# Expected Improvement of Penalty-based Boundary Intersection for Expensive Multiobjective Optimization

Nobuo Namura, *Member*, *IEEE*, Koji Shimoyama, *Member*, *IEEE*,
and Shigeru Obayashi, *Member*, *IEEE*

***Abstract*—Computationally expensive multiobjective optimization problems are difficult to solve using solely evolutionary algorithms (EA) and require surrogate models, such as the Kriging model. To solve such problems efficiently, we propose infill criteria for appropriately selecting multiple additional sample points for updating the Kriging model. These criteria correspond to the expected improvement of the penalty-based boundary intersection (PBI) and the inverted PBI. These PBI-based measures are increasingly applied to EAs due to their ability to explore better nondominated solutions than those that are obtained by the Tchebycheff function. In order to add sample points uniformly in the multiobjective space, we assign territories and niche counts to uniformly distributed weight vectors for evaluating the proposed criteria. We investigate these criteria in various test problems and compare them with established infill criteria for multiobjective surrogate-based optimization. Both of the proposed criteria yield better diversity and convergence than those obtained with other criteria for most of the test problems.**

***Index Terms*—Efficient global optimization (EGO), expected improvement (EI), penalty-based boundary intersection (PBI), expensive optimization, multiobjective optimization.**

## I. INTRODUCTION

IN real-world problems, optimization is typically formulated as a multiobjective problem with tradeoffs between objective functions [1,2]. Evolutionary algorithms (EA) have been developed to solve these multiobjective optimization problems successfully and obtain diverse and converging nondominated solutions (NDS) [3-5]. They have recently been applied to many-objective optimization problems [6-10]. However, real-world problems can be time-consuming and computationally expensive in terms of the objective function evaluation [11,12], which is generally limited to a small number of function evaluations (e.g., 10–1,000 [13,14]) by computational resources. EAs usually operate with many individuals through many generations and the total number of function evaluations amounts to 1,000–1,000,000. Hence, EAs cannot singlehandedly tackle expensive multiobjective optimization problems.

Surrogate models are often useful for alleviating this difficulty. They are constructed for promptly estimating the values of the objective functions at any point from a set of sample points where the real values of the objective functions have been obtained by expensive computations. The optimal solution is explored efficiently, since EAs search on the surrogate model using estimated values instead of real objective function values. However, the optimal solution on the model does not always correspond to the real optimal solution due to the insufficient accuracy of the model with the initial sample points. In this case, the model should be updated by adding sample points in order to improve its accuracy. In the simplest case, we can choose the optimal solution on the model as an additional sample point. We call this infill criterion EST (estimation). However, EST does not always improve accuracy, since approximation errors are not considered.

The Kriging model [15] is often used to approximate complex functions and compute the approximation errors. It can adapt well to complex functions, since it approximates a function by a weighted superposition of basis functions, such as the Gaussian function. The Kriging model yields not only estimated function values but also approximation errors that help users determine the locations of the additional sample points in order to improve accuracy. Mockus et al. derived the expected improvement (EI) value [16] from the estimated function value and the approximation error. EI corresponds to the expected value of the objective function improvement that is obtained by adding a sample point at a certain position. Additional sample points are selected by maximizing the EI value. The EI value enables the Kriging model to explore efficiently the optimum as well as improve the model accuracy in single-objective optimization. Jones et al. [17] called this method efficient global optimization (EGO).

For multiobjective optimization, Knowles proposed ParEGO [18]. This method employs a randomly selected weight vector

This work is supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI 14J07397 through JSPS Research Fellowships for Young Scientists.

N. Namura is with the Institute of Fluid Science, Tohoku University, Sendai, Japan. He is a JSPS Research Fellow (e-mail: namura@edge.ifs.tohoku.ac.jp).

K. Shimoyama and S. Obayashi are with the Institute of Fluid Science, Tohoku University, Sendai, Japan (e-mail: shimoyama@edge.ifs.tohoku.ac.jp; obayashi@ifs.tohoku.ac.jp).

and the Tchebycheff function in order to aggregate objective functions into a scalar function for each iteration. Thus, maximization of the EI value for the scalar function enables us to select the next additional sample point as in single-objective optimization. However, this method is a single-point infill criterion and provides only one candidate point in each iteration. Multiple sample points may be added simultaneously if their objective functions can be evaluated concurrently using parallel computation.

The multiobjective evolutionary algorithm that is based on decomposition (MOEA/D) [19] was connected to the EGO framework by Zhang et al. [20] in order to create MOEA/D-EGO. The problem can then be tackled in MOEA/D using multiple scalar functions with different weight vectors that are uniformly distributed in the objective space. The EI values for each weight vector are maximized and some of them are selected as additional sample points. MOEA/D-EGO achieved comparable results with ParEGO with parallel evaluation. To aggregate objective functions into scalar functions, the authors used the Tchebycheff function as well. MOEA/D-EGO divides sample points into clusters containing a small number of sample points and local Kriging models are constructed for each cluster in order to reduce the computational cost of model construction.

Jeong et al. [21] proposed EGO for multiobjective problems (EGOMOP). This approach uses the NDSs of EI values for each objective function. EA determines numerous NDSs of EI values, and an arbitrary number of solutions are selected as additional sample points. This selection is one of the drawbacks of the method. In [22], the authors apply k-means clustering [22] to divide the solutions and select those that are closest to the centers of each cluster as additional sample points. This approach sometimes fails to cover the entire Pareto front (PF).

Another subset of methods focuses on hypervolume. Ponweiser et al. proposed EGO that is based on S-metric selection (SMS-EGO) [23,24] and uses an estimate for the gain in hypervolume. This gain can be rapidly computed from the estimated function values and the approximation errors in the Kriging models. In order to efficiently lead dominated solutions to a nondominated region, SMS-EGO employs a penalty. The authors compared SMS-EGO with ParEGO and EGOMOP and demonstrated that SMS-EGO attains better performance. Expected hypervolume improvement (EHVI) [25-29] is available for surrogate-based multiobjective optimization as well. The EHVI approach maximizes the expected value of the increase in hypervolume corresponding to an additional sample point. However, SMS-EGO and EHVI cannot add multiple sample points simultaneously, because maximizing EHVI is always a single-objective optimization problem independent of the number of objective functions. Moreover, they entail a high computational cost for evaluating the hypervolume in many-objective problems. This is another drawback of these approaches.

Shimoyama et al. [30] proposed the combination of EHVI with estimated objective function values (this is called EHVI+EST) and explored the NDSs of EHVI and $M$ objective functions in an ($M$+1)-dimensional objective space. EHVI+EST is as capable of adding multiple sample points simultaneously, as in the case of EGOMOP. However, the authors added the solutions closest to the center of ($M$+1)-dimensional NDSs individually and applied EHVI+EST only to two-objective optimization problems. Horn et al. [31] recently proposed a method that enables single-point infill criteria, such as SMS-EGO and EHVI, to add multiple sample points simultaneously. They virtually add a sample point, using the lower confidence bound [32] instead of expensive objective function evaluation, to the location that is suggested by the single-point infill criteria. The infill criteria provide additional sample points until a desirable number of points has been generated.

The penalty-based boundary intersection (PBI) approach [19] has recently become an alternative to the approach for multiobjective EAs that are represented by various MOEA/D-based algorithms [7,10]. PBI approaches have increased in popularity due to their uniformly-distributed NDSs, despite the required tuning of a penalty parameter in the PBI function. The inverted PBI (IPBI) approach, which uses the distance from the nadir point, has been investigated as well [33].

In this study, we propose the expected PBI/IPBI improvement (EPBII/EIPBII) as new infill criteria for efficiently obtaining well-distributed NDSs even though multiple sample points are added simultaneously. The PBI and IPBI approaches enable EPBII and EIPBII to obtain well-distributed NDSs. Fitness functions that are derived from EPBII and EIPBII are used as multipoint infill criteria for efficiently adding multiple sample points by considering the diversity of NDSs. PBI and IPBI improvements can be defined for each weight vector independently if multiple weight vectors are prepared. Thus, we can maximize EPBII and EIPBII for many different weight vectors covering the PF and obtain a large number of candidate points. We introduce territories and niche counts for each weight vector in order to evaluate the fitness functions and select additional sample points from the set of candidates. Moreover, these territories eliminate the need to tune the penalty parameter in PBI and IPBI. We apply the EPBII and EIPBII approaches to various test problems and compare them with other EGO approaches in order to investigate their capabilities.

## II. Efficient Global Optimization

We compare EPBII and EIPBII with EST, EI, EHVI, and EHVI+EST in multiobjective problems. All these infill criteria are implemented in the same EGO framework that is described by the flowchart in Fig. 1. Our approach generates the initial sample points uniformly in the design space using the Latin hypercube sampling (LHS) method [34] and evaluates them on the expensive objective functions. We construct the Kriging models for each objective function by interpolating these sample points. The nondominated sorting genetic algorithm II (NSGA-II) [35] explores NDSs on the Kriging models using estimated objective functions. Some of these NDSs are chosen

as additional sample points as in the EST approach. Moreover, we define a nadir point (for EI, EHVI, EHVI+EST, and EIPBII) or utopia point (for EPBII) in these NDSs as a reference point for computing an infill criterion, as shown in Fig. 2. One of these criteria is maximized by NSGA-II and candidates for additional sample points are identified. We reconstruct the Kriging models with the initial and additional sample points after evaluating the objective functions of the candidates at additional sample points. EGO iterates the procedure that is illustrated in Fig. 1 until a termination condition is satisfied.

In this study, we use NSGA-II not only for multiobjective optimizations but also for single-objective ones (EHVI, EPBII and EIPBII maximizations) in order to simplify our EGO code. However, other single-objective EAs may solve single-objective optimization more efficiently. Fitness in NSGA-II as a single-objective optimizer corresponds to objective function values and crowding distance is never used.

### A. Ordinary Kriging Model

The Kriging model expresses the unknown function $f(\mathbf{x})$ as

$$f(\mathbf{x}) = \mu(\mathbf{x}) + \varepsilon(\mathbf{x}), \tag{1}$$

where $\mathbf{x}$ is an $m$-dimensional vector ($m$ design variables), $\mu(\mathbf{x})$ is a global model, and $\varepsilon(\mathbf{x})$ represents a local deviation from the global model that is defined as a Gaussian process following $\mathcal{N}(0, \sigma^2)$. The correlation between $\varepsilon(\mathbf{x}^i)$ and $\varepsilon(\mathbf{x}^j)$ is strongly related to the distance between the corresponding points, $\mathbf{x}^i$ and $\mathbf{x}^j$. Various functions are available to define the correlation. In this study, we employ the Gaussian function with a weighted distance. The correlation between $\mathbf{x}^i$ and $\mathbf{x}^j$ is defined as

$$Corr\left(\varepsilon(\mathbf{x}^i), \varepsilon(\mathbf{x}^j)\right) = \exp\left(-\sum_{k=1}^{m} \theta_k \left(x_k^i - x_k^j\right)^2\right), \tag{2}$$

where $\theta_k$ $(0 \le \theta_k < \infty)$ is the weight coefficient and the $k$-th element of an $m$-dimensional weight vector $\boldsymbol{\theta}$. These weights provide the Kriging model with anisotropy and enhance its accuracy.

The Kriging predictor and uncertainty, which are proposed by Jones et al. [17], are expressed as

$$\hat{f}(\mathbf{x}) = \hat{\mu}(\mathbf{x}) + \mathbf{r}(\mathbf{x})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{f} - \hat{\boldsymbol{\mu}}), \tag{3}$$

$$s^2(\mathbf{x}) = \hat{\sigma}^2 \left(\mathbf{1} - \mathbf{r}(\mathbf{x})^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{r}(\mathbf{x}) + \frac{(1 - \mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{r}(\mathbf{x}))^2}{\mathbf{1}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{1}}\right), \tag{4}$$

where $\hat{\mu}(\mathbf{x})$ is the estimated value of $\mu(\mathbf{x})$, $\mathbf{R}$ denotes the $n \times n$ matrix whose $(i, j)$ entry is $Corr\left(\varepsilon(\mathbf{x}^i), \varepsilon(\mathbf{x}^j)\right)$, $\mathbf{r}(\mathbf{x})$ is an $n$-dimensional vector whose $i$-th element is $Corr\left(\varepsilon(\mathbf{x}), \varepsilon(\mathbf{x}^i)\right)$, and $\mathbf{f}$ and $\hat{\boldsymbol{\mu}}$ are defined as follows (for $n$ sample points):

$$\mathbf{f} = \left(f(\mathbf{x}^1) \quad \cdots \quad f(\mathbf{x}^n)\right)^{\mathrm{T}}, \tag{5}$$

$$\hat{\boldsymbol{\mu}} = \left(\hat{\mu}(\mathbf{x}^1) \quad \cdots \quad \hat{\mu}(\mathbf{x}^n)\right)^{\mathrm{T}}. \tag{6}$$

The unknown parameters in the Kriging model are $\boldsymbol{\theta}$, $\hat{\mu}(\mathbf{x})$, and $\hat{\sigma}^2$ (estimated $\sigma^2$). They are obtained by maximizing the likelihood function. Details may be found in [21]

### B. Expected Improvement

The estimated values in the Kriging model are not deterministic, but rather follow the Gaussian distribution. This is denoted by $\mathcal{N}(\hat{f}(\mathbf{x}), s^2(\mathbf{x}))$. From this, we can calculate the probability that the solution at point $\mathbf{x}$ will achieve a new global optimum. The EI value, which corresponds to the expected value of the objective function improvement from the current optimal solution among the sample points, may be derived using this probability. In the single-objective minimization problem of $f(\mathbf{x})$, the improvement value $I(\mathbf{x})$ and the EI value $E(I(\mathbf{x}))$ of $f(\mathbf{x})$ are expressed as

$$I(\mathbf{x}) = \max\left(f_{ref} - f, 0\right), \tag{7}$$

$$E(I(\mathbf{x})) = \int_{-\infty}^{f_{ref}} (f_{ref} - f)\varphi(f) df, \tag{8}$$

respectively, where $\varphi$ is the probability density function that is denoted by $\mathcal{N}(\hat{f}(\mathbf{x}), s^2(\mathbf{x}))$ and represents uncertainty about $f$. $f_{ref}$ is the reference value of $f$ and corresponds to the minimum value of $f(\mathbf{x})$ among the existing sample points. The integral in Eq. (8) may be computed analytically using the error function.

## III. Expected PBI and IPBI Improvement

EPBII and EIPBII are derived by nearly the same procedure. Hence, we first introduce EPBII. Subsequently, we extend it to EIPBII. EPBII is the expected value of the PBI improvement that is derived from the multidimensional Gaussian distribution of the Kriging models. Fig. 3 shows the flowchart of our entire multiobjective EGO framework using EPBII. Fig. 4 shows the definition of EPBII. EPBII should be maximized for each weight vector in Fig 4. EPBII becomes a single-point infill criterion if we maximize EPBII for a weight vector that is randomly chosen from a set of vectors and use it as an additional sample point as in ParEGO. In order to add multiple sample points simultaneously, a multipoint infill criterion that is derived from EPBII must be employed.

We first introduce EPBII as a single-point infill criterion. Subsequently, we extend EPBII to a multipoint infill criterion using the procedure in Fig. 3. Finally, we derive EIPBII from EPBII.

### A. EPBII as a single-point infill criterion

#### 1) Definition of PBI

The PBI function is one of the scalarizing functions, such as the weighted sum and the Tchebycheff function. It is used to aggregate objective functions into a scalar function. Minimizing the PBI function along a weight vector crossing the PF in the multiobjective space, we obtain one of the Pareto-optimal solutions (POS).

