM$_h$ and VFEMHVIN, it can be concluded that the reference values and the adopted aggregation scheme for the variable-fidelity expected improvement matrix affects the performance of the corresponding criterion. The VFEMHVI is statistically comparable to VFEMHVIN over all analytic test cases. As VFEMHVI and VFEMHVIN only differ at the way to generate the reference point, this indicates that the normalization-based strategy to generate the reference point preserves the ability of the VFEMHVI criterion in finding solutions with a high level of convergence and diversity, while is superior on without the need to specify the reference point by the user/designer. Though the effects of reference point specification strategy are not obvious in above test results, it should be admitted that the strategy does affect the performance of the method relying on the hypervolume concept based on our preliminary studies. Such a phenomenon has been pointed out before and more suggestions on specifying the reference point can be found in [59–61]. Inspired by this, a dynamic reference point specification strategy will be investigated the future work. With the IGD statistic results shown in Table 6, the comparison conclusion is identical, i.e. VFEMHVIN performs better or comparative to all the rest five methods over all test problems. While the VFEMHVIN is now superior to EIM$_h$, EMHVI, EMHVIN, and VFEIM$_h$ over 4, 3, 3, and 2 problems.

The number of HF solutions to approximate the PF of each method is counted in Table 7. Generally, the single- and variable-fidelity expected improvement matrix-based methods, i.e. EIM$_h$ and VFEIM$_h$ can obtain more solutions

Table 7 Statistics of the number of HF solutions for the final PF approximation obtained by the compared methods

Problem	ZDT1	ZDT2	ZDT3	POL	+/–/=
EIM$_h$					
Mean	1.9091e+1	2.7273e+0	1.5545e+1	5.7273e+1	1/3/0
STD	3.65e+0 (–)	1.10e+0 (–)	1.69e+0 (–)	2.83e+0 (+)	
EMHVI					
Mean	1.9273e+1	6.0909e+0	1.2909e+1	4.8091e+1	0/4/0
STD	2.87e+0 (–)	2.17e+0 (–)	2.63e+0 (–)	4.35e+0 (–)	
EMHVIN					
Mean	1.9545e+1	5.0000e+0	1.1909e+1	4.9273e+1	0/4/0
STD	1.86e+0 (–)	1.90e+0 (–)	2.63e+0 (–)	2.41e+0 (–)	
VFEIM$_h$					
Mean	3.1182e+1	2.2000e+1	2.4727e+1	5.5818e+1	2/0/2
STD	4.12e+0 (=)	2.97e+0 (+)	3.90e+0 (=)	3.79e+0 (+)	
VFEMHVI					
Mean	2.7909e+1	1.7636e+1	2.6727e+1	4.8727e+1	0/0/4
STD	6.79e+0 (=)	1.86e+0 (=)	4.29e+0 (=)	4.13e+0 (=)	
VFEMHVIN					
Mean	2.8364e+1	1.7091e+1	2.5909e+1	5.1636e+1	
STD	5.18e+0	2.81e+0	3.02e+0	3.20e+0	

to approximate the PF on one or two problems. While the EMHVI, EMHVIN performs worst in terms of the number of solutions on the approximated PF. The ability of VFEM-HVI and VFEMHVIN in this aspect is still comparable. Figure 9 shows the final PF approximation by each algorithm with the median IGD value on the 11 runs on all the analytic test problems. The differences of the number of HF solutions approximating the PF can be observed in Fig. 9 institutively. Moreover, it can be also seen that all methods can obtain promising PF approximations on the ZDT1 and ZDT2 problems with continuous PF. While, on the two problems with discrete PF (ZDT3 and POL), the VF surrogate-based methods approximate the PF better as the single-fidelity method generally missed part of the discrete PF. Especially in the ZDT3 problem, it can be noted that the rightmost part of the discrete PF is missed by the three single-fidelity methods.

Table 8 compares the number of LF/HF samples added during the iterative optimization process of each method. From Table 8, the ability of the VF surrogate-based method in adaptively selecting LF/HF samples during the iterative optimization process can be observed. In terms of the portion of LF or HF samples on the whole sample set added, it depends on the problem. For example, the LF samples chosen by the three variable-fidelity methods took up nearly half of the total added samples on the three ZDT problems. While, on the POL problem, the three variable-fidelity methods spent most computational cost on HF samples. Generally, this is determined by the correlation between LF and HF functions. If the LF function correlates well to the HF function, added more LF samples is beneficial to improve

the accuracy of the HK model and hence the optimization performance.

The Wilcoxon rank-sum tests among the three single-fidelity methods are also conducted to reveal the best one as the comparison among the three single-fidelity methods has never been reported in the open literature. Table 9 shows the test results using the IGD values. It can be noted that the EIM$_h$ is significantly worse than EMHVIN on three cases over the four test problems and performs comparably to EMHVIN on the ZDT3 problem. Again, normalization-based reference point specification strategy still maintains the ability of EMHVI-based method as the comparable performance between EMHVI and EMHVIN can be observed in Table 9.

Finally, the effect of cost ratio on the three variable-fidelity methods is investigated. Though no cost-control strategy is incorporated in the proposed criteria, the cost ratio affects the number of initial LF samples hence the initial surrogates and optimization process. Before presenting the experiment results, it is expected that the variable-fidelity method will perform better if the cost ratio increases. This is due to more LF samples can be used with a larger cost ratio within the setting context. Providing the LF and HF functions correlate well, more LF samples will generally improve the accuracy of the VF surrogate models which will benefit the optimization task. In detail, additional experiments with CR = 3 and CR = 8 are conducted on the ZDT2 and ZDT3 problems with the obtained results shown in Table 10. The results of CR = 4 are taken from Table 5 and 6. Indeed, the statistic results of NHV and IGD values follows the above expectation. For instance,
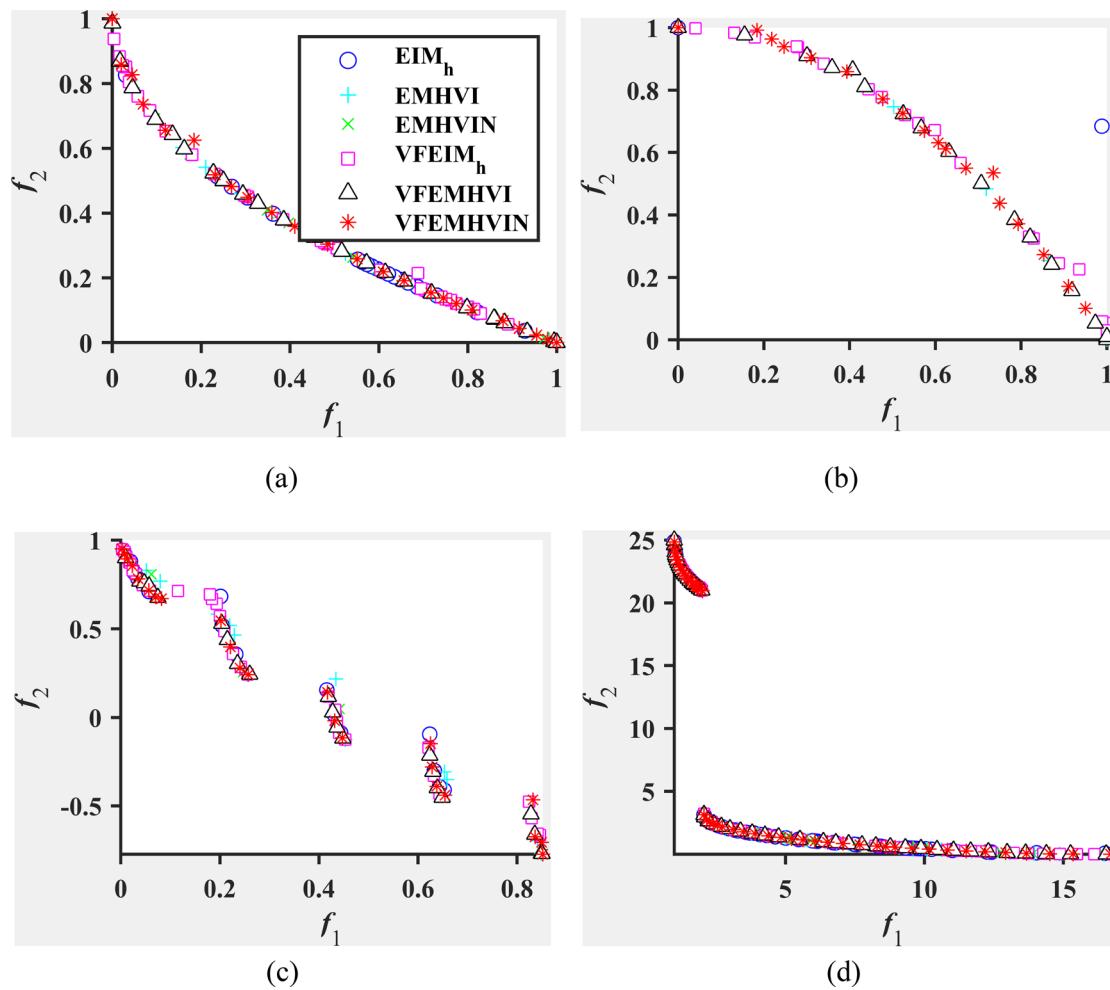
Fig. 9 Plots of the final PF approximation obtained by each algorithm with the median IGD value in the 11 runs on different problems: a ZDT1, b ZDT2, c ZDT3, d POL

Table 8 Added LF and HF samples of each method

	Sizes	ZDT1	ZDT2	ZDT3	POL
EIM_h	LF	–	–	–	–
	HF	100	100	100	70
EMHVI	LF	–	–	–	–
	HF	100	100	100	70
EMHVIN	LF	–	–	–	–
	HF	100	100	100	70
$VFEIM_h$	LF	77.4	65.2	96.0	12.2
	HF	81.8	84.1	76.7	67.4
VFEMHVI	LF	78.3	59.0	93.8	30.4
	HF	80.9	85.7	77.0	62.8
VFEMHVIN	LF	73.5	69.1	99.0	25.4
	HF	82.1	83.1	75.7	64.1

Table 9 Wilcoxon rank sum test results of IGD metric among three single-fidelity methods

Problem	ZDT1	ZDT2	ZDT3	POL	+/–/=
EIM vs EMHVIN	–	–	=	–	0/3/1
EMHVI vs EMHVIN	=	=	=	=	0/0/4

the mean IGD values of $VFEIM_h$ on ZDT3 problem with $CR = 3$, 4, and 8 are 2.6102e-2, 2.6038e-2, and 2.3487e-2, respectively, indicating the method performs better with the increase of CR.

Table 10 Statistic results of variable-fidelity methods with different cost ratios

Metric	Problem	CR		VFEIM$_h$	VFEMHVI	VFEMHVIN
NHV		3				
			Mean	4.1056e–1	4.2022e–1	4.2087e–1
			STD	9.72e–3	4.70e–3	5.41e–3
	ZDT2	4				
			Mean	4.1127e–1	4.2088e–1	4.1928e–1
			STD	9.40e–3	3.28e–3	3.93e–3
		8				
			Mean	4.1451e–1	4.2043e–1	4.2125e–1
			STD	7.91e–3	3.87e–3	5.76e–3
		3				
			Mean	5.7966e–1	5.7446e–1	5.7616e–1
			STD	1.04e–2	1.85e–3	1.05e–2
	ZDT3	4				
			Mean	5.8321e–1	5.7471e–1	5.7546e–1
			STD	1.15e–2	1.98e–3	2.43e–3
		8				
			Mean	5.7353e–1	5.7506e–1	5.7552e–1
			STD	4.92e–3	3.39e–3	2.61e–3
		3				
			Mean	3.5346e–2	2.8590e–2	2.6191e–2
			STD	1.11e–2	4.68e–3	4.79e–3
	ZDT2	4				
			Mean	3.5003e–2	2.8651e–2	2.9102e–2
			STD	7.48e–3	2.89e–3	4.05e–3
		8				
IGD			Mean	3.6385e–2	2.9090e–2	2.8402e–2
			STD	1.36e–2	3.23e–3	4.74e–3
		3				
			Mean	2.6102e–2	2.7918e–2	3.1732e–2
			STD	6.25e–3	3.51e–3	8.51e–3
	ZDT3	4				
			Mean	2.6038e–2	2.8579e–2	2.6137e–2
			STD	3.95e–3	5.41e–3	4.08e–3
		8				
			Mean	2.3487e–2	2.6629e–2	2.4100e–2
			STD	3.44e–3	7.82e–3	7.19e–3