Along a weight vector $\boldsymbol{\lambda}$ crossing the PF in the multiobjective space, the PBI function for a certain point in the $M$-dimensional objective space $\mathbf{f}(\mathbf{x})$, which is a vector of $M$ objective functions

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2017.2693320, IEEE Transactions on Evolutionary Computation



with *m* design variables **x**, is defined as follows:

$$PBI(\mathbf{f}(\mathbf{x}), \boldsymbol{\lambda}, \mathbf{z}^*) = d_1 + \theta_{PBI} d_2, \tag{9}$$

where

$$d_1 = \frac{\left\| \left(\mathbf{z}^* - \mathbf{f}(\mathbf{x})\right)^{\mathrm{T}} \boldsymbol{\lambda} \right\|}{\|\boldsymbol{\lambda}\|}, \tag{10}$$

$$d_2 = \begin{cases} \left\| \mathbf{f}(\mathbf{x}) - \left( \mathbf{z}^* + d_1 \dfrac{\boldsymbol{\lambda}}{\|\boldsymbol{\lambda}\|} \right) \right\| & \|\mathbf{z}^*\| \le \|\mathbf{f}(\mathbf{x})\| \\ \left\| \mathbf{f}(\mathbf{x}) - \left( \mathbf{z}^* - d_1 \dfrac{\boldsymbol{\lambda}}{\|\boldsymbol{\lambda}\|} \right) \right\| & \|\mathbf{z}^*\| > \|\mathbf{f}(\mathbf{x})\| \end{cases}. \tag{11}$$

$\theta_{PBI}$ is a penalty parameter. PBI corresponds to the distance $d_1$ between the reference point $\mathbf{z}^*$ (utopia point in Fig. 2) and a certain point in the objective space $\mathbf{f}(\mathbf{x})$ with a penalty that is the product of $\theta_{PBI}$ and the distance between $\mathbf{f}(\mathbf{x})$ and the weight vector $\boldsymbol{\lambda}$, as shown in Fig. 5. Therefore, smaller PBI implies that the point is closer to PF and the weight vector. Suitably diverse and converging NDSs can be obtained by minimizing PBI for uniformly distributed weight vectors in the objective space.

2) *Definition of EPBII*

Considering the *M* objective minimization of $f_1, f_2, \cdots, f_M$, EPBII for **x** is derived from the multidimensional Gaussian distribution of the Kriging model. This distribution is denoted by $\mathcal{N}(\hat{f}_i(\mathbf{x}), s_i^2(\mathbf{x}))$ ($i = 1,2,\cdots,M$). After evaluating PBI for each NDS $\tilde{\mathbf{x}}^i$ ($i$ = 1, ⋯, $n^{NDS}$ | $n^{NDS}$ is the number of NDSs among the sample points) for the weight vector $\boldsymbol{\lambda}$, we define the minimum PBI among them as the current reference PBI. This is denoted by $PBI_{ref}$. The PBI improvement (PBII) for the weight vector $\boldsymbol{\lambda}$ is defined as the difference between the reference PBI and PBI of a certain point **f*** (indicated by a cross) for $\boldsymbol{\lambda}$, as shown in Fig. 4. This can be written as follows:

$$PBII(\mathbf{F}, \mathbf{f}^*, \boldsymbol{\lambda}, \mathbf{z}^*) = \max\left(PBI_{ref}(\mathbf{F}, \boldsymbol{\lambda}, \mathbf{z}^*) - PBI(\mathbf{f}^*, \boldsymbol{\lambda}, \mathbf{z}^*), 0\right), \tag{12}$$

where $\mathbf{F} = \left(\mathbf{f}(\tilde{\mathbf{x}}^1), \mathbf{f}(\tilde{\mathbf{x}}^2), \cdots, \mathbf{f}(\tilde{\mathbf{x}}^{n^{NDS}})\right)$ denotes the objective functions of the existing sample points in the current EGO iteration. The dashed lines in Fig. 4 show the PBI contour when $\theta_{PBI}$ = 1. PBII has a positive value inside the reference PBI contour line on which the projected point is located, whereas PBI is not improved and PBII = 0 outside this contour. EPBII is derived from the following integral of PBII and the multidimensional Gaussian distribution in the objective space:

$$EPBII(\mathbf{x}, \mathbf{F}, \boldsymbol{\lambda}, \mathbf{z}^*) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} PBII(\mathbf{F}, \mathbf{f}^*, \boldsymbol{\lambda}, \mathbf{z}^*) \times \varphi(f_1)\varphi(f_2)\cdots\varphi(f_M) df_1 df_2 \cdots df_M, \tag{13}$$

where $\varphi(f_i)$ is the probability density function that is denoted by $\mathcal{N}(\hat{f}_i(\mathbf{x}), s_i^2(\mathbf{x}))$. Integration in Eq. (13) is numerically performed using 500 points that are generated by a Monte Carlo sampling following $\mathcal{N}(\hat{f}_i(\mathbf{x}), s_i^2(\mathbf{x}))$ ($i = 1,2,\cdots,M$). EPBII is the average of PBII for each point.

A candidate additional sample point is obtained by maximizing EPBII. This becomes a single-objective optimization problem. In this study, we maximize EPBII using NSGA-II.

### *B. Multipoint infill criterion derived from EPBII*

Following the procedure in Fig. 3, EPBII may become a multipoint infill criterion.

1) *Definition and clustering of weight vectors*

Weight vectors that are distributed evenly in the objective space must be constructed for the multipoint infill criterion. We generate a weight vector $\boldsymbol{\lambda}$ according to the method in [37], as in the case of MOEA/D [19] and NSGA-III [8]. Thus, the weights for *M* objective functions are selected from

$$\left\{ \frac{0}{H}, \ \frac{1}{H}, \ \cdots, \ \frac{H}{H} \right\}, \tag{14}$$

ensuring that the sum of the weights is one, where *H* is a natural number that defines the number of weight vectors. Considering all combinations of such weights, the number of weight vectors should be

$$N_{ref} = C_{M-1}^{H+M-1}. \tag{15}$$

$C_{M-1}^{H+M-1}$ is the number of ways of choosing (*M* - 1) objects from a set of (*H* + *M* – 1) objects ((*M*-1)-combinations of (*H*+*M*-1) objects). For instance, if *H* = 10 and *M* = 4, then $N_{ref}$ = 286 and one of the weight vectors is written as (1/10, 3/10, 3/10, 3/10). The weight vectors are predefined and are not updated during the EGO iterations in Fig. 3. When we use EPBII as a multipoint infill criterion and an arbitrary number ($n_{add}$) of sample points are added simultaneously during one iteration, the weight vectors are classified into $n_{add}$ clusters by the k-means method according to their weight values.

2) *Definition of a reference point*

As shown in Fig. 3, a reference point for computing EPBII is a utopia point of the NDSs that NSGA-II found on the Kriging models in Fig. 2. Thus, the reference point should be updated at every iteration. The reference point $\mathbf{z}^* = \left(f_{1ref}, f_{2ref}, \cdots, f_{Mref}\right)^{\mathrm{T}}$ is the origin of the weight vectors in Fig. 4. The objective space is normalized by the utopia and nadir points in Fig. 2.

3) *Assignment of NDSs and definition of reference PBI*

For each NDS $\tilde{\mathbf{x}}^i$, we compute $d_2$ values in Eq. (11) to all weight vectors and assign that solution to the closest weight vector with the minimum $d_2$ value, as shown in Fig. 4. After evaluating PBI for all NDSs for the assigned weight vectors, we define the minimum PBI corresponding to each weight vector as the current reference PBI for the vector. Let $\tilde{\mathbf{x}}^i \in \Lambda_j$ ($j$ = 1,⋯, $N_{ref}$) denote that $\tilde{\mathbf{x}}^i$ is assigned to a weight vector $\boldsymbol{\lambda}^j$. Then, the reference PBI for $\boldsymbol{\lambda}^j$ is expressed as

$$PBI_{ref}(\mathbf{F}, \boldsymbol{\lambda}^j, \mathbf{z}^*) = \min_{\tilde{\mathbf{x}}^i \in \Lambda_j} PBI(\mathbf{F}(\tilde{\mathbf{x}}^i), \boldsymbol{\lambda}^j, \mathbf{z}^*), \tag{16}$$

The NDSs with reference PBI are projected to the points (projected points in Fig. 4) on their weight vector, maintaining their PBI. To the weight vectors that have no corresponding NDS (the center vector in Fig. 4), we assign a larger reference PBI than the maximum reference PBI among the vectors with NDSs (1.1 times as large in this study).

4) *Maximization of EPBII in territories*

The exploration territories that are assigned to each weight vector are hyper circular cones (sector forms in two-dimensional objective space). They are defined by the following equations:

$$T(\mathbf{f}(\mathbf{x}), \boldsymbol{\lambda}^j, \mathbf{z}^*) = d_1 - \theta_{ref} d_2 \geq 0, \tag{17}$$

where

$$\theta_{ref} = \frac{1}{\tan\dfrac{\pi}{4H}}. \tag{18}$$

Territories enable EPBII to explore convex and concave PFs. Each weight vector explores only within its territory, i.e., the EPBII is assumed to be zero when $\hat{\mathbf{f}} = \left(\hat{f}_1, \hat{f}_2, \cdots, \hat{f}_M\right)^{\mathrm{T}}$ (estimated objective function values by the Kriging models) that is used in $\mathcal{N}(\hat{f}_i(\mathbf{x}), s_i^2(\mathbf{x}))$ ($i = 1,2,\cdots,M$) does not satisfy Eq. (17). Using this procedure, individuals whose EPBII is zero are frequently generated. In this case, the left-hand side of Eq. (17) is assigned to EPBII as negative fitness in order to lead individuals to the region where positive EPBII can be assigned. We can write this approach as

$$EPBII(\mathbf{x}, \mathbf{F}, \boldsymbol{\lambda}^j, \mathbf{z}^*) = \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} PBII(\mathbf{F}, \mathbf{f}^*, \boldsymbol{\lambda}^j, \mathbf{z}^*) \\ \times \varphi(f_1)\varphi(f_2)\cdots\varphi(f_M) df_1 df_2 \cdots df_M & T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) \geq 0 \\ T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) & T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) < 0 \end{cases}. \tag{19}$$

This approach has the same effect as the penalty that is used in SMS-EGO.

Territories in two-objective problems (sector forms) connect to neighboring territories and fill the objective space that is dominated by the reference point, as shown in Fig. 4. Fig. 6 shows territories and weight vectors on a hyper plane in three-objective space. Gray circles are cross-sections of territories (hyper circular cones) on the hyper plane. In problems involving more than two objectives, neighboring territories contact in a region that is not included in any territories. Solutions in these gaps are led to the interior of territories using Eq. (19). These gaps may cause an issue when EPBII is applied to a many-objective problem with discontinuous PF. Using a higher value for $\theta_{ref}$ than that in Eq. (18) can resolve this issue. However, territories may now overlap.

Candidate additional sample points are obtained by maximizing EPBII for each weight vector. This amount to solving a large number of single-objective optimization problems. In this study, we maximize EPBII for each weight vector individually using NSGA-II. MOEA/D may be used as well. It can obtain all candidates in only one run. However, MOEA/D cannot evaluate the objective function of all individuals in the same generation simultaneously, since the reference point must be updated at every evaluation. As objective function evaluation by the Kriging models includes large matrix operations, the execution speed of MOEA/D is comparable with that of NSGA-II, which can evaluate individuals simultaneously. Additionally, MOEA/D fails to obtain suitably converging NDSs in some test problems. Therefore, we use NSGA-II, which can robustly obtain better NDSs, to eliminate the influence of the optimizer on the exploration capability of EPBII.

5) *Evaluation of multipoint infill criterion*

The multipoint infill criterion is required to be derived from EPBII in order to select $n_{add}$ additional sample points from $N_{ref}$ candidates. Since the weight vectors were previously classified into $n_{add}$ clusters, the sample points are uniformly distributed in the objective space by choosing the candidates corresponding to the maximum EPBII in each cluster. Nevertheless, increasing the number of weight vectors may result in additional sample points converging on vectors in certain regions (e.g., regions where the Kriging models overestimate function values) in each cluster. In order to tackle this problem, we introduce a niche count that is assigned to each weight vector. The niche count for $\boldsymbol{\lambda}^j$ is defined as

$$nc_j = \sum_{i=1}^{N_{ref}} \frac{n_i^{NDS}}{d_{ij}/d_{min} + 1}, \tag{20}$$

where $n_i^{NDS}$ is the number of NDSs among the sample points belonging to $\boldsymbol{\lambda}^i$, $d_{ij}$ is the distance between $\boldsymbol{\lambda}^i$ and $\boldsymbol{\lambda}^j$, and $d_{min}$ is the minimum non-zero $d_{ij}$. Hence, a weight vector passing near NDSs has a high niche count.

To handle a discontinuous PF, we conduct a Pareto ranking among $N_{ref}$ candidates using the estimated objective function values $\hat{\mathbf{f}}^j$ ($j = 1, \cdots, N_{ref}$) of the optimal solutions that are obtained by maximizing EPBII for each weight vector. Fig. 7 illustrates the EPBII approach for a problem with discontinuous PF, where five weight vectors belong to the same cluster ($n_{add} = 1$). We obtain five candidates $\hat{\mathbf{f}}^j$ ($j = 1,2,\cdots,5$) along five vectors by maximizing EPBII, even though one of them ($\hat{\mathbf{f}}^3$) is dominated by another ($\hat{\mathbf{f}}^4$). In this case, we preferentially add candidates whose rank is one in order to obtain NDSs rapidly by decreasing the fitness (EPBII) of a dominated candidate. Moreover, weight vectors whose territories do not cross PF cannot obtain NDSs and become useless in our EPBII approach. Thus, candidates that are found on these vectors should be excluded. The Pareto ranking among candidates is used to achieve this.

Let the rank of the $j$-th candidate be $rank_j$. The fitness of the $j$-th candidate as a multipoint infill criterion is expressed as

$$fitness_j = \frac{EPBII_j}{nc_j \cdot rank_j}, \tag{21}$$

where $EPBII_j$ is the maximized EPBII of the $j$-th candidate. Suitably diverse NDSs can be explored even with a few additional sample points by choosing candidates with the highest fitness in each cluster.

6) *Discussion on the arbitrariness of the parameters*

In the EPBII approach, the only parameters that are chosen arbitrarily are $\theta_{PBI}$ for the PBI definition and $H$ for the weight vector definition (and the parameters in NSGA-II). The other parameters are strictly defined by the aforementioned procedures.

The parameter $\theta_{PBI}$ defines a contour line of PBI in Fig. 5. The angle between the contour line and a weight vector equals $\arccot\theta_{PBI}$. The contour line is perpendicular to the weight vector when $\theta_{PBI} = 0$. In this case, the PBI approach behaves as a weighted sum that cannot obtain well-distributed NDSs in optimization problems with concave PFs. Hence, the PBI approach with small $\theta_{PBI}$ may fail to explore suitable POSs on concave PFs. Additionally, the PBI approach with $\theta_{PBI} = 1$ cannot obtain NDSs around the edges of a convex PF (see [36]). [19] recommends solving optimization problems using $\theta_{PBI} = 5$. However, the region where PBII has positive values (reference region) decreases as $\theta_{PBI}$ increases, since $\arccot\theta_{PBI}$ in Fig. 5 decreases and the reference region is the reference PBI contour line. $\theta_{PBI}$ not only is the penalty parameter but also defines the reference region for evaluating EPBII. If $\theta_{PBI}$ is overly large and the reference region, as the effective integral region, decreases, then the EPBII attains its maximum where the uncertainty defined by Eq. (4) is low. Our preliminary tests showed that most of the multidimensional Gaussian distribution corresponding to the uncertainty fits inside the reference region at the position with the maximum EPBII when an overly large $\theta_{PBI}$ was used ($\theta_{PBI} = 12$ for a two-objective problem). EGO is never achieved with such additional sample points, since the accuracy of the model is not improved. Thus, we recommend $\theta_{PBI} = 1$ or smaller values to balance exploration and exploitation by EPBII, since the territories enable EPBII to handle both convex and concave PF even with such small $\theta_{PBI}$ values. Furthermore, the choice of $H$ depends on computational resources (the total number of additional sample points), since $H$ defines the number of weight vectors $N_{ref}$ by Eq. (17). Thousands of weight vectors are useless if, for instance, only 50 additional sample points are permitted. Therefore, we choose $H$ in order that $N_{ref}$ has the same order as the total number of additional sample points. If EPBII maximization for each weight vector is parallelized or solved by MOEA/D-like algorithms, which are more efficient and can maximize all EPBII for each vector simultaneously, a large value of $H$ may be permissible.