5 Engineering optimization design cases

5.1 NACA0012 airfoil optimization in transonic inviscid flow

The first engineering design case is the drag minimization and lift maximization of the NACA0012 airfoil. The 2-objective minimization problem is formulated as:

$$\min_{\mathbf{x} \in \Omega} F(\mathbf{x}) = \left\{ C_D, -C_L \right\}, \tag{38}$$

where C_D, C_L is the drag and lift coefficient respectively at the angle of attack 1.25 and free-stream Mach number 0.8. The HF and LF analysis are defined as CFD simulations solving the Euler equations with different convergence criteria. The transonic inviscid flow simulation is conducted by the open-source CFD simulator SU2 [62]. The grid corresponding to the simulation for the baseline airfoil is presented in Fig. 10. The LF CFD simulation terminates after 120 iteration and the HF simulation stops until the density residual is lower than 10^{-8}. The histories of residual, drag, and lift coefficients for the baseline airfoil is shown in Fig. 11. The lift and drag coefficients for the baseline NACA0012 airfoil from the inviscid simulation are 0.3269 and 0.0214, respectively. It should be mentioned that this problem is not a much expensive problem to allow multiple runs to obtain statistical performance of each method on this problem. From Fig. 11, it can also be observed that the computational cost ratio of the HF simulation to the LF simulation for the baseline airfoil is approximately 2.6. However, it is found the number of iteration for an HF simulation to reach the specified residual is dependent on the spatial position of a sample point in the optimization process, resulting in cost ratio varying between 2 and 60 over the design space. Varying cost for expensive engineering design problems has been observed before and a specified strategy to treat the varying cost in design space exploration for a single-fidelity method is proposed in [52]. While the inclusion of a scheme to deal with varying cost over the design space is out of the scope of this work and will be conducted in the future. To this end, the cost ratio is specified to be 3 for the NACA0012 airfoil optimization problem.

Hicks–Henne bump functions [63] with a total of 18 design variables are used in the geometry parameterization of the NACA0012 airfoil. The new airfoil is obtained by adding disturbance (the summation of bump functions) on the lower and upper surfaces of the baseline shape. The lower and upper profile disturbance Δ_l and Δ_u are expressed as:

$$\begin{cases} \Delta_l = -\sum_{i=1}^{9} \delta_{li} f_i(x), \ \delta_{li} \in [-0.01, 0.01] \\ \Delta_u = \sum_{i=1}^{9} \delta_{ui} f_i(x), \ \delta_{ui} \in [-0.01, 0.01] \end{cases},$$
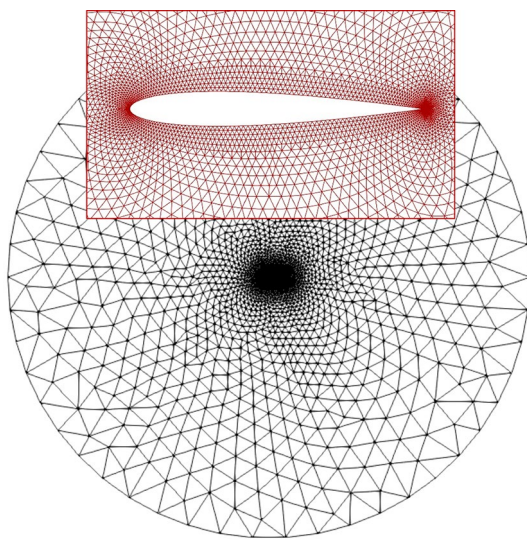
$$f_i(x) = \sin^3 \left(\pi x^{e(i)} \right), \quad e(i) = \log(0.5)/\log \left(x_i \right), \ i = 1, \ldots, 9$$
$$x = [0.05, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.95] \tag{39}$$

The amplitudes of bumps δ_{li} and δ_{ui}, $(i = 1, \cdots, 9)$ are the 18 design variables of the optimization. The initial sample size for single-fidelity methods is set as 5d. For variable-fidelity methods, the number of initial HF and LF samples
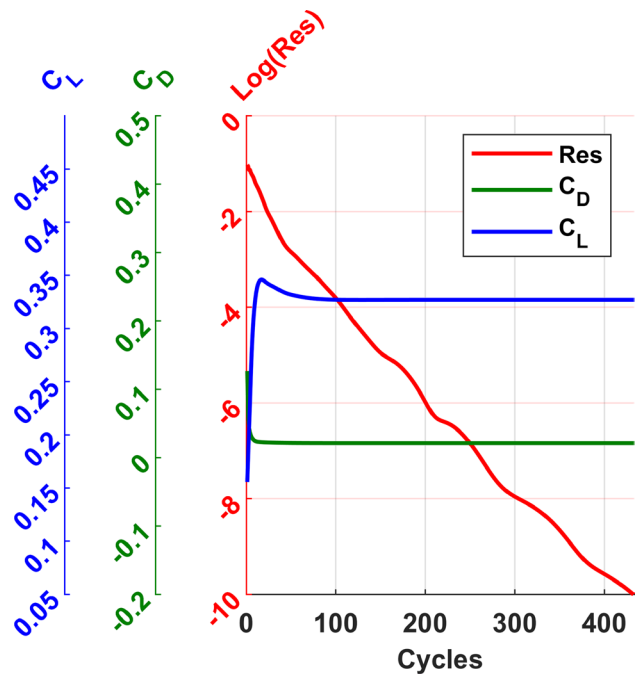
Table 11 Statistic results of compared methods on NACA0012 airfoil optimization

	NHV	No. of final HF solutions for PF		No. of samples
EMHVI				
Mean	8.6963e−1	1.5545e+1	LF	–
STD	3.39e−2 (−)	3.83e+0	HF	180
VFEIM$_h$				
Mean	8.7642e−1	1.5000e+1	LF	175.5
STD	5.14e−2 (=)	4.20e+0	HF	122.2
VFEMHVI				
Mean	8.5421e−1	1.4545e+1	LF	172.2
STD	3.83e−2 (−)	2.16e+0	HF	123.2
VFEMHVIN				
Mean	9.0534e−1	1.2727e+1	LF	170.8
STD	2.33e−2	2.24e+0	HF	123.7

Fig. 11 Convergence histories for the baseline NACA0012 airfoil

Fig. 10 Sketch of grid for CFD simulation for baseline NACA0012 airfoil


Fig. 12 PF approximations with median NHV value by each method on NACA0012 airfoil optimization problem

is set as 2d and 3d × CR, respectively. All methods stop once the 90 allowed equivalent iteration is used out. Only NHV is adopted in the performance assessment, as no true PF is available for the calculation of IGD of the practical problem. The reference point used for NHV estimation is [0.1, 0]. To demonstrate the effectiveness of the variable-fidelity methods as well as release the computational burden of the experiments, the EMHVI method is chosen among the three single-fidelity methods to solve the practical engineering design problem. The EMHVI is chosen as it performs better than EIM$_h$ and comparative to EMHVIN in the experiments over the analytic problems.

**Fig. 13** Comparison of baseline NACA0012 and optimal airfoils

Fig. 14 Mach number contour comparison of baseline NACA0012 and optimal airfoils
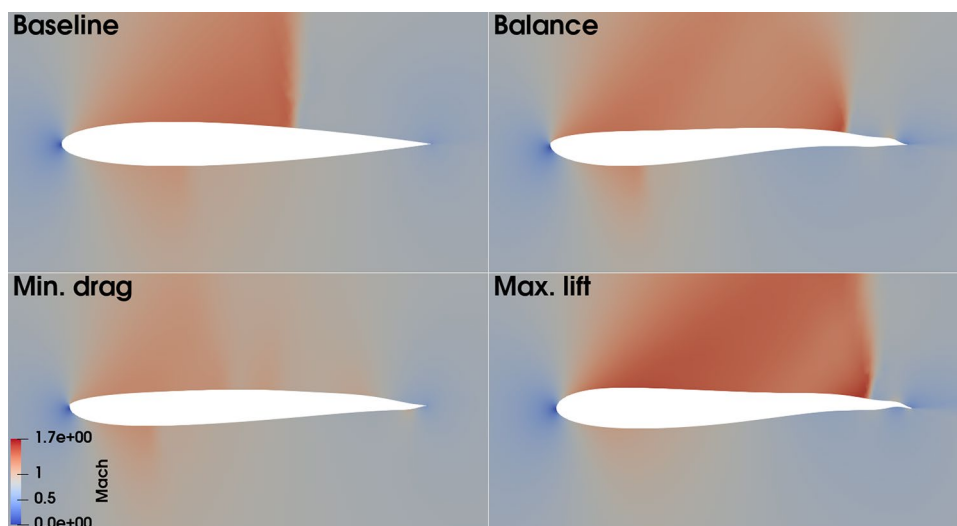
The NACA0012 airfoil optimization problem is solved by each method with 11 times and Table 11 presents the statistic results of the compared methods over this problem. As observed, the VFEMHVIN performs statistically better than EMHVI and VFEMHVI, and is comparable to the VFEIM$_h$ method. The slightly worse performance of VFEMHVI compared with the VFEMHVIN indicates that an inappropriate reference point for the hypervolume calculation in the VFEMHVI criterion will deteriorate its performance and the normalization-based strategy to define the reference point is still applicable in practical problem. From the number of LF/HF samples for each method shown in Table 11, the ability to adaptively determining LF/HF sample in the iterative optimization phase is well demonstrated in the practical engineering design problem. PF approximations with median NHV value among the 11 runs by each method are shown in Fig. 12. Recall the objective values of the baseline airfoil (0.0214, −0.3269), it can be noted that all methods can obtain improved solutions. Representative optimal solutions are selected to show the details. The objective values for the three optimal solutions (denoted as the Min. drag, Balance, Max. lift) are (0.001612, −0.2909), (0.0163, -0.7506), (0.05986, −0.9263), respectively. Figures 13 and 14 compare the shapes and the Mach number contours of the baseline NACA0012 and optimal airfoils. It can be seen the optimal airfoils are all asymmetrical and generally thinner especially at the backward parts.

5.2 RAE2822 airfoil optimization in viscous turbulent flow

The second engineering design case is the drag minimization and lift maximization of the RAE2822 airfoil. The 2-objective minimization problem is formulated as:
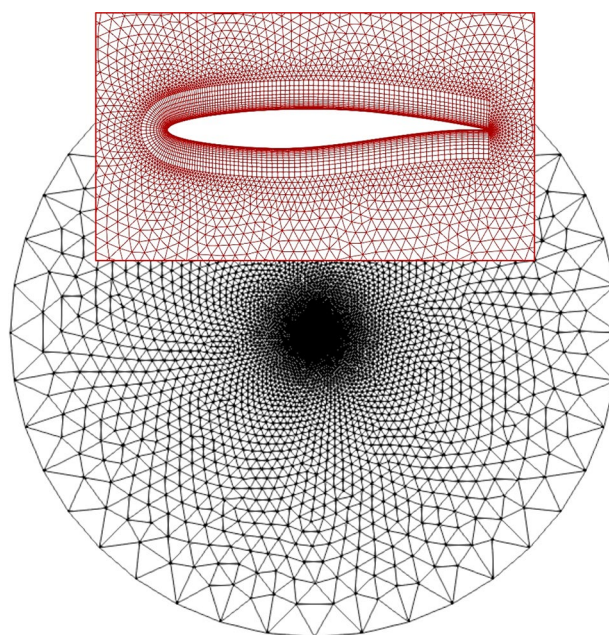
Fig. 15 Sketch of grid for CFD simulation for baseline RAE2822 airfoil

$$\min_{\mathbf{x} \in \Omega} F(\mathbf{x}) = \left\{ C_D, -C_L \right\}, \tag{40}$$

where C_D, C_L are the drag and lift coefficient respectively at the angle of attack 2.9 and free-stream Mach number 0.734. The HF and LF analysis are defined as CFD simulations solving the Reynolds-averaged Naiver–Stokes equations with different convergence criteria in SU2. The grid corresponding to the simulation for the baseline airfoil is presented in Fig. 15. The LF CFD simulation terminates after 800 iterations and the HF simulation stops until the density
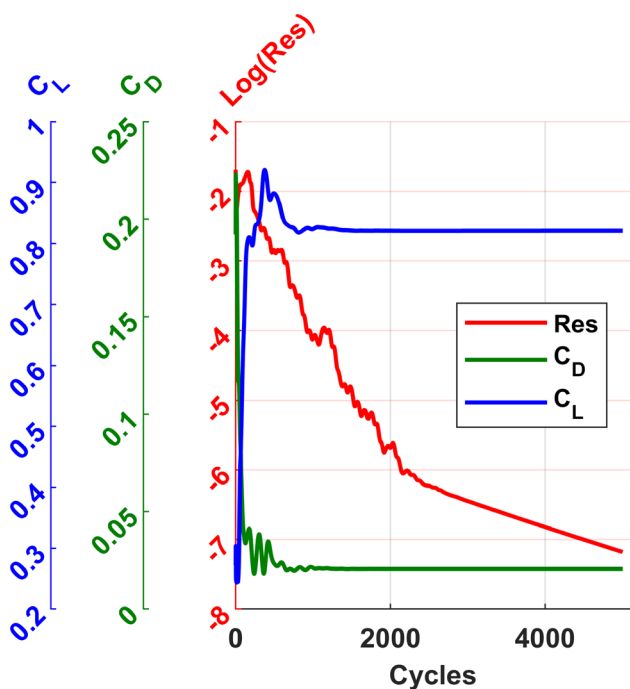
Fig. 16 Convergence histories for the baseline RAE2822 airfoil

Fig. 17 PF approximations with median NHV value by each method on RAE2822 airfoil optimization problem