### C. Expected IPBI improvement

The EGO framework with EIPBII is the same as that with EPBII with the exception that we use IPBI instead of PBI. The IPBI function is defined as follows:

$$IPBI(\mathbf{f}(\mathbf{x}), \boldsymbol{\lambda}, \mathbf{z}^*) = d_1 - \theta_{PBI} d_2 . \tag{22}$$

The reference point for IPBI is the nadir point of NDS on the Kriging models, which are obtained by NSGA-II, as shown in Fig. 2. It has the maximum values of the estimated objective functions if multiobjective minimization is conducted. The weight vectors that are used in IPBI can be derived from those used in PBI by multiplying each vector by -1. A point with larger IPBI is closer to PF and the weight vector, as shown in Fig.8. We use $\theta_{PBI} = 1$ in IPBI as well.

The definition of EIPBII is shown in Fig. 9. Since IPBI should be maximized, Eqs. (16) and (12) are replaced by the reference IPBI and IPBI improvement (IPBII), respectively, that are defined as follows:

$$IPBI_{ref}(\mathbf{F}, \boldsymbol{\lambda}^j, \mathbf{z}^*) = \max_{\tilde{\mathbf{x}}^i \in \Lambda_j} IPBI(\mathbf{f}(\tilde{\mathbf{x}}^i), \boldsymbol{\lambda}^j, \mathbf{z}^*), \tag{23}$$

$$IPBII(\mathbf{F}, \mathbf{f}^*, \boldsymbol{\lambda}^j, \mathbf{z}^*) = \max\left(IPBI(\mathbf{f}^*, \boldsymbol{\lambda}^j, \mathbf{z}^*) - IPBI_{ref}(\mathbf{f}, \boldsymbol{\lambda}^j, \mathbf{z}^*), 0\right). \tag{24}$$

EIPBII is computed as EPBII in Eq. (19):

$$EIPBII(\mathbf{x}, \mathbf{F}, \boldsymbol{\lambda}^j, \mathbf{z}^*) = \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} IPBII(\mathbf{F}, \mathbf{f}^*, \boldsymbol{\lambda}^j, \mathbf{z}^*) \\ \times \varphi(f_1)\varphi(f_2)\cdots\varphi(f_M) df_1 df_2 \cdots df_M & T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) \geq 0 \\ T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) & T(\hat{\mathbf{f}}, \boldsymbol{\lambda}^j, \mathbf{z}^*) < 0 \end{cases} \tag{25}$$

Integration in Eq. (25) is numerically performed as in the case of EPBII.

We employ territories, niche count, Pareto ranking, and fitness in the EIPBII approach as well, using the same equations as in the EPBII approach. Candidates with the highest fitness, which is the maximized EIPBII divided by the niche count and the rank, are selected as the additional sample points in each cluster.

## IV. Implementation of Other Infill Criteria

Four other infill criteria (EST, EI, EHVI, and EHVI+EST) are implemented in our EGO framework in order to compare EPBII and EIPBII with them. Overview and implementation of these four criteria are as follows.

### A. Estimation (EST) in multiobjective problems

EST only uses estimated objective function values from the Kriging models. In. Fig. 1, the definition of a reference point and the maximization of the infill criterion are omitted when we use EST. NSGA-II explores NDSs on the Kriging models using estimated objective functions. These NDSs are classified into $n_{add}$ clusters by the k-means method and the solutions closest to the centers of each cluster form the additional sample points.

### B. Expected improvement (EI) in multiobjective problems

According to EGOMOP [21], we compute EI values for each objective function in multiobjective problems. $f_{ref}$ in Eq. (8) becomes the value of each objective function at the nadir point that is defined by the NDSs on the Kriging models, which were found by NSGA-II, as shown in Fig. 2. Moreover, NSGA-II is employed to maximize EI values for multiple objective functions and obtain numerous NDSs of EI values from which additional sample points are selected. Thus, the NDSs are classified into $n_{add}$ clusters by the k-means method and the solutions closest to the centers of each cluster form the additional sample points.

### C. Expected hypervolume improvement (EHVI and EHVI+EST)

We compute EHVI as in [30] due to its simplicity, using numerical integration. However, analytical integration methods for efficiently computing EHVI were recently proposed in [27-29]. Numerical integration in EHVI is performed using 500 points that are generated by a Monte Carlo sampling following $\mathcal{N}(\hat{f}_i(\mathbf{x}), s_i^2(\mathbf{x}))$ $(i = 1,2,\cdots,M)$, as with EPBII and EIPBII. A reference point for calculating the hypervolume corresponds to the nadir point of NDSs on the Kriging models, which were obtained by NSGA-II, as shown in Fig. 2. NSGA-II is used as a single-objective optimizer for EHVI maximization as well.

In the EHVI computation using numerical integration, individuals whose EHVI is zero are frequently generated, since all numerical integration points are likely to be dominated by sample points in an early stage of EHVI maximization with EA. In this case, we propose to use negative fitness that is based on the Pareto ranking (product of the Pareto ranking and −1). This leads individuals to the region in which a positive EHVI can be assigned.

Only one additional sample point, which maximizes the EHVI, is obtained in the EHVI approach, since maximizing EHVI is a single-objective optimization problem. Moreover, we employ Shimoyama's EHVI+EST [30] to explore the NDSs of EHVI and estimated objective functions and add multiple sample points immediately. The EHVI computation method in EHVI+EST is the same as above.

## V. Experimental Setup

We compare the following six criteria for two-objective, three-objective and four-objective test problems (with the exception of EHVI, which we apply to two-objective functions only) in order to investigate the capabilities of EPBII and EIPBII.

1) EST: Involves classifying the NDSs of the estimated objective functions on the Kriging models into $n_{add}$ clusters and simultaneously adding the $n_{add}$ solutions that are closest to the centers of each cluster at every iteration.
2) EI: Involves classifying the NDSs of the EI values for each objective function into $n_{add}$ clusters and simultaneously adding the $n_{add}$ solutions that are closest to the centers of each cluster at every iteration according to EGOMOP [21].
3) EHVI: Involves adding one solution with maximum EHVI at every iteration. Negative fitness is employed to enhance EHVI maximization. This criterion is only applied to two-objective problems.
4) EHVI+EST [30]: Involves classifying the NDSs of EHVI and the estimated objective functions on the Kriging models into $n_{add}$ clusters, then simultaneously adding the $n_{add}$ solutions that are closest to the centers of each cluster at every iteration.
5) EPBII: We pre-classify weight vectors into $n_{add}$ clusters and simultaneously add the $n_{add}$ solutions that have the highest fitness related to EPBII in each cluster at every iteration.
6) EIPBII: We pre-classify weight vectors into $n_{add}$ clusters and simultaneously add the $n_{add}$ solutions that have the highest fitness related to EIPBII in each cluster at every iteration.

### A. Test Problem Definition

We select the test problems from previous MOEA/D-EGO studies. As these methods can add multiple sample points simultaneously, we compare the results of our experiments with those of MOEA/D-EGO that are reported in [20]: ZDT1-4, 6 [38], LZ08-F1-4 [39] ($M$ = 2, $m$ = 8), and DTLZ2 [40] ($M$ = 2, 3, $k$ = 4, $m = M + k - 1$). In all these problems, the objective functions should be minimized. Moreover, we include three types of modified DTLZ2 problems (DTLZ2max1-3), where maximization is required. The objective functions of these problems can be written as follows:

$$\begin{aligned}
f_1(\mathbf{x}) &= g(\mathbf{x}_M)\cos(x_1\pi/2)\cdots\cos(x_{M-2}\pi/2)\cos(x_{M-1}\pi/2)\\
f_2(\mathbf{x}) &= g(\mathbf{x}_M)\cos(x_1\pi/2)\cdots\cos(x_{M-2}\pi/2)\sin(x_{M-1}\pi/2)\\
f_3(\mathbf{x}) &= g(\mathbf{x}_M)\cos(x_1\pi/2)\cdots\cos(x_{M-3}\pi/2)\sin(x_{M-2}\pi/2)\\
&\cdots\\
f_{M-1}(\mathbf{x}) &= g(\mathbf{x}_M)\cos(x_1\pi/2)\sin(x_2\pi/2)\\
f_M(\mathbf{x}) &= g(\mathbf{x}_M)\sin(x_1\pi/2),
\end{aligned} \tag{26}$$

where $g(\mathbf{x}_M)$ in Eq. (26) and the domains of the design variables are defined as follows:

1) For DTLZ2max1:

$$g(\mathbf{x}_M) = \frac{1}{k}\sum_{i=M}^{m}\left[1 - 4(x_i - 0.5)^2\right] \tag{27}$$

$$0 \le x_i \le 1 \quad (i = 1,2,\cdots,m)$$

$$M = 2, 3,\ k = 4,\ m = M + k - 1.$$

2) For DTLZ2max2:

$$g(\mathbf{x}_M) = \frac{1}{k}\sum_{i=M}^{m}\left[1 - 4(x_i - 0.5)^2\right] \tag{28}$$

$$0.25 \le x_i \le 0.75 \quad (i = 1,2,\cdots,M-1)$$

$$0 \le x_i \le 1 \quad (i = M, M+1,\cdots,m)$$

$M$ = 2, 3, 4, $k$ = 4, $m = M + k$ - 1.

3) For DTLZ2max3:

$$g(\mathbf{x}_M) = \frac{1}{k}\sum_{i=M}^{m}\left[1-(x_i-0.5)^2+\frac{1}{3}\{\cos(4\pi(x_i-0.5))-1\}\right] \quad (29)$$

$$0.25 \le x_i \le 0.75 \ \ (i = 1,2,\cdots,M-1)$$

$$0 \le x_i \le 1 \ \ (i = M, M+1,\cdots,m)$$

$M$ = 2, 3, 4, $k$ = 4, 8, $m = M + k - 1$.

Fig. 10 shows PF (black line) and the feasible region (gray region) of the two-objective DTLZ2max1-3 problems. We include DTLZ2max1, since it does not have weak POSs and real-world problems may not have these either. The reference point strongly affects the exploration capability of EPBII and EIPBII. However, it is difficult for our EGO framework with EPBII and EIPBII to identify a suitable reference point in DTLZ2 due to its weak POSs. Since DTLZ2 is a test problem, we know that $f_i$ = 0 and $f_i$ = 1 ($i = 1,2,\cdots,M$) are the suitable reference points for EPBII and EIPBII, respectively. Nevertheless, we use the utopia and nadir points of the estimated PF as the reference points for EPBII and EIPBII in consideration of real-world applications. Moreover, PFs of real-world problems usually cover only a portion of the feasible region, as shown in Fig. 11. This describes the result of high-lift airfoil optimization (maximization of lift coefficients at different angles of attack) [41]. DTLZ2max2 mimics this characteristic. Finally, we include DTLZ2max3, since this problem is difficult to approximate owing to its local PFs. The difference between DTLZ2max2 and DTLZ2max3 appears in $g(\mathbf{x}_M)$, whose responses with respect to $x_M$ ($k$ = 1) are shown in Fig. 12.

### B. Parameter settings

The number of initial sample points that is generated by LHS is set to $11m - 1$, as suggested in [17]. The number of additional sample points in each iteration $n_{add}$ is one for EHVI and five for the other criteria. Sample points are added until the total number of initial and additional sample points reaches 200 for two-objective problems (with the exception of DTLZ2max3, where $k$ = 8) and 300 for the other problems.

The number of weight vectors for EPBII and EIPBII is set to $N_{ref}$ = 101 ($H$ = 100) for two-objective problems, $N_{ref}$ = 231 ($H$ = 20) for three-objective problems, and $N_{ref}$ = 286 ($H$ = 10) for four-objective problems.

The population and generation count in NSGA-II for optimizing each criterion are 500 and 100, respectively, for EST, EI, EHVI, and EHVI+EST. In order to maximize EPBII and EIPBII for each weight vector, the population count is set to 200 and the generation count to 50, respectively. We set the other parameters of NSGA-II as follows.

1) Crossover probability: 0.9.
2) Crossover parameter: $\eta_c$ = 10.
3) Mutation probability: 1/$m$.
4) Mutation parameter: $\eta_m$ = 20.

### C. Performance Metrics

In order to compare the exploration capability of the criteria, we compute the hypervolume and the inverted generational distance (IGD) [19] consisting of sample points at every iteration. The reference point for the hypervolume is ($f_1$, $f_2$) = (20, 20) for ZDT3, (100, 100) for ZDT4, $f_i$ = 0 ($i = 1,2,\cdots,M$) for DTLZ2max1-3, and $f_i$ = 10 ($i = 1,2,\cdots,M$) for the other problems. We subtract the computed hypervolume from the hypervolume consisting of the POSs and denote the result by $I_H^-$ [42]. The number of evaluation points for computing IGD is 1,000 for two-objective problems, $31^2$ = 961 for three-objective problems, and $21^3$ = 9,261 for four-objective problems. Both $I_H^-$ and IGD decrease as better NDSs are found. We conduct 10 independent runs starting with different initial sample points in order to evaluate the mean, standard deviation, best, and worst $I_H^-$ and IGD values for comparison. The mean values are mainly used in the following discussion. The other values are shown in Appendices A-D.

## VI. Results and Discussion

We use both $I_H^-$ and IGD in the following discussion. IGD is suitable for evaluating EPBII and EIPBII, since optimization methods that are based on weight vectors, including EPBII and EIPBII, are suitable for obtaining well-distributed NDSs and IGD can measure the distribution of NDSs. Optimization methods that are based on hypervolume, including EHVI and EHVI+EST, should be evaluated by $I_H^-$, as hypervolume maximization is generally not compatible with IGD maximization. Hence, we show both $I_H^-$ and IGD values in the tables in order to clarify the correlation and difference between these two metrics. Fortunately, $I_H^-$ and IGD values indicate nearly the same trends, as shown in the following discussion.

### A. Performance comparison on well-known problems

Tables I and II show the mean $I_H^-$ and IGD values for the six criteria and MOEA/D-EGO at the end of EGO iterations. For each test problem, the lowest (top) values among the six criteria are highlighted in bold face with gray background and the second lowest values are highlighted in bold face. We cite the MOEA/D-EGO results [20] and directly list them in Tables I and II for comparison, even though MOEA/D-EGO was executed with different conditions from those of our EGO with the six criteria. That is, MOEA/D-EGO uses multiple Kriging models for overlapping clusters of sample points, a utopia point among the sample points as a reference point, differential evolution with neighboring weight vector information for EI maximization, more weight vectors than our EGO and different initial sample sets.

1) *ZDT1-4, 6*

EPBII, EIPBII, and EHVI+EST show better performance than the other three criteria and comparable with that of MOEA/D-EGO in most of the ZDT problems, as shown in Tables I and II. In ZDT1-3, the diversity of the NDSs that are

obtained by EST and EI is poor, even though all six criteria, including these two, can find suitably converging NDSs due to the high approximation accuracy of the Kriging models. Figs. 13-15 show the NDSs for ZDT1-3 obtained by EPBII and EIPBII. NDSs with the values of $I_H^-$ that are closest to the mean value in 10 runs are presented. Both EPBII and EIPBII with $\theta_{PBI}$ = 1 successfully solve ZDT1 with convex PF and ZDT2 with concave PF, since territories perform well. Moreover, the Pareto ranking among candidates enables EPBII and EIPBII to efficiently explore NDSs for ZDT3 with discontinuous PF. The $I_H^-$ values of EIPBII and EHVI+EST are significantly lower than the other values for ZDT3, since these criteria can robustly find all five discontinuous PFs in 10 runs.

By contrast, no criterion can reach PF in ZDT4 and 6 due to the poor approximation accuracy of the Kriging models. However, EIPBII and EPBII demonstrate equal or superior exploration capability compared with MOEA/D-EGO (comparable or better IGD and $I_H^-$) in the above ZDT problems, including ZDT4. Moreover, they achieve the lowest mean $I_H^-$ and IGD for ZDT6 among the six criteria. The NDSs corresponding to the mean $I_H^-$ that is obtained by EPBII for ZDT4 and EIPBII for ZDT6 are shown in Fig. 16. Overall, MOEA/D-EGO achieves significantly lower $I_H^-$ and IGD for ZDT6 than the six criteria, since its model accuracy is high, as mentioned in [20]. This is due to the fact that MOEA/D-EGO uses multiple Kriging models for overlapping clusters, whereas $I_H^-$ and IGD of the six criteria are comparable with those of ParEGO [18].