Fig. 18 Comparison of baseline RAE2822 and optimal airfoils

Table 12 Statistic results of compared methods on RAE2822 airfoil optimization

	NHV	No. of final HF solutions for PF		No. of samples
EMHVI				
Mean	7.0741e−1	1.3091e+1	LF	–
STD	2.62e−2 (−)	2.39e+0	HF	180
VFEIM$_h$				
Mean	7.3248e−1	1.3455e+1	LF	353.0
STD	1.80e−2 (=)	2.34e+0	HF	121.5
VFEMHVI				
Mean	6.8278e−1	1.2636e+1	LF	363.2
STD	1.67e−2 (−)	2.66e+0	HF	119.8
VFEMHVIN				
Mean	7.2659e−1	1.0364e+1	LF	365.2
STD	1.42e−2	2.06e+0	HF	119.5

residual is lower than 10^{-7}, resulting in the CR $\approx$ 6. The histories of residual, drag, and lift coefficients for the baseline airfoil is shown in Fig. 16. The lift and drag coefficients for the baseline RAE2822 airfoil from the viscous turbulent flow simulation are 0.8210 and 0.0206, respectively. Airfoil parameterization strategy in the NACA0012 case is employed. The amplitudes of bumps δ_{li} and $\delta_{ui}, (i = 1, \ldots, 9)$ are the 18 design variables of the optimization but with a reduced range for each variable as [− 0.0001, 0.0001] to ensure the possibility of encountering simulation failures
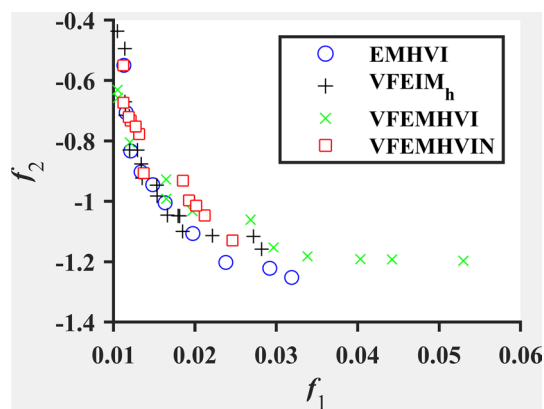
as low as possible. It should be mentioned that such treatment will limit the optimization gain, but it is necessary to conduct the comparative study smoothly. Strategy to deal with the simulation failures in our previous work [64] can be incorporated to potentially improve the robustness and efficiency of the proposed methods in solving practical engineering problems in presence of simulation failures and is left for future work. The settings of the initial sample size and termination criterion hold identical to that in the NACA0012 case. The reference point used for NHV estimation is [0.05, − 0.4] (See Fig. 16).

The RAE airfoil optimization problem is solved by each method with 11 times and Table 11 presents the statistic results of the compared methods over this problem. As observed, the VFEMHVIN performs statistically better than EMHVI and VFEMHVI, and is comparable to the VFEIM$_h$ method. Again, the slightly worse performance of VFEMHVI compared with the VFEMHVIN indicates that an inappropriate reference point for the hypervolume calculation in the VFEMHVI criterion will deteriorate its performance and the normalization-based strategy to define the reference point is still applicable in this more expensive practical problem. The ability to adaptively choosing LF/HF
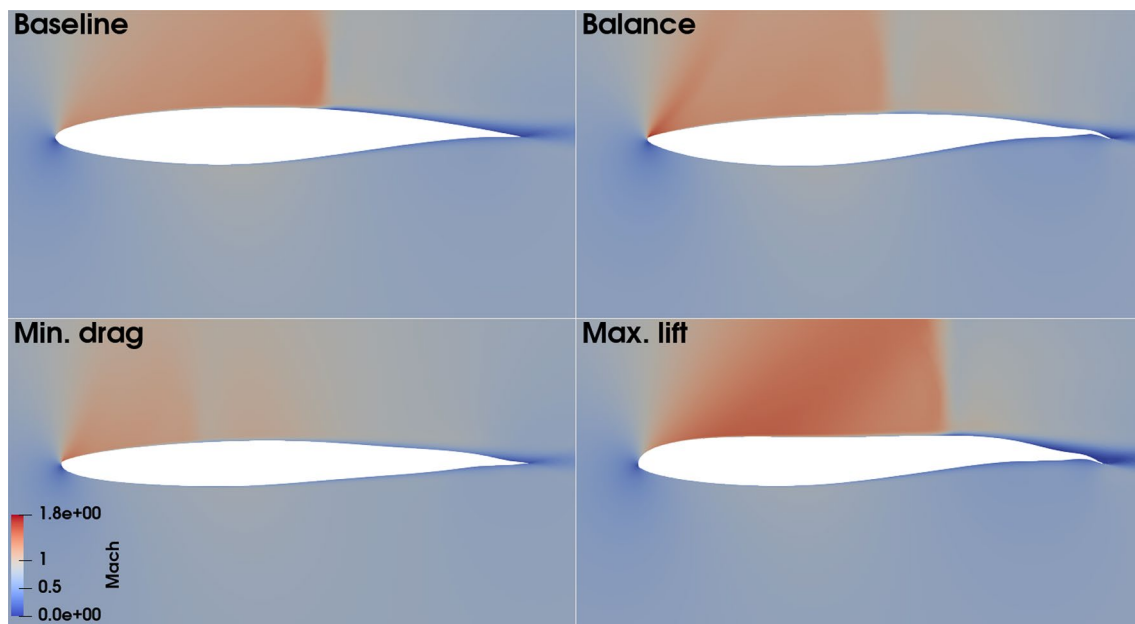
Fig. 19 Mach number contour comparison of baseline RAE2822 and optimal airfoils

sample in the iterative optimization phase is demonstrated from the number of LF/HF samples for each method shown in Table 12. PF approximations with median NHV value among the 11 runs by each method are shown in Fig. 17. It can be noted that all methods can improve the baseline design in at least one of the objectives. Representative optimal solutions are selected to show the details. The objective values for the three optimal solutions (denoted as the Min. drag, Balance, Max. lift) are $(0.0105, -0.6315)$, $(0.01649, -0.9918)$, $(0.05298, -1.197)$, respectively. Figures 18 and 19 compare the shapes and the Mach number contours of the baseline RAE2822 and optimal airfoils. It can be seen the optimal airfoils are all asymmetrical and generally thinner especially at the backward parts. The shock intensity is decreased significantly for the Min. drag optimal solution and increased sharply for the Max. lift design.

6 Discussion

This section presents a discussion on the merits, limitations, and avenues of improvements of the proposed approaches. The discussion is based on the introduction in Sect. 3 and above experimental results but also extends to broader comments.

Comparison between the two criteria Both criteria are formulated explicitly and cheap-to-evaluate. The VFEIM$_h$ criterion is obtained heuristically via replacing the improvement in the simplified hypervolume improvement function by the corresponding expected improvement. The idea is simple and straightforward. While the VFEMHVI is derived analytically based on the modified hypervolume improvement. The modified hypervolume improvement shares the identical properties with the original hypervolume improvement function in MOO but the properties of the simplified hypervolume improvement function behind VFEIM$_h$ is not clear. From the test results, it can be concluded that the VFEMHVI incorporating the normalization-based strategy to define the reference point is the more efficient one. Though the VFEIM$_h$ performs slightly worse in obtaining solutions with high quality, it still has an advantage of being able to obtain more solutions to approximate the PF.