2) *LZ08-F1-4*

All six criteria find suitably diverse and converging NDSs for the LZ08 problems. These are more difficult for the Kriging models to approximate than ZDT1, which has the same PF. EPBII and EIPBII obtain such NDSs, as well (as shown in Fig. 17), even though EHVI achieves the lowest $I_H^-$ and IGD in most problems. This is due to the fact that only EHVI adds sample points individually, whereas the other criteria add five points simultaneously. The POSs of the LZ08 problems have different optimal values of the design variables. This implies that we must improve the model accuracy in the entire design space, where POSs are scattered, to obtain diverse NDSs. Adding one sample point in each iteration enhances model accuracy, since the number of model updates increases. By contrast, all POSs of the ZDT problems have the same optimal value $x_i$ = 0 ($i = 2, \cdots, m$) except $x_1$. Thus, obtaining diverse NDSs is easy if one NDS near the PF is found and the model accuracy around the POSs in a small region of the design space is improved.

MOEA/D-EGO has an extremely high $I_H^-$ and IGD compared with the six criteria for LZ08-F2-4, despite showing comparable results for LZ08-F1 (the simplest problem among these four problems). Even EST and EI are better than MOEA/D-EGO. Zhang et al. reported that MOEA/D-EGO and ParEGO show almost the same exploration capability in these problems [20]. This suggests that aggregation by the Tchebycheff function or its approximation may cause some problems for LZ08-F2-4, otherwise the reference point, which is defined by the NDSs on the Kriging model in this study and by the NDSs among the sample points in MOEA/D-EGO may affect their exploration capability.

3) *DTLZ2*

EPBII and EIPBII achieve the lowest values of $I_H^-$ and IGD in the two-objective and three-objective DTLZ2, except for the three-objective DTLZ2, where IGD has the second lowest value. Fig. 18 presents the NDSs for the three-objective DTLZ2 that are obtained by EPBII and EIPBII. EPBII accomplishes the lowest $I_H^-$ and IGD in both the two-objective and three-objective DTLZ2, since the weight vector distribution of EPBII fits the concave PF of DTLZ2.

By contrast, EIPBII has the highest IGD among the criteria, except for MOEA/D-EGO, for the three-objective DTLZ2, even though EIPBII has the second lowest $I_H^-$. EIPBII tends to add many sample points to the region where weak POSs exist, since the outermost weight vectors lose their assigned NDSs at every iteration with the movement of the reference points. These reference points are defined as the nadir points of the NDSs on the Kriging models and are easily affected by weak POSs. These NDSs enable EIPBII to robustly decrease $I_H^-$. However, they prevent it from adding sample points around the center of the PF and decrease IGD. Therefore, the exploration capability of EIPBII depends on the selection of reference points, especially in problems involving three or more objectives with weak POSs. Even in two-objective problems, EIPBII is less likely to add sample points around the center of the PF if only one or two additional sample points are available in each iteration. Hence, more than two sample points should be added simultaneously when EIPBII is applied to two-objective problems with weak POSs.

Compared with MOEA/D-EGO, all six criteria achieve lower IGD due to their reference point definition, which increases the update speed of the reference point. By contrast, the reference point in MOEA/D-EGO is gradually updated, since it is defined by the NDSs among the sample points.

### *B. Performance comparison on the DTLZ2max1-3 problems*

Tables III and IV summarize the $I_H^-$ and IGD values for the six criteria at the end of the EGO iterations, where the lowest values are in bold with a grey background and the second lowest values are in bold.

1) *DTLZ2max1*

Both EPBII and EIPBII obtain NDSs with low both $I_H^-$ and IGD for the two-objective and three-objective DTLZ2max1, which has no weak POS. Moreover, EHVI achieves nearly the same $I_H^-$ and IGD as EIPBII with two objectives. This demonstrates that EIPBII can explore suitably diverse and converging NDSs for problems with convex PFs without weak POSs as well as or better than EPBII, as shown in Fig. 19.

2) *DTLZ2max2*

EIPBII accomplishes the lowest $I_H^-$ and IGD in all DTLZ2max2 problems. EPBII and EHVI follow. EIPBII obtains NDSs on the weight vectors crossing the PF, even though the number of objective functions increases, as seen in Fig. 20. EIPBII and EPBII may be desirable criteria due to their low $I_H^-$ and IGD, even for the four-objective DTLZ2max2.

EIPBII has a higher IGD than the other criteria in an early stage of EGO in Fig. 21(a), even though its $I_H^-$ with a small number of sample points is comparable with that of the other criteria, as shown in Fig. 21(b). In DTLZ2max2, the number of weight vectors that do not cross the PF increases, as the number of objective functions increases. These vectors never cross the feasible regions in EPBII, whereas those in EIPBII cross the feasible regions, which do not include POSs, and make EIPBII add sample points at undesirable positions. Fitness in EIPBII is modified by the Pareto ranking and Eq. (21) in order to handle not only a discontinuous PF but also these regions, thereby helping to obtain well-distributed NDSs at the end of EGO. However, to obtain better NDSs with fewer sample points, the weight vector generation method should be modified in order to increase the vectors crossing such a PF in DTLZ2max2. We note that this is an issue not only for EPBII and EIPBII but also for EAs using weight vectors (e.g., MOEA/D and NSGA-III). Some approaches have already been proposed in [7,43].

EST and EI have higher $I_H^-$ and IGD. This is derived from the characteristic distributions of NDSs that are obtained by each criterion in Fig. 22(a), which shows the NDSs for the two-objective DTLZ2max2 with $I_H^-$ closest to the mean value that is obtained by EI. The NDSs that are obtained by EI converge at five particular positions on the PF, whereas EIPBII finds all suitably converging NDSs along all weight vectors, as shown in Fig. 20(a). We note that EI and EST fail to obtain suitably diverse NDSs in multiobjective optimization, since most of the sample points are added around the centers of each cluster that is obtained by the k-means method. Changing the number of clusters at each iteration and the position of the reference point for calculating EI values may resolve this. However, the former cannot fully utilize parallel computation and the latter may violate objectivity and repeatability. EHVI+EST shows the same phenomenon for three-objective DTLZ2max2, as seen in Fig. 22(b).

3) *DTLZ2max3*

EPBII and EIPBII show better performance than the other criteria in most DTLZ2max3 problems, which are difficult to approximate with the Kriging models. Fig. 23 shows the NDSs for the two-objective DTLZ2max3 ($k$ = 8) corresponding to the mean $I_H^-$ that is obtained by EPBII and EIPBII. Both criteria obtain NDSs closer to the true PF than the other criteria and reach the true PF in the best run, even though they fall into the local PF, as shown in Fig. 23. EST, EI, and EHVI+EST cannot obtain NDSs on the true PF in any run. EIPBII successfully avoids local PFs and attains better NDSs on next (local) PFs in many runs for the two-objective DTLZ2max3 ($k$ = 4 and 8), thereby balancing exploration and exploitation in the EGO process. By contrast, EPBII has higher $I_H^-$ and IGD than EHVI, EHVI+EST, and EIPBII. However, it achieves the second lowest best and worst $I_H^-$ and IGD value, respectively (see Appendices C and D), for the two-objective DTLZ2max3 ($k$ = 4). EPBII can obtain suitably diverse NDSs. However, it is likely to fall into the local PF, leading to high mean values of the two metrics.

By contrast, EPBII achieves the lowest $I_H^-$ and IGD for the four-objective DTLZ2max3 ($k$ = 8). The best $I_H^-$ and IGD (in Appendices C and D) that are obtained by EPBII for the four-objective DTLZ2max3 ($k$ = 8) are nearly the same as those for the four-objective DTLZ2max2 and DTLZ2max3 ($k$ = 4). This suggests that EPBII can robustly obtain diverse NDSs with respect to the number of objective functions. However, EIPBII has nearly the same $I_H^-$ and IGD as EST, EI, and EHVI+EST for the four-objective DTLZ2max3 ($k$ = 8) due to the increase in weight vectors that do not cross the PF, as discussed for DTLZ2max2.

EHVI has a significantly high mean $I_H^-$ and IGD for the two-objective DTLZ2max3 ($k$ = 8), even though it achieves the lowest $I_H^-$ and IGD ($k$ = 4). We note that when model accuracy is poor, EHVI cannot efficiently obtain good NDSs. However, EPBII and EIPBII robustly obtain better NDSs.

### C. Effect of number of additional sample points

In order to investigate the effect of the number of additional sample points in each iteration, exploration capabilities of EPBII and EIPBII with $n_{add}$ = 3, 5, and 10 are compared in LZ08-F2, as shown in Figs. 24. All their other settings are the same as EPBII and EIPBII with $n_{add}$ = 5 that are employed above. Since the number of model updates decreases and the model accuracy improvement becomes small, EPBII and EIPBII with larger $n_{add}$ have higher $I_H^-$ though EPBII with $n_{add}$ = 3 and 5 has nearly the same $I_H^-$. EPBII and EIPBII with large $n_{add}$ may be useful if our priority is decreasing the number of model updates and total time spent on optimization. EPBII and EIPBII with $n_{add}$ = 10 can obtain well-distributed NDSs and decrease $I_H^-$ in a short time with a small number of model updates when parallel evaluation is available, as shown in Fig. 25. This is an advantage of EPBII and EIPBII as multipoint infill criteria.

### D. Choice between EPBII and EIPBII

We recommend using EIPBII for two-objective problems, since it can robustly solve problems with discontinuous PFs and local PFs. All weight vectors should cross PF in two-objective problems unless PF is discontinuous. This implies that the number of useless weight vectors that do not cross PF is limited in two-objective problems. Hence, the drawback of EIPBII is not serious. Moreover, EIPBII can overcome weak POSs in two-objective problems if more than two sample points are added simultaneously.

For problems with more than two objectives, using EPBII is

preferable, since it can robustly obtain diverse NDSs in up to four-objective problems. In particular, we recommend using EPBII and against using EIPBII when the shape of PF is not known and the total number of additional sample points is small.

### *E. Superiority of EPBII/EIPBII and EHVI*

The features of NDSs that are obtained by each criterion have already been discussed above. We focus on other aspects of EPBII/EIPBII and EHVI in order to suggest criteria that are suitable for specific problems.

EPBII and EIPBII can add multiple sample points immediately and retain the quality of NDSs at comparable level with that of EHVI as a single point infill criterion. The number of sample points that is added simultaneously does not affect the computational cost for calculating EPBII/EIPBII, since only the number of weight vector clusters changes. Thus, EPBII and EIPBII are suitable for problems where multiple sample points can be evaluated at the same time. Additionally, EPBII and EIPBII can obtain well-distributed NDSs according to their weight vector distribution. EPBII and EIPBII are useful if our priority is the diversity of NDSs.

Moreover, EHVI can be computed analytically and efficiently by the method proposed in [27-29] when the number of objective functions is small, even though EHVI is computed numerically in this study. Analytic computations prevent numerical noise and may improve exploration capability. Furthermore, EHVI can efficiently improve hypervolume consisting of NDSs. Hence, EHVI is preferred as a single-point infill criterion if our priority is the hypervolume and the number of objective functions is not overly large.

## VII. CONCLUSION

We proposed the new infill criteria EPBII and EIPBII for efficiently solving multiobjective optimization problems with expensive objective functions. These criteria appropriately determine multiple candidate additional sample points for updating the Kriging surrogate model. The definition corresponds to the EI of PBI and IPBI, which are defined for uniformly-distributed weight vectors in the objective space and enable EGO to explore well-distributed NDSs. We introduce territories and niche counts for each weight vector and the Pareto ranking of candidates in order to select suitable additional sample points from a large number of candidates. Moreover, these territories eliminate the need to tune the penalty parameter in PBI and IPBI. We investigated the exploration capabilities of EPBII and EIPBII by comparing them with other EGO approaches for various test problems.

For most of the two-objective test problems, both EPBII and EIPBII could find NDSs with better diversity and convergence, compared with conventional criteria, such as EI and EHVI. Moreover, they demonstrated an exploration capability that met or exceeded that of MOEA/D-EGO in these two-objective problems and the three-objective DTLZ2. Territories and the Pareto ranking enable EPBII and EIPBII with small values of the penalty parameter in PBI and IPBI to efficiently explore NDSs on convex, concave, and discontinuous PFs. EIPBII obtained desirable NDSs for problems with discontinuous PFs and local PFs more robustly than EPBII. Therefore, we recommend using EIPBII for two-objective problems, even though analytically integrated EHVI may be another choice for problems where sample points are evaluated individually.

Furthermore, EPBII and EIPBII could obtain suitably diverse and converging NDSs even in three-objective and four-objective problems, whereas EST, EI, and EHVI+EST failed to add suitably diverse NDSs. Moreover, EPBII and EIPBII robustly achieved optimization for problems that were difficult to approximate with the Kriging models, whereas EHVI did not perform well when the model accuracy was poor. The exploration capability of EIPBII depends heavily on the definition of the reference point. We discovered that weak POSs in the test problems hinder the ability of EIPBII to identify the suitable reference point, especially in problems with more than two objectives. Moreover, in some problems, the number of weight vectors that cross the feasible regions and not the PF increased with the number of objective functions. These vectors made EIPBII add sample points at undesirable positions. Thus, EPBII is suitable for solving problems with more than two objectives.

Compared with EHVI-based infill criteria, EPBII and EIPBII are essentially applicable to many-objective problems without high computational cost, even though only EPBII was found to be suitable for many-objective problems. Hence, the method for generating weight vectors should be modified in order to decrease the number of useless weight vector that do not cross PF. Additionally, EHVI was computed by numerical integration using Monte Carlo sampling, even though the latest studies employed exact computation of EHVI with analytical integration. Comparison between EPBII/EIPBII with modified weight vectors and exactly computed EHVI should be conducted in future work.

**Nobuo Namura** (S'14) received a B.Sc. degree in mechanical and aerospace engineering, an M.Sc. degree in aerospace engineering, and a Ph.D. degree in aerospace engineering from Tohoku University, Sendai, Japan, in 2011, 2013, and 2016, respectively.

He was a Research Fellow of the Japan Society for the Promotion of Science from 2014 to 2016. His research interests include surrogate-assisted evolutionary computation, computational fluid dynamics, and aerodynamic design of aircraft.

**Koji Shimoyama** (M'10) received a B.Sc. degree in mechanical engineering, an M.Sc. degree in aeronautics and astronautics, and a Ph.D. degree in aeronautics and astronautics from The University of Tokyo, Japan, in 2001, 2003, and 2006, respectively.

He was a Research Fellow from 2006 to 2009 and an Assistant Professor from 2009 to 2014. Since 2014 he has been an Associate Professor with the Institute of Fluid Science, Tohoku University, Japan. In addition, he was concurrently a Visiting Assistant Professor with the Department of Aeronautics and Astronautics, Stanford University, CA, from 2012 to 2013 and an Invited Professor with Laboratoire de Tribologie et Dynamique des Systèmes (LTDS), École Centrale de Lyon, France, in 2013. His current research interests include multiobjective optimization, uncertainty quantification, and decision making for real-world engineering design.

**Shigeru Obayashi** (M'96) received a B.Sc. and M.Sc. degree from the University of Tsukuba in 1982 and 1984, respectively. He received a Doctorate in Engineering from The University of Tokyo in 1987.

He worked as a visiting scientist at NASA Ames Research Center starting in 1987 until he joined the faculty of engineering at Tohoku University as an associate professor in 1994. He has been a professor at the Institute of Fluid Science, Tohoku University, Sendai, Japan, since 2003. He has also served as the Director of the Institute since 2014. His current research interests include computational fluid dynamics, multidisciplinary design optimization, evolutionary computation, data mining, and their real-world applications.