Limitations As presented in Sect. 3, the current criteria are developed for application with two levels of fidelity. We believe that extending the two-level HK model to multi-fidelity one and developing dedicated criteria can be a remedy to tackle this problem.

Avenues for improvement Beside the development of criterion for multi-fidelity applications, several other avenues for improvements can be envisioned. For problems with varying cost over the design space, incorporation of a cost-control strategy might be helpful to potentially enhance the efficiency of the proposed method in such application. The simulation failures in practical engineer problems, especially for CFD-based applications, is critical to the proposed sequential method since it will lead to the premature halt of the iterative optimization process. Integration of an efficient strategy to attack the simulation failures might improve the robustness and optimization efficiency of the approaches. Parallelizing the proposed criteria can be an alternative to deal with simulation failures and accelerate the optimization process with the aid of parallel computation.

7 Summary and future works

In this paper, two variable-fidelity hypervolume-based expected improvement criteria for multi-objective efficient global optimization of expensive-to-evaluate functions with the assistance of cheaper-to-evaluate functions were proposed. Both criteria can be regarded as a multi-objective extension of the variable-fidelity expected improvement for the single-objective variable-fidelity EGO method. The first criterion, VFEIM$_h$, was obtained by aggregating the variable-fidelity expected improvement matrix using a simplified hypervolume-based aggregation scheme. The second criterion, VFEMHVI, was derived analytically based on a modified hypervolume definition. It was found the second criterion can also be obtained by using the variable-fidelity expected improvement matrix via replacing the reference values and utilizing the other aggregation scheme. Both criteria can adaptively select new LF/HF samples in the iterative optimal search process to update the variable-fidelity models towards the HF Pareto front, distinguishing the proposed methods to the rests in the open literature. The constrained versions of the two criteria were also derived. As both criteria are based on the hypervolume measure, a reference point for hypervolume calculation is necessary. Except for specifying it by the user/designer, a strategy based on normalization was also proposed and implemented. The effectiveness and efficiency of the proposed methods were then verified and validated over analytic multi-objective problems and demonstrated for aerodynamic shape optimizations of the NACA0012 airfoil in transonic inviscid flow and RAE2822 in a viscous turbulent flow. It was shown that the VFEMHVI combined with the normalization-based strategy to define the reference point is the most efficient method over the compared single-fidelity and variable-fidelity ones. Except for applying the proposed method to more complex test problems and real-world engineering problems from the application view, the performance of the general criterion (25), which allows the objectives of the newly selected sample to be evaluated as different fidelities, will be explored.

Appendix

MATLAB code of VFEIM

---

function y = **Infill_VFEIM**(x, HK, non_dominated_front, fidlty)

% VFEIM based criterion utilizing normalization procedure to specify the reference point

% x: value of design variable

% HK: Hierarchichal models for each objective

MATLAB code of VFEMHVI

---

function y = **Infill_VFMEHVI**(x, HK, non_dominated_front, fidlty, ReferencePoint)

% VFEMHVI based criterion with user inputting the reference point

% x: value of design variable

% HK: Hierarchichal models for each objective

% non_dominated_front: data of current HF approximation of PF

% fidlty: fidelity level, i.e. 1 or 2

% ReferencePoint: reference point

%--

% the input parameters

P_NDS = non_dominated_front;

% Local Upper Bounds (LUBs)

f = GenerateLUBs(P_NDS,ReferencePoint);

% number of LUB points and number of objectives

[p,M] = size(f);

% number of input designs

num_x = size(x,1);

%--

% the HK predictions and varinces

u = zeros(num_x,M);

mse = zeros(num_x,M);

for ii = 1:M

    [~, ~, y_MLHK, pred_mse] = pred_MLHK(x, HK{ii});

    if fidlty==1

        mse(:,ii) = pred_mse(:,fidlty)*(HK{ii}.belta0)^2;

    else

        mse(:,ii) = pred_mse(:,fidlty);

    end

    u(:,ii) = y_MLHK(:,fidlty);

end

s=sqrt(max(0,mse));

% the VFEI matrix (three dimensional matrix)

f_matrix   =   f .* ones(1,1,num_x);

u_matrix = reshape(u', 1, M, num_x).* ones(p,1);

s_matrix =   reshape(s', 1, M, num_x).* ones(p,1);

EI_matrix=(f_matrix-u_matrix).*Gaussian_CDF((f_matrix-u_matrix)./s_matrix)+s_matrix.*Gaussian_PDF((f_matrix-u_matrix)./s_matrix);

%   the modified VFEHVI

EIM = reshape(sum(prod(EI_matrix,2), 1), num_x, 1, 1);

% the objective is to be maximized

y = -EIM;

end

Gaussian_CDF and Gaussian_PDF are the Gaussian cumulative distribution function and probability density function respectively, and are defined as follows.

```
function y=Gaussian_CDF(x)
    y=0.5*(1+erf(x/sqrt(2)));
end
```

```
function y=Gaussian_PDF(x)
    y=1/sqrt(2*pi)*exp(-x.^2/2);
end
```

Availability of data and material The data of the numerical experiments will be provided after the paper being published.

Code availability The Appendix provides the MATLAB codes to calculate the proposed criteria. The executable code of the proposed method will be available online after the paper being published.