Prof. Obayashi is currently President of the Japanese Society for Evolutionary Computation (2015–2016), Associate Fellow of the American Institute of Aeronautics and Astronautics, Fellow of the Japan Society for Aeronautical and Space Sciences, Fellow of the Japan Society of Mechanical Engineers, and Fellow of the Japan Society of Fluid Mechanics. His awards and honors include the 1993 NASA Ames Honor Award, the 2012 Computational Mechanics Award, the Japan Society of Mechanical Engineers, and the 2014 Commendation for Science and Technology from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

## REFERENCES

[1] K. Sugimura, S. Jeong, S. Obayashi, and T. Kimura, "Kriging-model-based multi-objective robust optimization and trade-off rule mining of a centrifugal fan with dimensional uncertainty," *J. Comput. Sci. Tech.*, vol. 3, no. 1, pp. 196-211, Feb. 2009.

[2] K. Shimoyama, S. Yoshimizu, S. Jeong, S. Obayashi, and Y. Yokono, "Multi-objective design optimization for a steam turbine stator blade using LES and GA," *J. Comput. Sci. Tech.*, vol. 5, no. 3, pp. 134-147, Nov. 2011.

[3] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI: Univ. Michigan Press, 1975.

[4] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, MA: Addison-Wesley, 1989.

[5] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, U.K.: John Wiley and Sons, 2001.

[6] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: a short review," in *Proc. IEEE Congr. Evol. Comput.*, Hong Kong, China, 2008 pp.2424-2431.

[7] Y. Tan, Y. Jiao, H. Li, and X. Wang, "MOEA/D + uniform design: A new version of MOEA/D for optimization problems with many objectives," *Comput. Oper. Res.*, vol. 40, no. 6, pp. 1648-1660, Jun. 2013.

[8] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577-601, Aug. 2014.

[9] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 602-622, Aug. 2014.

[10] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition", *IEEE Trans. Evol. Comput.*, vol. 19, no. 5, pp. 694-716, Oct. 2015.

[11] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P.K. Tucker, "Surrogate-based analysis and optimization," *Prog. Aero. Sci.*, vol. 41, no. 1, pp. 1-28, Jan. 2005.

[12] A. Hashimoto, S. Jeong, and S. Obayashi, "Aerodynamic optimization of near-future high-wing aircraft," *Trans. Japan Soc. Aero. Space Sci.*, vol. 58, no. 2, pp. 73-82, Mar. 2015.

[13] S. Jeong, S. Hasegawa, K. Shimoyama, and S. Obayashi, "Development and investigation of efficient GA/PSO-hybrid algorithm applicable to real-world design optimization," *IEEE Comput. Intell. Mag.*, vol. 4, no. 3, pp. 36-44, Aug. 2009.

[14] K. Shimoyama, K. Seo, T. Nishiwaki, S. Jeong, and S. Obayashi, "Design Optimization of a sport shoe sole structure by evolutionary computation and finite element method analysis," *Proc. Inst. Mech. Eng., Part P: J. Sports Eng. Tech.*, vol. 225, no. 4, pp.179-188, Nov. 2011.

[15] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, no. 8, pp. 1246-1266, Dec. 1963.

[16] J. Mockus, V. Tiesis, and A. Zilinskas, "The application of Bayesian methods for seeking the extremum, " *Towards Global Optimization*, vol. 2, pp. 117-129, 1978.

[17] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box function," *J. Global Optim.*, vol. 13, no. 4, pp. 455-492, Dec. 1998.

[18] J. Knowles, "ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans. Evol. Comput.*, vol. 10, no. 1, pp. 50-66, Feb. 2006.

[19] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp.712-731, Dec. 2007.

[20] Q. Zhang, W. Liu, E. Tsang, and B. Virginas, "Expensive optimization by MOEA/D with Gaussian process model," *IEEE Trans. Evol. Comput.*, vol. 14, no. 3, pp. 456-474, Jun. 2010.

[21] S. Jeong, Y. Minemura, and S. Obayashi, "Optimization of combustion chamber for diesel engine using Kriging model," *J. Fluid Sci. Tech.*, vol. 1, no. 2, pp. 138-146, Dec. 2006.

[22] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264-323, Sep. 1999.

[23] W. Ponweiser, T. Wagner, D. Biermann, M. Vincze, "Multiobjective optimization on a limited budget of evaluations using model-assisted S-metric selection," in *Proc. Parallel Problem Solving from Nature (PPSN) X*, Dortmund, Germany, Sep. 2008, pp. 784-794.

[24] M. Emmerich, K. Giannakoglou, and B. Naujoks, "Single and multiobjective evolutionary optimization assisted by Gaussian random field metamodels," *IEEE Trans. Evol. Comput.*, vol. 10, no. 4, pp. 421-439, Aug. 2006.

[25] Ł. Łaniewski-Wołłk, S. Obayashi, and S. Jeong, "Development of expected improvement for multi-objective problem," in *Proc. 42nd Fluid Dynamics Conference/Aerospace Numerical Simulation Symposium*, Yonago, Japan, Jun. 2010.

[26] T. Wagner, M. Emmerich, A. Deutz, and W. Ponweiser, "On expected-improvement criteria for model-based multi-objective optimization," in *Proc. Parallel Problem Solving from Nature (PPSN) XI*, Kraków, Poland, Sep. 2010, pp. 718-727.

[27] M. T. M. Emmerich, A. H. Deutz, and J. W. Klinkenberg, "Hypervolume-based expected improvement: monotonicity properties and exact computation," in *Proc. IEEE Congr. Evol. Comput.*, New Orleans, LA, Jun 2011, pp.2147-2154.

[28] I. Hupkens, A. Deutz, K. Yang, and M. Emmerich, "Faster exact algorithms for computing expected hypervolume improvement," in *Proc. 8th Evol. Multi-criterion Optim. Part 2*, Guimarães, Portugal, 2015, pp. 65-79.

[29] M. Emmerich, K. Yang, A. Deutz, H. Wang, and C. M. Fonseca, "A multicriteria generalization of Bayesian global optimization," in *Advances in Stochastic and Deterministic Global Optim.*, vol. 107 of *Springer Optimization and Its Applications*, Switzerland: Springer International Publishing, 2016, pp. 229—242.

[30] K. Shimoyama, K. Sato, S. Jeong, and S. Obayashi, "Updating Kriging surrogate models based on the hypervolume indicator in multi-objective optimization," *J. Mech. Design*, vol. 135, no. 9, pp. 094503-1-7, Jul. 2013.

[31] D. Horn, T. Wagner, D. Biermann, C. Weihs, and B. Bischl, "Model-based multi-objective optimization: taxonomy, multi-point proposal, toolbox and benchmark," in *Proc. 8th Evol. Multi-criterion Optim. Part 1*, Guimarães, Portugal, 2015, pp. 64-78.

[32] D. R. Jones, "A Taxonomy of Global Optimization Methods Based on Response Surfaces," *J. Global Optim*., vol. 21, no. 4, pp. 345–383, Dec. 2001.

[33] H. Sato, "Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization," in *Proc. Genetic Evol. Comput. Conf*., Vancouver, Canada, Jul. 2014, pp. 645-652.

[34] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239-245, May 1979.

[35] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182-197, Apr. 2002.

[36] H. Ishibuchi, N. Akedo, Y. Nojima, "A study on the specification of a scalarizing function in MOEA/D for many-objective knapsack problems," in *Proc. 7th Learn. Intell. Optim. Catania*, Italy, 2013, pp. 231-246.

[37] I. Das and J. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optim.*, vol. 8, no. 3, pp. 631-657, Aug. 1998.

[38] K. Deb, "Multiobjective genetic algorithms: Problem difficulties and construction of test problems," *Evol. Comput.*, vol. 7, no. 3, pp. 205-230, Feb. 1999.

[39] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 12, no. 2, pp. 284-302, Apr. 2009.

[40] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable multiobjective optimization test problems," in *Proc. IEEE Congr. Evol. Comput.*, Honolulu, HI, May 2002, pp.825-830.

[41] M. Kanazaki, K. Tanaka, S. Jeong, and K. Yamamoto. "Multi-objective aerodynamic optimization of elements' setting for high-lift airfoil using Kriging model," in *Proc. 44th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, Jan. 2006, AIAA2006-1471.

[42] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 117-132, Apr. 2003.

[43] R. Cheng, Y. Jin, and K. Narukawa, "Adaptive reference vector generation for inverse model based evolutionary multiobjective optimization with degenerate and disconnected Pareto fronts," in *Proc. 8th Evol. Multi-criterion Optim. Part 1*, Guimarães, Portugal, 2015, pp. 127-140.

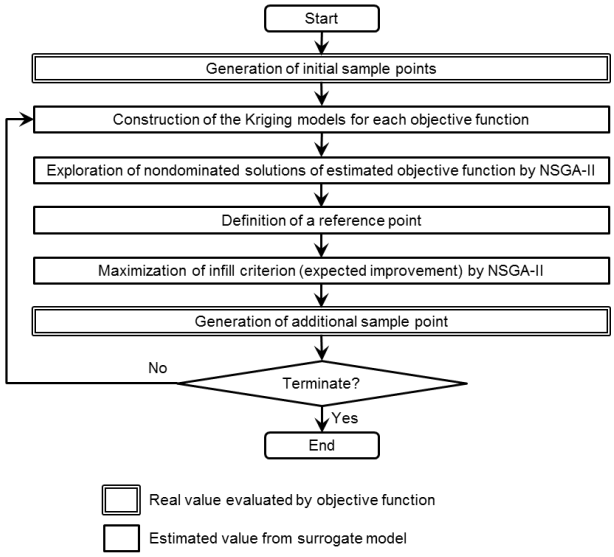



Fig. 1. Flowchart of the efficient global optimization.

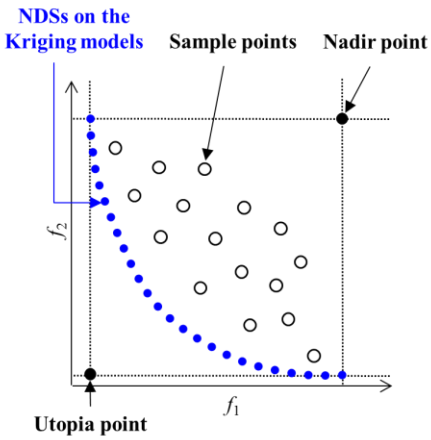



Fig. 2. Reference point definition using a utopia or nadir point of nondominated solutions (NDSs) on the Kriging models constructed from sample points.

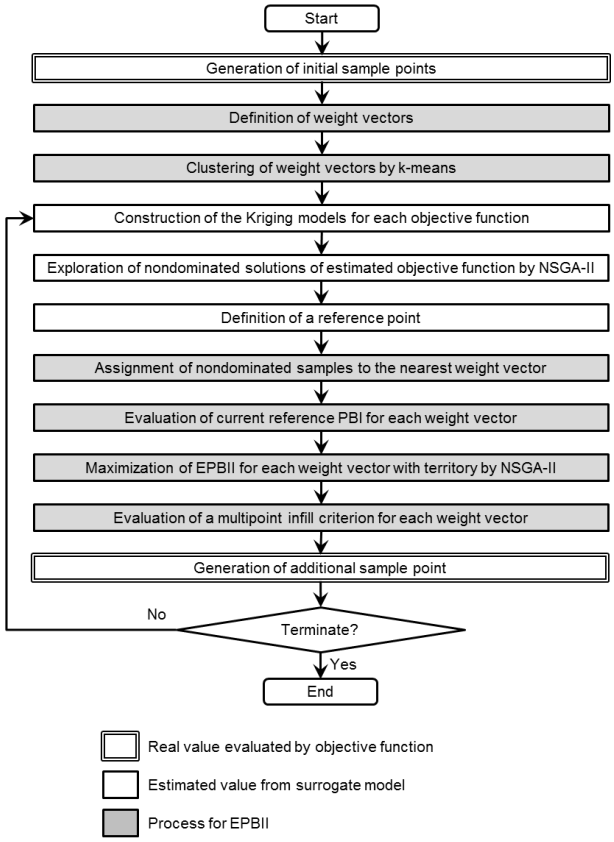



Fig. 3. Flowchart of the efficient global optimization with EPBII.

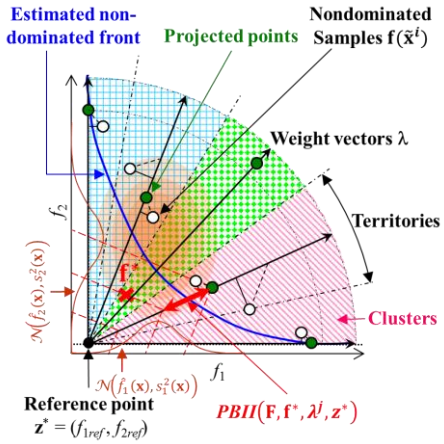



Fig. 4. Definition of EPBII from PBII (double-headed arrow) and the probability density function (fading ellipse) from the Kriging models. PBII is defined as the difference of PBI values between a certain point F* (cross) and a projected point derived from sample points (dashed lines are the PBI contours). Three types of shaded sector forms represent each cluster of weight vectors.

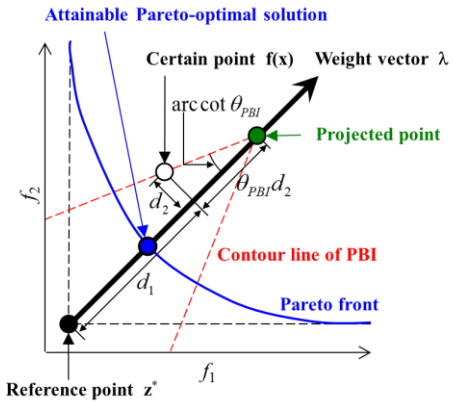



Fig. 5. Definition of the PBI function. A utopia point is used as a reference point for the PBI computation. One of the Pareto-optimal solution can be attained by minimizing the PBI function along the weight vector.

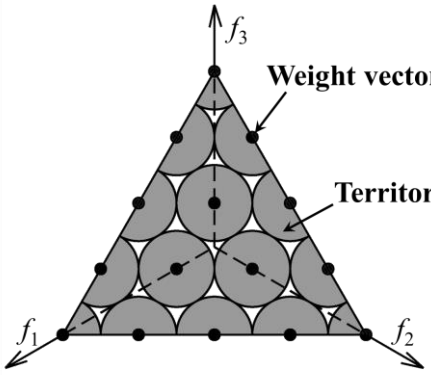



Fig. 6. Simplified illustration of territories for each weight vectors on hyper plane in three-objective space ($H$ = 4, $M$ = 3). Each territory forms a circular cone whose top is at the reference point (coordinate origin in this figure). The radius of the circular cone is proportional to the distance from the reference point (radii in this figure are not accurate).

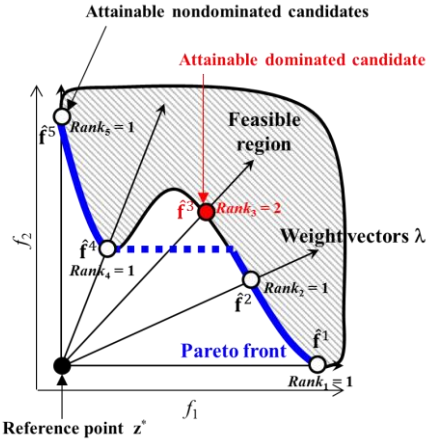



Fig. 7. Pareto ranking among candidates for a problem with discontinuous Pareto front in the EPBII approach. The gray shaded area represents feasible region. To nondominated candidates is assigned rank 1 and to a dominated candidate is assigned rank 2.

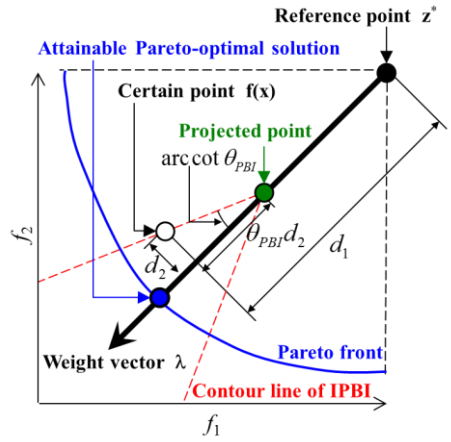



Fig. 8. Definition of the IPBI function. A nadir point is used as a reference point for the IPBI computation. One of the Pareto-optimal solutions can be attained by maximizing the IPBI function along the weight vector.