Declarations

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

References

1. Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. Stat Sci 4:409–423. https://doi.org/10.1214/ss/1177012413
2. Rasmussen CE (2004) Gaussian processes in machine learning. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). https://doi.org/10.1007/978-3-540-28650-9_4
3. Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Glob Optim 13:455–492. https://doi.org/10.1023/A:1008306431147
4. Song P, Sun J, Wang K (2014) Axial flow compressor blade optimization through flexible shape tuning by means of cooperative co-evolution algorithm and adaptive surrogate model. Proc Inst Mech Eng Part A J Power Energy 228:782–798. https://doi.org/10.1177/0957650914541647
5. Venturelli G, Benini E (2016) Kriging-assisted design optimization of S-shape supersonic compressor cascades. Aerosp Sci Technol 58:275–297. https://doi.org/10.1016/j.ast.2016.08.021
6. Song P, Sun J, Wang K, He Z (2011) Development of an optimization design method for turbomachinery by incorporating the Cooperative Coevolution Genetic Algorithm and adaptive approximate model. In: Proceedings of the ASME Turbo Expo. pp 1139–1153 https://doi.org/10.1115/GT2011-45411
7. Song L, Guo Z, Li J, Feng Z (2016) Research on metamodel-based global design optimization and data mining methods. J Eng Gas Turbines Power 138:92604–92614. https://doi.org/10.1115/1.4032653
8. Forrester AIJ, Keane AJ (2009) Recent advances in surrogate-based optimization. Prog Aerosp Sci 45:50–79. https://doi.org/10.1016/j.paerosci.2008.11.001
9. Bhosekar A, Ierapetritou M (2018) Advances in surrogate-based modeling, feasibility analysis, and optimization: a review. Comput Chem Eng 108:250–267. https://doi.org/10.1016/j.compchemeng.2017.09.017
10. Park C, Haftka RT, Kim NH (2017) Remarks on multi-fidelity surrogates. Struct Multidiscip Optim 55:1029–1050. https://doi.org/10.1007/s00158-016-1550-y
11. Giselle Fernández-Godino M, Park C, Kim NH, Haftka RT (2019) Issues in deciding whether to use multifidelity surrogates. AIAA J. https://doi.org/10.2514/1.J057750
12. Knowles J (2006) ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Trans Evol Comput 10:50–66. https://doi.org/10.1109/TEVC.2005.851274
13. He Y, Sun J, Song P et al (2020) Preference-driven Kriging-based multiobjective optimization method with a novel multipoint infill criterion and application to airfoil shape design. Aerosp Sci Technol. https://doi.org/10.1016/j.ast.2019.105555
14. Chen S, Jiang Z, Yang S, Chen W (2017) Multimodel fusion-based sequential optimization. AIAA J 55:241–254. https://doi.org/10.2514/1.J054729
15. Kennedy MC, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. Biometrika. https://doi.org/10.1093/biomet/87.1.1
16. Forrester AIJ, Sóbester A, Keane AJ (2007) Multi-fidelity optimization via surrogate modelling. Proc R Soc A Math Phys Eng Sci. https://doi.org/10.1098/rspa.2007.1900
17. Huang D, Allen TT, Notz WI, Miller RA (2006) Sequential kriging optimization using multiple-fidelity evaluations. Struct Multidiscip Optim 32:369–382. https://doi.org/10.1007/s00158-005-0587-0
18. Xiong Y, Chen W, Tsui KL (2008) A new variable-fidelity optimization framework based on model fusion and objective-oriented sequential sampling. J Mech Des Trans ASME. https://doi.org/10.1115/1.2976449
19. Zhang Y, Han Z-H, Zhang K-S (2018) Variable-fidelity expected improvement method for efficient global optimization of expensive functions. Struct Multidiscip Optim 58:1431–1451. https://doi.org/10.1007/s00158-018-1971-x
20. Liu Y, Chen S, Wang F, Xiong F (2018) Sequential optimization using multi-level cokriging and extended expected improvement criterion. Struct Multidiscip Optim 58:1155–1173. https://doi.org/10.1007/s00158-018-1959-6
21. Jiang P, Cheng J, Zhou Q et al (2019) Variable-fidelity lower confidence bounding approach for engineering optimization problems with expensive simulations. AIAA J. https://doi.org/10.2514/1.j058283
22. Hao P, Feng S, Li Y et al (2020) Adaptive infill sampling criterion for multi-fidelity gradient-enhanced kriging model. Struct Multidiscip Optim. https://doi.org/10.1007/s00158-020-02493-8
23. Forrester A, Sobester A, Keane A (2008) Engineering design via surrogate modelling: a practical guide. Wiley, Hoboken. https://doi.org/10.2514/4.479557

24. Emmerich MTM, Deutz AH, Klinkenberg JW (2011) Hypervolume-based expected improvement: Monotonicity properties and exact computation. In: 2011 IEEE Congress of evolutionary computation, CEC 2011. pp 2147–2154 /https://doi.org/10.1109/CEC.2011.5949880

25. Couckuyt I, Deschrijver D, Dhaene T (2014) Fast calculation of multiobjective probability of improvement and expected improvement criteria for Pareto optimization. J Glob Optim 60:575–594. https://doi.org/10.1007/s10898-013-0118-2

26. Bautista DC (2009) A Sequential design for approximating the pareto front using the expected pareto improvement function. Ph.D. dissertation, Ohio State University, Ohio, USA

27. Svenson J, Santner T (2016) Multiobjective optimization of expensive-to-evaluate deterministic computer simulator models. Comput Stat Data Anal 94:250–264. https://doi.org/10.1016/j.csda.2015.08.011

28. Namura N, Shimoyama K, Obayashi S (2017) Expected improvement of penalty-based boundary intersection for expensive multiobjective optimization. IEEE Trans Evol Comput 21:898–913. https://doi.org/10.1109/TEVC.2017.2693320

29. Zhan D, Cheng Y, Liu J (2017) Expected improvement matrix-based infill criteria for expensive multiobjective optimization. IEEE Trans Evol Comput 21:956–975. https://doi.org/10.1109/TEVC.2017.2697503

30. van der Herten J, Knudde N, Couckuyt I, Dhaene T (2020) Multiobjective Bayesian optimization for engineering simulation. In: Bartz-Beielstein T, Filipič B, Korošec P, Talbi E-G (eds) High-performance simulation-based optimization. Springer International Publishing, Cham, pp 47–68. https://doi.org/10.1007/978-3-030-18764-4_3

31. Zuhal LR, Palar PS, Shimoyama K (2019) A comparative study of multi-objective expected improvement for aerodynamic design. Aerosp Sci Technol 91:548–560. https://doi.org/10.1016/j.ast.2019.05.044

32. Li Z, Wang X, Ruan S et al (2018) A modified hypervolume-based expected improvement for multi-objective efficient global optimization method. Struct Multidiscip Optim 58:1961–1979. https://doi.org/10.1007/s00158-018-2006-3

33. Cheng S, Zhan H, Shu Z et al (2019) Effective optimization on Bump inlet using meta-model multi-objective particle swarm assisted by expected hyper-volume improvement. Aerosp Sci Technol 87:431–447. https://doi.org/10.1016/j.ast.2019.02.039

34. Koziel S, Bekasiewicz A, Couckuyt I, Dhaene T (2014) Efficient multi-objective simulation-driven antenna design using co-kriging. IEEE Trans Antennas Propag 62:5900–5905. https://doi.org/10.1109/TAP.2014.2354673

35. Koziel S, Bekasiewicz A (2013) Multi-objective design of antennas using Variable-Fidelity Simulations and Surrogate surrogate models. IEEE Trans Antennas Propag 61:5931–5939. https://doi.org/10.1142/q0043

36. Bekasiewicz A, Koziel A (2014) Rapid multi-objective optimization of a MIMO antenna for UWB applications. Loughbrgh Antennas Propag Conf LAPC 2014:500–503. https://doi.org/10.1109/LAPC.2014.6996434

37. Leifsson L, Koziel S, Tesfahuneng YA, Hosder S (2015) Multi-objective aeroacoustic shape optimization by variable-fidelity models and response surface surrogates. In: 56th AIAA/ASCE/AHS/ASC Struct Struct Dyn Mater Conf, pp 1–12. https://doi.org/10.2514/6.2015-1800

38. Koziel S, Leifsson L (2014) Multi-objective airfoil design using variable-fidelity CFD simulations and response surface surrogates. In: 10th AIAA Multidiscip Des Optim Spec Conf, pp 1–9. https://doi.org/10.2514/6.2014-0289

39. Liu Y, Collette M (2014) Improving surrogate-assisted variable fidelity multi-objective optimization using a clustering algorithm.