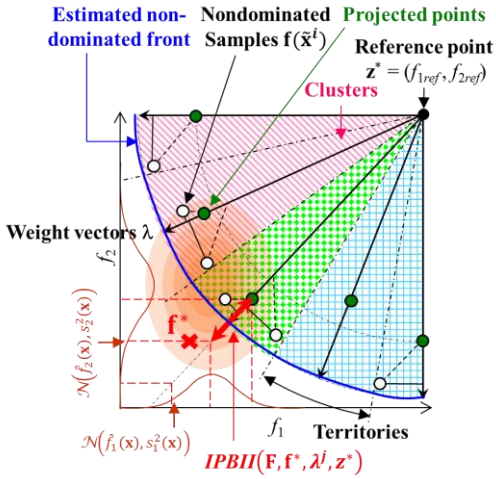
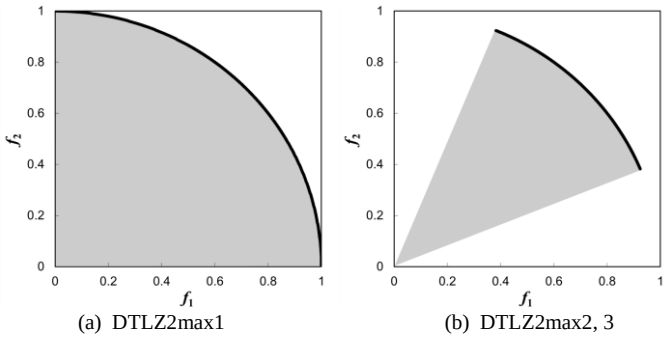



Fig. 9. Definition of EIPBII from IPBII (double-headed arrow) and the probability density function (fading ellipse) from the Kriging models. IPBII is defined as the difference of IPBI values between a certain point F* (cross) and a projected point derived from sample points (dashed lines are the IPBI contour). Three types of shaded sector forms represent each cluster of weight vectors.




(a) DTLZ2max1 (b) DTLZ2max2, 3

Fig. 10. Pareto fronts (black line) and feasible regions (gray zone) of the two-objective DTLZ2max1-3 problems.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2017.2693320, IEEE Transactions on Evolutionary Computation



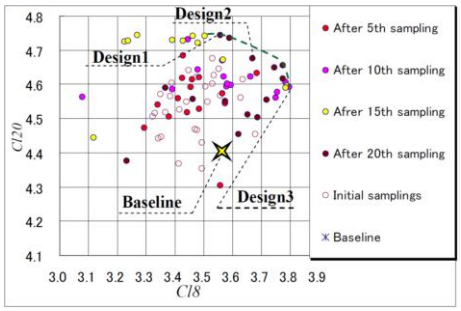



Fig. 11. Pareto front of a real-world problem: two-objective maximization of lift coefficients of a high-lift airfoil at two different angles of attack [41]. Circles are solutions.

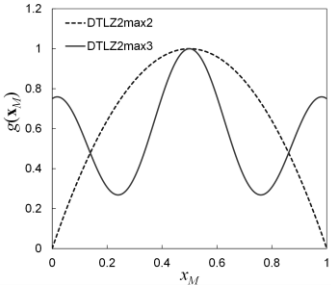
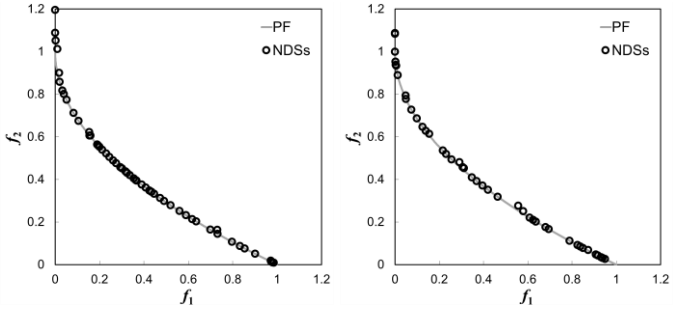
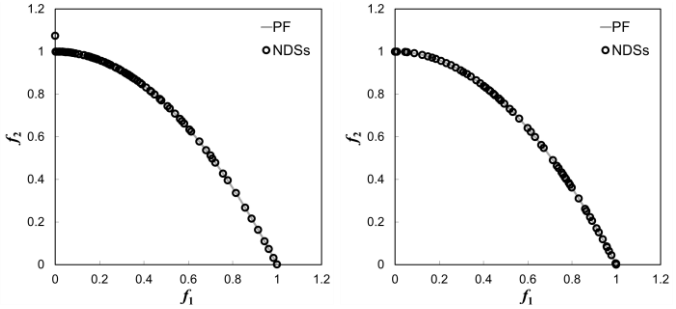



Fig. 12. Difference of $g(\mathbf{x}_M)$ functions between DTLZ2max2 and DTLZ2max3. Local maxima generate a large number of local Pareto fronts in DTLZ2max3 problems.



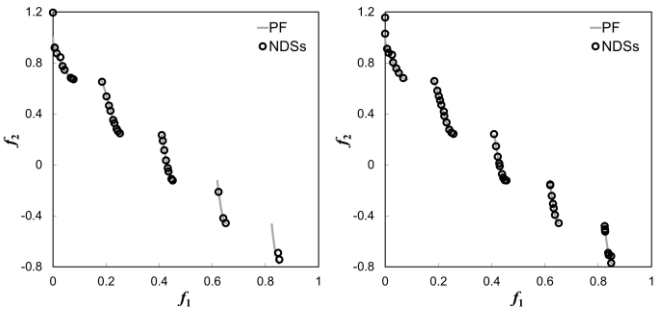

(a) EPBII (b) EIPBII

Fig. 13. Pareto front (gray line) of ZDT1 and nondominated solutions (black circles) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



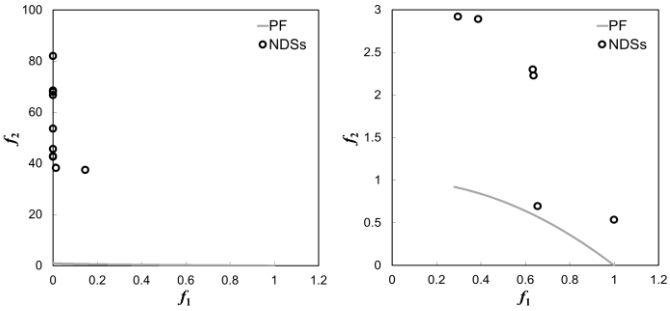

(a) EPBII (b) EIPBII

Fig. 14. Pareto front (gray line) of ZDT2 and nondominated solutions (black circles) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



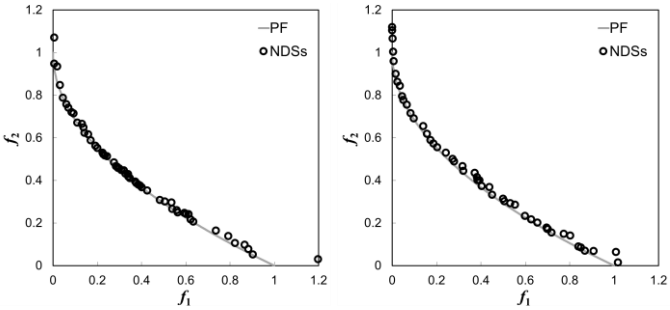

(a) EPBII (b) EIPBII

Fig. 15. Pareto front (PF) of ZDT3 and nondominated solutions (NDSs) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



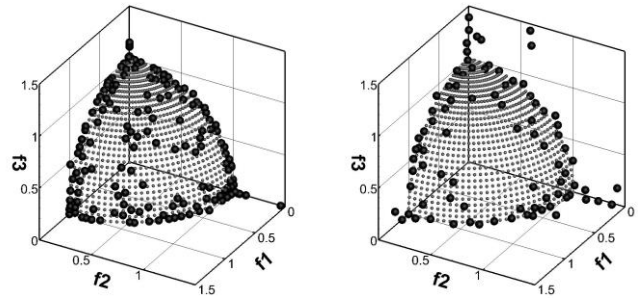

(a) EPBII for ZDT4 (b) EIPBII for ZDT6

Fig. 16. Pareto front (gray line) of ZDT4 and 6 and nondominated solutions (black circles) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII, respectively.



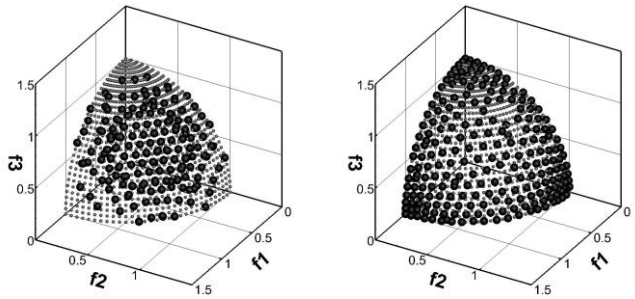

(a) EPBII (b) EIPBII

Fig. 17. Pareto front (gray line) of LZ08-F2 and nondominated solutions (black circles) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



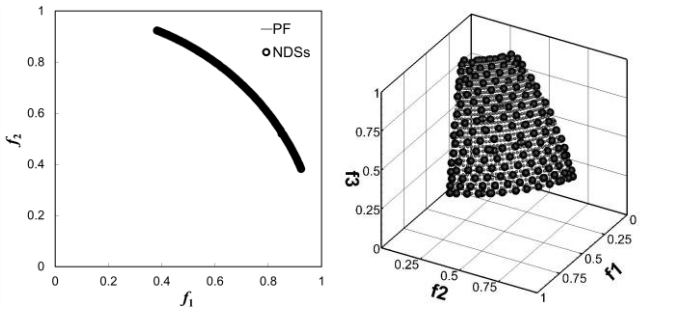

(a) EPBII (b) EIPBII

Fig. 18. Pareto optimal solutions (gray dots) of the three-objective DTLZ2 and nondominated solutions (black dots) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



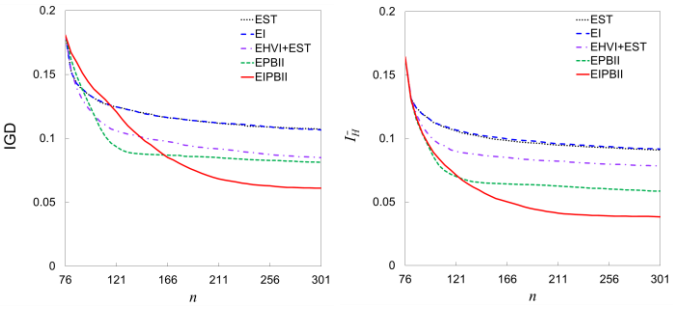

(a) EPBII (b) EIPBII

Fig. 19. Pareto optimal solutions (gray dots) of the three-objective DTLZ2max1 and nondominated solutions (black dots) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.



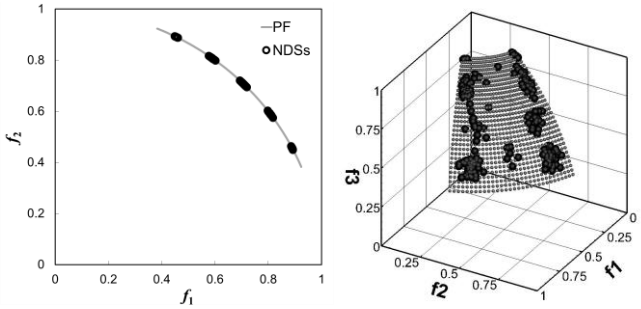

(a) Two-objective DTLZ2max2 (b) Three-objective DTLZ2max2

Fig. 20. Nondominated solutions (black circles and dots) with $I_H^-$ closest to the mean value for the two-objective and three-objective DTLZ2max2 problems obtained by EIPBII. Gray line and dots are the Pareto front and Pareto optimal solutions, respectively.



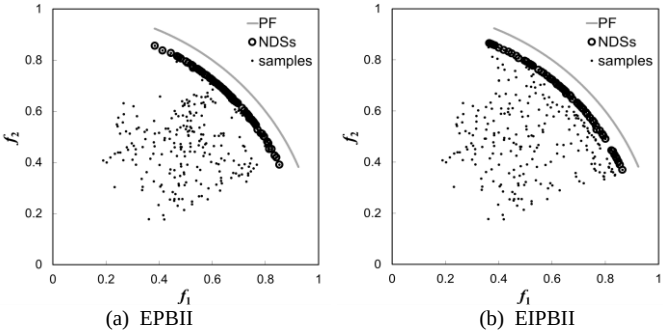

(a) IGD (b) $I_H^-$

Fig. 21. Mean value of performance metrics obtained by five criteria against number of sample points for the four-objective DTLZ2max2.



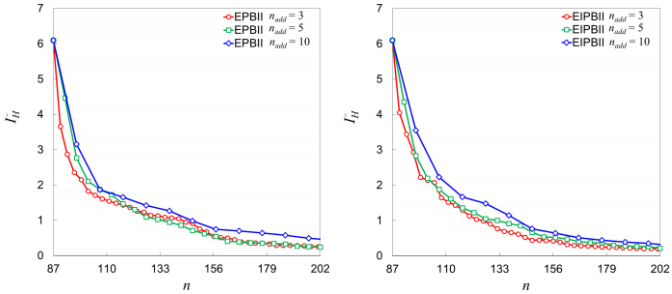

(a) EI (b) EHVI+EST

Fig. 22. Nondominated solutions (black circles and dots) with $I_H^-$ closest to the mean value for the two-objective and three-objective DTLZ2max2 problems obtained by EI and EHVI+EST, respectively. Gray line and dots are the Pareto front and Pareto optimal solutions, respectively.



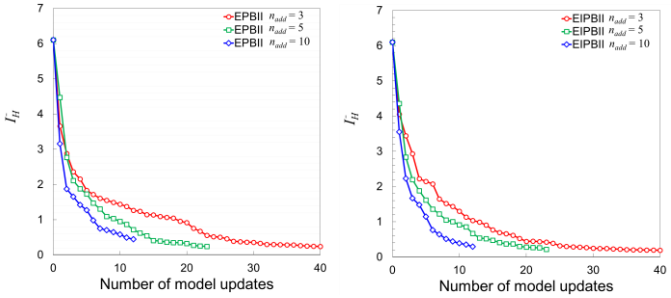

(a) EPBII (b) EIPBII

Fig. 23. Pareto front (gray line) of the two-objective DTLZ2max3 ($k$ = 8), all sample points (black dots), and nondominated solutions (black circles) with $I_H^-$ closest to the mean value obtained by EPBII and EIPBII.




(a) EPBII (b) EIPBII

Fig. 24. Mean value of $I_H^-$ obtained by EPBII and EIPBII with $n_{add}$ = 3, 5, and 10 against number of sample points for LZ08-F2.




(a) EPBII (b) EIPBII

Fig. 25. Mean value of $I_H^-$ obtained by EPBII and EIPBII with $n_{add}$ = 3, 5, and 10 against number of model updates for LZ08-F2.

TABLE I

MEAN VALUE OF $I_H^-$ FOR WELL-KNOWN TEST PROBLEMS OBTAINED BY SIX CRITERIA AND MOEA/D-EGO

| Problem | *M* | *M* | *n* | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII | MOEA/D-EGO |
|---|---|---|---|---|---|---|---|---|---|---|
| ZDT1 | 2 | 8 | 202 | 0.71090 | 0.51047 | 0.38085 | **0.05643** | **0.10366** | 0.31290 | 0.0672 |
| ZDT2 | 2 | 8 | 202 | 1.25972 | 0.95688 | 0.59756 | 0.03977 | **0.02264** | **0.01041** | 0.0198 |
| ZDT3 | 2 | 8 | 202 | 3.77783 | 2.62827 | 1.57737 | **0.04588** | 0.79416 | **0.07255** | 0.1178 |
| ZDT4 | 2 | 8 | 202 | 2242.9 | 2499.7 | **1839.1** | 2161.2 | **1796.4** | 2187.0 | 1709.52 |
| ZDT6 | 2 | 8 | 202 | 17.168 | 13.204 | 14.108 | 17.574 | **11.564** | **8.2978** | 0.1000 |
| LZ08-F1 | 2 | 8 | 202 | 0.03746 | 0.35950 | **0.01502** | 0.02259 | 0.08526 | **0.02049** | 0.1547 |
| LZ08-F2 | 2 | 8 | 202 | 0.31918 | 0.39875 | **0.15701** | **0.14900** | 0.24191 | 0.21558 | 2.8784 |
| LZ08-F3 | 2 | 8 | 202 | 0.35168 | 0.59933 | **0.11341** | **0.14125** | 0.27914 | 0.25515 | 2.5061 |
| LZ08-F4 | 2 | 8 | 202 | 0.39915 | 0.56362 | **0.11313** | **0.25962** | 0.45486 | 0.28381 | 1.8971 |
| DTLZ2 | 2 | 5 | 204 | 0.01715 | 0.01573 | 0.00905 | 0.01464 | **0.00499** | **0.00579** | - |
| | 3 | 6 | 300 | 0.51940 | 0.80711 | - | 0.24526 | **0.06169** | **0.10192** | 0.4845 |

$I_H^-$ of MOEA/D-EGO are from [20].

TABLE II

MEAN VALUES OF IGD FOR WELL-KNOWN TEST PROBLEMS OBTAINED BY SIX CRITERIA AND MOEA/D-EGO

| Problem | *M* | *m* | *n* | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII | MOEA/D-EGO |
|---|---|---|---|---|---|---|---|---|---|---|
| ZDT1 | 2 | 8 | 202 | 0.02663 | **0.01772** | 0.01904 | 0.02125 | **0.01061** | 0.01853 | 0.0148 |
| ZDT2 | 2 | 8 | 202 | 0.02090 | 0.01565 | 0.02465 | 0.02299 | **0.00642** | **0.00807** | 0.0156 |
| ZDT3 | 2 | 8 | 202 | 0.05042 | 0.04606 | 0.03016 | 0.03280 | **0.01652** | **0.01305** | 0.0665 |
| ZDT4 | 2 | 8 | 202 | 47.075 | 41.549 | **36.455** | 44.163 | **36.818** | 46.397 | 33.346 |
| ZDT6 | 2 | 8 | 202 | 0.99999 | 0.65096 | 0.81383 | 1.05801 | **0.51945** | **0.30128** | 0.0585 |
| LZ08-F1 | 2 | 8 | 202 | 0.01445 | 0.02586 | **0.00509** | 0.01424 | **0.00503** | 0.00843 | 0.0100 |
| LZ08-F2 | 2 | 8 | 202 | 0.02787 | 0.02390 | **0.01384** | 0.02722 | 0.01818 | **0.01759** | 0.1704 |
| LZ08-F3 | 2 | 8 | 202 | 0.02207 | 0.02400 | **0.01000** | 0.01833 | **0.01233** | 0.01337 | 0.1488 |
| LZ08-F4 | 2 | 8 | 202 | 0.02127 | 0.02211 | **0.00917** | 0.02143 | 0.01559 | **0.01378** | 0.0844 |
| DTLZ2 | 2 | 5 | 204 | 0.00770 | 0.00807 | 0.00867 | 0.00941 | **0.00479** | **0.00493** | - |
| | 3 | 6 | 300 | 0.07725 | **0.07476** | - | 0.07601 | **0.05680** | 0.09016 | 0.1398 |

IGD of MOEA/D-EGO are from [20].

TABLE III

MEAN VALUES OF $I_H^-$ FOR DTLZ2MAX1-3 PROBLEMS OBTAINED BY SIX CRITERIA

| Problem | *M* | *m* | *n* | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII |
|---|---|---|---|---|---|---|---|---|---|
| DTLZ2max1 | 2 | 5 | 204 | 0.05361 | 0.05029 | **0.00405** | 0.01514 | 0.00975 | **0.00394** |
| | 3 | 6 | 300 | 0.07207 | 0.07135 | - | 0.06583 | **0.04794** | **0.04660** |
| DTLZ2max2 | 2 | 5 | 204 | 0.04044 | 0.04165 | **0.00432** | 0.00768 | 0.00534 | **0.00148** |
| | 3 | 6 | 300 | 0.07555 | 0.07538 | - | 0.06323 | **0.03349** | **0.01818** |
| | 4 | 7 | 301 | 0.09099 | 0.09168 | - | 0.07848 | **0.05857** | **0.03843** |
| DTLZ2max3 | 2 | 5 | 204 | 0.07193 | 0.06156 | **0.03574** | 0.04219 | 0.05879 | **0.03907** |
| | | 9 | 303 | 0.14061 | 0.15107 | 0.20781 | 0.12709 | **0.10331** | **0.09340** |
| | 3 | 6 | 300 | 0.09754 | 0.08488 | - | 0.08619 | **0.04506** | **0.03151** |
| | | 10 | 304 | 0.19321 | 0.18676 | - | 0.21288 | **0.14011** | **0.11778** |
| | 4 | 7 | 301 | 0.09384 | 0.08885 | - | 0.08140 | **0.06330** | **0.04593** |
| | | 11 | 300 | 0.15037 | 0.15251 | - | 0.14661 | **0.11474** | **0.14049** |

TABLE IV

MEAN VALUES OF IGD FOR DTLZ2MAX1-3 PROBLEMS OBTAINED BY SIX CRITERIA

| Problem | *M* | *m* | *n* | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII |
|---|---|---|---|---|---|---|---|---|---|
| DTLZ2max1 | 2 | 5 | 204 | 0.04907 | 0.04634 | **0.00378** | 0.00941 | 0.00501 | **0.00399** |
| | 3 | 6 | 300 | 0.09393 | 0.09372 | - | 0.08295 | **0.06409** | **0.03351** |
| DTLZ2max2 | 2 | 5 | 204 | 0.02725 | 0.02801 | **0.00219** | 0.00712 | 0.00322 | **0.00191** |
| | 3 | 6 | 300 | 0.05022 | 0.04991 | - | 0.04607 | **0.02490** | **0.02170** |
| | 4 | 7 | 301 | 0.10720 | 0.10667 | - | 0.08503 | **0.08139** | **0.06117** |
| DTLZ2max3 | 2 | 5 | 204 | 0.04269 | 0.03564 | **0.02037** | 0.02676 | 0.03813 | **0.02613** |
| | | 9 | 303 | 0.07865 | 0.08833 | 0.13686 | 0.07674 | **0.06239** | **0.06024** |
| | 3 | 6 | 300 | 0.05963 | 0.05346 | - | 0.05343 | **0.03575** | **0.02931** |
| | | 10 | 304 | 0.13082 | 0.12417 | - | 0.14535 | **0.08893** | **0.08254** |
| | 4 | 7 | 301 | 0.10494 | 0.10268 | - | 0.08651 | **0.08196** | **0.06721** |
| | | 11 | 300 | 0.17344 | 0.17973 | - | **0.17149** | **0.13801** | 0.17256 |

## APPENDIX A

STATISTICS OF $I_H^-$ FOR WELL-KNOWN TEST PROBLEMS OBTAINED BY SIX CRITERIA AND MOEA/D-EGO

| Problem | *M* | *m* | *n* | Statistics | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII | MOEA/D-EGO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ZDT1 | 2 | 8 | 202 | Mean | 0.71090 | 0.51047 | 0.38085 | **0.05643** | **0.10366** | 0.31290 | 0.0672 |
| | | | | Std. dev. | 0.24491 | 0.20471 | 0.18581 | **0.04760** | **0.09492** | 0.42415 | 0.0801 |
| | | | | Best | 0.46448 | 0.09792 | 0.17487 | **0.02407** | 0.03391 | **0.01716** | 0.0182 |
| | | | | Worst | 1.41447 | 0.86532 | 0.81559 | **0.18762** | **0.37726** | 1.18376 | - |
| ZDT2 | 2 | 8 | 202 | Mean | 1.25972 | 0.95688 | 0.59756 | 0.03977 | **0.02264** | **0.01041** | 0.0198 |
| | | | | Std. dev. | 0.14796 | 0.22236 | 0.18650 | **0.00947** | 0.02805 | **0.00096** | 0.0035 |
| | | | | Best | 0.89187 | 0.43533 | 0.46434 | 0.02432 | **0.01033** | **0.00971** | 0.0160 |
| | | | | Worst | 1.44285 | 1.30311 | 1.00432 | **0.05424** | 0.10647 | **0.01309** | - |
| ZDT3 | 2 | 8 | 202 | Mean | 3.77783 | 2.62827 | 1.57737 | **0.04588** | 0.79416 | **0.07255** | 0.1178 |
| | | | | Std. dev. | 2.24293 | 1.26404 | 1.49953 | **0.02483** | 1.79163 | **0.06694** | 0.0707 |
| | | | | Best | 1.47564 | 1.46608 | 0.56530 | 0.02257 | **0.01516** | **0.01082** | 0.0308 |
| | | | | Worst | 7.62204 | 6.17488 | 5.65501 | **0.09626** | 6.14987 | **0.20163** | - |
| ZDT4 | 2 | 8 | 202 | Mean | 2242.9 | 2499.7 | **1839.1** | 2161.2 | **1796.4** | 2187.0 | 1709.5 |
| | | | | Std. dev. | 388.11 | - | 436.72 | **353.36** | 581.62 | **348.10** | 593.51 |
| | | | | Best | 1207.9 | 2499.7 | **998.29** | 1448.2 | **780.41** | 1491.8 | 785.08 |
| | | | | Worst | **2499.7** | **2499.7** | **2499.7** | **2499.7** | **2499.7** | **2499.7** | - |
| ZDT6 | 2 | 8 | 202 | Mean | 17.168 | 13.204 | 14.108 | 17.574 | **11.564** | **8.2978** | 0.1000 |
| | | | | Std. dev. | 2.8771 | **2.1173** | 4.2048 | 3.6530 | 4.2334 | **1.6404** | 0.1009 |
| | | | | Best | 13.426 | 9.6010 | **4.8966** | 12.550 | 7.3899 | **5.1129** | 0.0386 |
| | | | | Worst | 21.913 | **15.351** | 18.767 | 21.690 | 22.196 | **10.938** | - |
| LZ08-F1 | 2 | 8 | 202 | Mean | 0.03746 | 0.35950 | **0.01502** | 0.02259 | 0.08526 | **0.02049** | 0.1547 |
| | | | | Std. dev. | 0.01508 | 0.27491 | **0.00506** | **0.00392** | 0.01803 | 0.00660 | 0.0405 |
| | | | | Best | 0.02420 | 0.03927 | **0.00950** | 0.01731 | 0.05179 | **0.01312** | 0.0993 |
| | | | | Worst | 0.07706 | 0.86627 | **0.02415** | **0.02961** | 0.10730 | 0.03625 | - |
| LZ08-F2 | 2 | 8 | 202 | Mean | 0.31918 | 0.39875 | **0.15701** | **0.14900** | 0.24191 | 0.21558 | 2.8784 |
| | | | | Std. dev. | 0.26634 | 0.22686 | **0.10884** | **0.12552** | 0.18294 | 0.17162 | 0.6833 |
| | | | | Best | 0.06810 | 0.11597 | **0.02207** | **0.04608** | 0.07754 | 0.06162 | 1.8166 |
| | | | | Worst | 1.01001 | 0.69917 | **0.39777** | **0.46432** | 0.56247 | 0.57771 | - |
| LZ08-F3 | 2 | 8 | 202 | Mean | 0.35168 | 0.59933 | **0.11341** | **0.14125** | 0.27914 | 0.25515 | 2.5061 |
| | | | | Std. dev. | 0.16092 | 0.23700 | **0.07818** | **0.09602** | 0.15878 | 0.17808 | 0.4284 |
| | | | | Best | 0.08714 | 0.26480 | **0.01931** | 0.05172 | 0.10476 | **0.02611** | 1.5683 |
| | | | | Worst | 0.73883 | 1.00676 | **0.27958** | **0.37602** | 0.58320 | 0.58058 | - |
| LZ08-F4 | 2 | 8 | 202 | Mean | 0.39915 | 0.56362 | **0.11313** | **0.25962** | 0.45486 | 0.28381 | 1.8971 |
| | | | | Std. dev. | 0.16807 | 0.19416 | **0.07718** | 0.19451 | 0.31402 | **0.10855** | 0.6182 |
| | | | | Best | 0.13340 | 0.19085 | **0.04658** | **0.04992** | 0.05443 | 0.09137 | 0.7402 |
| | | | | Worst | 0.69103 | 0.88509 | **0.32265** | 0.60443 | 1.16507 | **0.42972** | - |
| DTLZ2 | 2 | 5 | 204 | Mean | 0.01715 | 0.01573 | 0.00905 | 0.01464 | **0.00499** | **0.00579** | - |
| | | | | Std. dev. | 0.00268 | 0.00183 | 0.00056 | 0.00200 | **0.00021** | **0.00038** | - |
| | | | | Best | 0.01388 | 0.01141 | 0.00779 | 0.01126 | **0.00477** | **0.00519** | - |
| | | | | Worst | 0.02289 | 0.01824 | 0.00984 | 0.01842 | **0.00542** | **0.00636** | - |
| | 3 | 6 | 300 | Mean | 0.51940 | 0.80711 | - | 0.24526 | **0.06169** | **0.10192** | 0.4845 |
| | | | | Std. dev. | 0.26354 | 0.48157 | - | 0.09786 | **0.00203** | **0.00645** | 0.1244 |
| | | | | Best | 0.20566 | 0.26746 | - | 0.14181 | **0.05939** | **0.09414** | 0.2801 |
| | | | | Worst | 1.04990 | 1.71934 | - | 0.49835 | **0.06586** | **0.11330** | - |

Statistics of MOEA/D-EGO are from [20].