40. Zhu J, Wang YJ, Collette M (2014) A multi-objective variable-fidelity optimization method for genetic algorithms. Eng Optim 46:521–542. https://doi.org/10.1080/0305215X.2013.786063

41. Shu L, Jiang P, Zhou Q et al (2018) An online variable fidelity metamodel-assisted multi-objective genetic algorithm for engineering design optimization. Appl Soft Comput J 66:438–448. https://doi.org/10.1016/j.asoc.2018.02.033

42. Shu L, Jiang P, Zhou Q, Xie T (2019) An online variable-fidelity optimization approach for multi-objective design optimization. Struct Multidiscip Optim. https://doi.org/10.1007/s00158-019-02256-0

43. Zhou Q, Wu J, Xue T, Jin P (2019) A two-stage adaptive multi-fidelity surrogate model-assisted multi-objective genetic algorithm for computationally expensive problems. Eng Comput. https://doi.org/10.1007/s00366-019-00844-8

44. Jiang P, Zhou Q, Liu J, Cheng Y (2019) A three-stage surrogate model-assisted multi-objective genetic algorithm for computationally expensive problems. In: 2019 IEEE Congress on evolutionary computation, CEC 2019–Proceedings. pp 1680–1687 https://doi.org/10.1109/CEC.2019.8790241

45. Yi J, Gao L, Li X et al (2019) An on-line variable-fidelity surrogate-assisted harmony search algorithm with multi-level screening strategy for expensive engineering design optimization. Knowl-Based Syst 170:1–19. https://doi.org/10.1016/j.knosys.2019.01.004

46. Habib A, Singh KH, Ray HT (2019) A multiple surrogate-assisted multi/many-objective multi-fidelity evolutionary algorithm. Inf Sci (Ny) 502:537–557. https://doi.org/10.1016/j.ins.2019.06.016

47. Belakaria S, Deshwal A, Doppa JR (2020) Multi-fidelity multi-objective bayesian optimization: an output space entropy search approach. In: Proceedings of the AAAI Conference on artificial intelligence, pp 10035–10043 https://doi.org/10.1609/aaai.v34i06.6560

48. Han ZH, Görtz S (2012) Hierarchical kriging model for variable-fidelity surrogate modeling. AIAA J 50:1885–1896. https://doi.org/10.2514/1.J051354

49. Xu J, Han Z, Song W, Li K (2020) Efficient aerodynamic optimization of propeller using hierarchical kriging models. J Phys Conf Ser 1519:12019. https://doi.org/10.1088/1742-6596/1519/1/012019

50. Bu Y, Song W, Han Z et al (2020) Aerodynamic/aeroacoustic variable-fidelity optimization of helicopter rotor based on hierarchical Kriging model. Chin J Aeronaut 33:476–492. https://doi.org/10.1016/j.cja.2019.09.019

51. Han Z, Xu C, Zhang L et al (2020) Efficient aerodynamic shape optimization using variable-fidelity surrogate models and multi-level computational grids. Chin J Aeronaut 33:31–47. https://doi.org/10.1016/j.cja.2019.05.001

52. Zhang Y, Neelakantan A, Park C et al (2019) Adaptive sampling with varying sampling cost for design space exploration. AIAA J. https://doi.org/10.2514/1.J057470

53. Klamroth K, Lacour R, Vanderpooten D (2015) On the representation of the search region in multi-objective optimization. Eur J Oper Res 245:767–778. https://doi.org/10.1016/j.ejor.2015.03.031

54. Tian Y, Cheng R, Zhang X, Jin Y (2017) PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [Educational Forum]. IEEE Comput Intell Mag. https://doi.org/10.1109/MCI.2017.2742868

55. Zitzler E, Thiele L (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans Evol Comput 3:257–271. https://doi.org/10.1109/4235.797969

56. Coello CAC, Cortés NC (2005) Solving multiobjective optimization problems using an artificial immune system. Genet

Appl Soft Comput J 24:482–493. https://doi.org/10.1016/j.asoc.2014.07.022

Program Evol Mach 6:163–190. https://doi.org/10.1007/s10710-005-6164-x

57. Zhang Q, Liu W, Tsang E, Virginas B (2010) Expensive multiobjective optimization by MOEA/D with Gaussian process model. IEEE Trans Evol Comput 14:456–474. https://doi.org/10.1109/TEVC.2009.2033671

58. Han Z, Liu F, Xu C, et al (2019) Efficient multi-objective evolutionary algorithm for constrained global optimization of expensive functions. In: 2019 IEEE Congress on Evolutionary Computation, CEC 2019–Proceedings. pp 2026–2033 https://doi.org/10.1109/CEC.2019.8789986

59. Ishibuchi H, Imada R, Setoguchi Y, Nojima Y (2018) How to specify a reference point in hypervolume calculation for fair performance comparison. Evol Comput. https://doi.org/10.1162/EVCO_a_00226

60. Yang K, Emmerich M, Deutz A, Bäck T (2019) Efficient computation of expected hypervolume improvement using box

decomposition algorithms. J Glob Optim. https://doi.org/10.1007/s10898-019-00798-7

61. Yang K, Emmerich M, Deutz A, Bäck T (2018) Multi-objective bayesian global optimization using expected hypervolume improvement gradient. Swarm Evol Comput. https://doi.org/10.1016/j.swevo.2018.10.007

62. Economon TD, Palacios F, Copeland SR et al (2016) SU2: an open-source suite for multiphysics simulation and design. AIAA J 54:828–846. https://doi.org/10.2514/1.J053813

63. Hicks RM, Henne PA (1978) Wing design by numerical optimization. J Aircr doi 10(2514/3):58379

64. He Y, Sun J, Song P, Wang X (2020) Dual Kriging-assisted efficient global optimization of expensive problems with evaluation failures. Aerosp Sci Technol 105:106006. https://doi.org/10.1016/j.ast.2020.106006