APPENDIX B
STATISTICS OF IGD FOR WELL-KNOWN TEST PROBLEMS OBTAINED BY THE SIX CRITERIA AND MOEA/D-EGO

| Problem | *M* | *m* | *n* | Statistics | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII | MOEA/D-EGO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ZDT1 | 2 | 8 | 202 | Mean | 0.02663 | **0.01772** | 0.01904 | 0.02125 | **0.01061** | 0.01853 | 0.0148 |
| | | | | Std. dev. | 0.01089 | 0.00466 | 0.00462 | **0.00422** | **0.00197** | 0.00915 | 0.0016 |
| | | | | Best | 0.01682 | 0.01089 | 0.01447 | 0.01605 | **0.00750** | **0.01035** | 0.0120 |
| | | | | Worst | 0.05639 | **0.02692** | 0.02890 | 0.03132 | **0.01421** | 0.03909 | - |
| ZDT2 | 2 | 8 | 202 | Mean | 0.02090 | 0.01565 | 0.02465 | 0.02299 | **0.00642** | **0.00807** | 0.0156 |
| | | | | Std. dev. | 0.00252 | 0.00350 | 0.00121 | 0.00465 | **0.00115** | **0.00062** | 0.0022 |
| | | | | Best | 0.01754 | 0.01185 | 0.02174 | 0.01552 | **0.00509** | **0.00742** | 0.0121 |
| | | | | Worst | 0.02421 | 0.02324 | 0.02636 | 0.03175 | **0.00838** | **0.00976** | - |
| ZDT3 | 2 | 8 | 202 | Mean | 0.05042 | 0.04606 | 0.03016 | 0.03280 | **0.01652** | **0.01305** | 0.0665 |
| | | | | Std. dev. | 0.01104 | 0.01516 | 0.00990 | **0.00667** | 0.00775 | **0.00133** | 0.0189 |
| | | | | Best | 0.03323 | 0.02659 | 0.02245 | 0.02409 | **0.01128** | **0.01119** | 0.0435 |
| | | | | Worst | 0.06997 | 0.07766 | 0.05809 | 0.04572 | **0.03918** | **0.01493** | - |
| ZDT4 | 2 | 8 | 202 | Mean | 47.075 | 41.549 | **36.455** | 44.163 | **36.818** | 46.397 | 33.346 |
| | | | | Std. dev. | 10.020 | 9.1344 | **8.8453** | **8.7994** | 13.408 | 9.9592 | 12.943 |
| | | | | Best | 23.837 | 26.393 | **19.646** | 28.359 | **15.291** | 29.514 | 14.869 |
| | | | | Worst | 59.266 | 59.989 | **50.172** | 59.714 | 60.430 | **59.110** | - |
| ZDT6 | 2 | 8 | 202 | Mean | 0.99999 | 0.65096 | 0.81383 | 1.05801 | **0.51945** | **0.30128** | 0.0585 |
| | | | | Std. dev. | 0.30145 | **0.17310** | 0.21044 | 0.36357 | 0.39054 | **0.06990** | 0.0421 |
| | | | | Best | 0.61401 | 0.39724 | 0.44875 | 0.60079 | **0.25074** | **0.22190** | 0.0273 |
| | | | | Worst | 1.49281 | **0.86181** | 1.12421 | 1.48168 | 1.53050 | **0.43136** | - |
| LZ08-F1 | 2 | 8 | 202 | Mean | 0.01445 | 0.02586 | **0.00509** | 0.01424 | **0.00503** | 0.00843 | 0.0100 |
| | | | | Std. dev. | 0.00317 | 0.00739 | 0.00090 | 0.00302 | **0.00022** | **0.00087** | 0.0011 |
| | | | | Best | 0.00953 | 0.01597 | **0.00410** | 0.01087 | **0.00462** | 0.00696 | 0.0083 |
| | | | | Worst | 0.02145 | 0.03749 | **0.00725** | 0.02042 | **0.00549** | 0.00935 | - |
| LZ08-F2 | 2 | 8 | 202 | Mean | 0.02787 | 0.02390 | **0.01384** | 0.02722 | 0.01818 | **0.01759** | 0.1704 |
| | | | | Std. dev. | 0.00607 | 0.00436 | 0.00517 | **0.00359** | **0.00185** | 0.00401 | 0.0241 |
| | | | | Best | 0.02183 | 0.01791 | **0.00958** | 0.02250 | 0.01510 | **0.01118** | 0.1297 |
| | | | | Worst | 0.03890 | 0.03182 | **0.02633** | 0.03525 | **0.02066** | 0.02659 | - |
| LZ08-F3 | 2 | 8 | 202 | Mean | 0.02207 | 0.02400 | **0.01000** | 0.01833 | **0.01233** | 0.01337 | 0.1488 |
| | | | | Std. dev. | 0.00344 | 0.00631 | **0.00210** | 0.00250 | **0.00195** | 0.00231 | 0.0297 |
| | | | | Best | 0.01706 | 0.01469 | **0.00706** | 0.01419 | **0.00962** | 0.01029 | 0.1010 |
| | | | | Worst | 0.02728 | 0.03967 | **0.01388** | 0.02360 | **0.01534** | 0.01686 | - |
| LZ08-F4 | 2 | 8 | 202 | Mean | 0.02127 | 0.02211 | **0.00917** | 0.02143 | 0.01559 | **0.01378** | 0.0844 |
| | | | | Std. dev. | 0.00450 | 0.00497 | **0.00087** | 0.00376 | 0.00519 | **0.00291** | 0.0144 |
| | | | | Best | 0.01614 | 0.01788 | **0.00723** | 0.01460 | 0.01086 | **0.00990** | 0.0659 |
| | | | | Worst | 0.03142 | 0.03281 | **0.01044** | 0.02706 | 0.02947 | **0.01881** | - |
| DTLZ2 | 2 | 5 | 204 | Mean | 0.00770 | 0.00807 | 0.00867 | 0.00941 | **0.00479** | **0.00493** | - |
| | | | | Std. dev. | 0.00080 | 0.00142 | 0.00058 | 0.00086 | **0.00024** | **0.00018** | - |
| | | | | Best | 0.00680 | 0.00653 | 0.00722 | 0.00831 | **0.00457** | **0.00469** | - |
| | | | | Worst | 0.00961 | 0.01016 | 0.00965 | 0.01128 | **0.00540** | **0.00527** | - |
| | 3 | 6 | 300 | Mean | 0.07725 | **0.07476** | - | 0.07601 | **0.05680** | 0.09016 | 0.1398 |
| | | | | Std. dev. | 0.00464 | **0.00379** | - | 0.00457 | **0.00195** | 0.00549 | 0.0085 |
| | | | | Best | 0.07205 | **0.06912** | - | 0.07016 | **0.05368** | 0.07897 | 0.1306 |
| | | | | Worst | 0.08613 | **0.08157** | - | 0.08421 | **0.05902** | 0.09784 | - |

Statistics of MOEA/D-EGO are from [20].

## APPENDIX C

STATISTICS OF $I_H^-$ FOR DTLZ2MAX1-3 PROBLEMS OBTAINED BY THE SIX CRITERIA

| Problem | *M* | *m* | *n* | Statistics | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII |
|---|---|---|---|---|---|---|---|---|---|---|
| DTLZ2max1 | 2 | 5 | 204 | Mean | 0.05361 | 0.05029 | **0.00405** | 0.01514 | 0.00975 | **0.00394** |
| | | | | Std. dev. | 0.00534 | 0.00696 | **0.00022** | 0.00160 | 0.00134 | **0.00010** |
| | | | | Best | 0.04562 | 0.03526 | **0.00370** | 0.01287 | 0.00815 | **0.00378** |
| | | | | Worst | 0.06437 | 0.05837 | **0.00445** | 0.01752 | 0.01219 | **0.00411** |
| | 3 | 6 | 300 | Mean | 0.07207 | 0.07135 | - | 0.06583 | **0.04794** | **0.04660** |
| | | | | Std. dev. | 0.00257 | 0.00199 | - | 0.00241 | **0.00036** | **0.00134** |
| | | | | Best | 0.06821 | 0.06732 | - | 0.06175 | **0.04753** | **0.04453** |
| | | | | Worst | 0.07671 | 0.07414 | - | 0.06996 | **0.04843** | **0.04917** |
| DTLZ2max2 | 2 | 5 | 204 | Mean | 0.04044 | 0.04165 | **0.00432** | 0.00768 | 0.00534 | **0.00148** |
| | | | | Std. dev. | 0.00160 | 0.00126 | 0.00029 | 0.00323 | **0.00015** | **0.00027** |
| | | | | Best | 0.03799 | 0.03957 | **0.00376** | 0.00414 | 0.00507 | **0.00135** |
| | | | | Worst | 0.04356 | 0.04371 | **0.00476** | 0.01277 | 0.00565 | **0.00227** |
| | 3 | 6 | 300 | Mean | 0.07555 | 0.07538 | - | 0.06323 | **0.03349** | **0.01818** |
| | | | | Std. dev. | 0.00157 | 0.00206 | - | 0.00590 | **0.00031** | **0.00026** |
| | | | | Best | 0.07257 | 0.07285 | - | 0.05311 | **0.03290** | **0.01773** |
| | | | | Worst | 0.07835 | 0.07936 | - | 0.07074 | **0.03400** | **0.01867** |
| | 4 | 7 | 301 | Mean | 0.09099 | 0.09168 | - | 0.07848 | **0.05857** | **0.03843** |
| | | | | Std. dev. | 0.00205 | 0.00214 | - | 0.00343 | **0.00064** | **0.00060** |
| | | | | Best | 0.08835 | 0.08788 | - | 0.07410 | **0.05762** | **0.03712** |
| | | | | Worst | 0.09464 | 0.09431 | - | 0.08489 | **0.05963** | **0.03947** |
| DTLZ2max3 | 2 | 5 | 204 | Mean | 0.07193 | 0.06156 | **0.03574** | 0.04219 | 0.05879 | **0.03907** |
| | | | | Std. dev. | 0.03729 | **0.03589** | **0.03476** | 0.03926 | 0.04132 | 0.04157 |
| | | | | Best | 0.03720 | 0.03444 | 0.01046 | **0.00371** | 0.00544 | **0.00218** |
| | | | | Worst | 0.12295 | 0.12170 | 0.10457 | 0.10208 | **0.09315** | **0.09070** |
| | | 9 | 303 | Mean | 0.14061 | 0.15107 | 0.20781 | 0.12709 | **0.10331** | **0.09340** |
| | | | | Std. dev. | **0.03952** | 0.03960 | 0.07863 | 0.04984 | 0.06301 | **0.03654** |
| | | | | Best | 0.07971 | 0.08246 | 0.02379 | 0.05951 | **0.00638** | **0.01249** |
| | | | | Worst | **0.19781** | 0.22867 | 0.27710 | 0.22660 | 0.22048 | **0.14728** |
| | 3 | 6 | 300 | Mean | 0.09754 | 0.08488 | - | 0.08619 | **0.04506** | **0.03151** |
| | | | | Std. dev. | 0.03103 | **0.02143** | - | 0.02710 | 0.02412 | **0.02275** |
| | | | | Best | 0.07034 | 0.07031 | - | 0.06509 | **0.03352** | **0.01879** |
| | | | | Worst | 0.14810 | 0.13395 | - | 0.13187 | **0.11478** | **0.09583** |
| | | 10 | 304 | Mean | 0.19321 | 0.18676 | - | 0.21288 | **0.14011** | **0.11778** |
| | | | | Std. dev. | 0.03898 | **0.03318** | - | 0.04191 | **0.02735** | 0.03570 |
| | | | | Best | 0.13116 | 0.13734 | - | 0.11462 | **0.08508** | **0.04985** |
| | | | | Worst | 0.24249 | 0.26791 | - | 0.25098 | **0.18771** | **0.19142** |
| | 4 | 7 | 301 | Mean | 0.09384 | 0.08885 | - | 0.08140 | **0.06330** | **0.04593** |
| | | | | Std. dev. | 0.01451 | **0.00711** | - | 0.01393 | **0.01162** | 0.01363 |
| | | | | Best | 0.08231 | 0.08222 | - | 0.06718 | **0.05715** | **0.04013** |
| | | | | Worst | 0.12616 | 0.10875 | - | 0.10942 | **0.09706** | **0.08672** |
| | | 11 | 300 | Mean | 0.15037 | 0.15251 | - | 0.14661 | **0.11474** | **0.14049** |
| | | | | Std. dev. | **0.01940** | **0.01566** | - | 0.01997 | 0.02309 | 0.02513 |
| | | | | Best | 0.09846 | 0.12081 | - | 0.10031 | **0.06454** | **0.09417** |
| | | | | Worst | **0.16714** | 0.16867 | - | 0.16961 | **0.15196** | 0.18174 |

APPENDIX D
STATISTICS OF IGD FOR DTLZ2MAX1-3 PROBLEMS OBTAINED BY THE SIX CRITERIA

| Problem | *M* | *m* | *n* | Statistics | EST | EI | EHVI | EHVI+EST | EPBII | EIPBII |
|---|---|---|---|---|---|---|---|---|---|---|
| DTLZ2max1 | 2 | 5 | 204 | Mean | 0.04907 | 0.04634 | **0.00378** | 0.00941 | 0.00501 | **0.00399** |
| | | | | Std. dev. | 0.00509 | 0.00658 | **0.00019** | 0.00086 | 0.00051 | **0.00009** |
| | | | | Best | 0.04180 | 0.03217 | **0.00356** | 0.00831 | 0.00445 | **0.00389** |
| | | | | Worst | 0.05896 | 0.05460 | **0.00422** | 0.01128 | 0.00579 | **0.00420** |
| | 3 | 6 | 300 | Mean | 0.09393 | 0.09372 | - | 0.08295 | **0.06409** | **0.03351** |
| | | | | Std. dev. | **0.00563** | 0.00592 | - | 0.00745 | 0.00580 | **0.00065** |
| | | | | Best | 0.08314 | 0.08258 | - | 0.07387 | **0.05428** | **0.03248** |
| | | | | Worst | 0.10264 | 0.10078 | - | 0.09752 | **0.07130** | **0.03442** |
| DTLZ2max2 | 2 | 5 | 204 | Mean | 0.02725 | 0.02801 | **0.00219** | 0.00712 | 0.00322 | **0.00191** |
| | | | | Std. dev. | 0.00164 | 0.00151 | **0.00007** | 0.00109 | 0.00013 | **0.00003** |
| | | | | Best | 0.02504 | 0.02604 | **0.00208** | 0.00492 | 0.00295 | **0.00186** |
| | | | | Worst | 0.03034 | 0.03076 | **0.00231** | 0.00852 | 0.00339 | **0.00198** |
| | 3 | 6 | 300 | Mean | 0.05022 | 0.04991 | - | 0.04607 | **0.02490** | **0.02170** |
| | | | | Std. dev. | 0.00220 | 0.00146 | - | 0.00231 | **0.00007** | **0.00010** |
| | | | | Best | 0.04601 | 0.04698 | - | 0.04357 | **0.02481** | **0.02149** |
| | | | | Worst | 0.05373 | 0.05154 | - | 0.05181 | **0.02502** | **0.02186** |
| | 4 | 7 | 301 | Mean | 0.10720 | 0.10667 | - | 0.08503 | **0.08139** | **0.06117** |
| | | | | Std. dev. | 0.00353 | 0.00329 | - | 0.00230 | **0.00097** | **0.00020** |
| | | | | Best | 0.10117 | 0.10166 | - | 0.08039 | **0.07938** | **0.06078** |
| | | | | Worst | 0.11329 | 0.11121 | - | 0.08788 | **0.08291** | **0.06150** |
| DTLZ2max3 | 2 | 5 | 204 | Mean | 0.04269 | 0.03564 | **0.02037** | 0.02676 | 0.03813 | **0.02613** |
| | | | | Std. dev. | **0.01995** | **0.02044** | 0.02165 | 0.02403 | 0.02760 | 0.02793 |
| | | | | Best | 0.02202 | 0.01734 | 0.00578 | 0.00510 | **0.00322** | **0.00195** |
| | | | | Worst | 0.06909 | 0.06797 | 0.06360 | 0.06202 | **0.06107** | **0.06032** |
| | | 9 | 303 | Mean | 0.07865 | 0.08833 | 0.13686 | 0.07674 | **0.06239** | **0.06024** |
| | | | | Std. dev. | **0.02366** | 0.02723 | 0.05807 | 0.03621 | 0.04157 | **0.02447** |
| | | | | Best | 0.04168 | 0.04168 | 0.01062 | 0.03230 | **0.00404** | **0.00670** |
| | | | | Worst | **0.11120** | 0.13791 | 0.20267 | 0.16123 | 0.14075 | **0.09654** |
| | 3 | 6 | 300 | Mean | 0.05963 | 0.05346 | - | 0.05343 | **0.03575** | **0.02931** |
| | | | | Std. dev. | 0.01594 | **0.01193** | - | 0.01608 | 0.02174 | **0.01300** |
| | | | | Best | 0.04431 | 0.04373 | - | 0.03852 | **0.02461** | **0.02197** |
| | | | | Worst | 0.08732 | 0.07916 | - | **0.07768** | 0.09661 | **0.06520** |
| | | 10 | 304 | Mean | 0.13082 | 0.12417 | - | 0.14535 | **0.08893** | **0.08254** |
| | | | | Std. dev. | 0.03094 | 0.02708 | - | 0.03472 | **0.02081** | **0.02477** |
| | | | | Best | 0.08307 | 0.09393 | - | 0.06648 | **0.04719** | **0.03956** |
| | | | | Worst | 0.17206 | 0.19500 | - | 0.18003 | **0.12394** | **0.13581** |
| | 4 | 7 | 301 | Mean | 0.10494 | 0.10268 | - | 0.08651 | **0.08196** | **0.06721** |
| | | | | Std. dev. | 0.01232 | **0.00821** | - | 0.01140 | **0.00585** | 0.00973 |
| | | | | Best | 0.08734 | 0.09142 | - | **0.07313** | 0.07640 | **0.06162** |
| | | | | Worst | 0.13323 | 0.12192 | - | 0.10832 | **0.09880** | **0.09591** |
| | | 11 | 300 | Mean | 0.17344 | 0.17973 | - | **0.17149** | **0.13801** | 0.17256 |
| | | | | Std. dev. | 0.02700 | **0.02316** | - | **0.02585** | 0.03508 | 0.03886 |
| | | | | Best | 0.11461 | 0.14201 | - | 0.12612 | **0.08317** | **0.10687** |
| | | | | Worst | 0.21402 | **0.21158** | - | 0.21243 | **0.19813** | 0.23859 |