# On Expected-Improvement Criteria for Model-based Multi-objective Optimization


Tobias Wagner[1], Michael Emmerich[2], André Deutz[2], and Wolfgang Ponweiser[3]

[1] Institute of Machining Technology (ISF)
Technische Universität Dortmund, 44227 Dortmund, Germany
wagner@isf.de

[2] Leiden Institute of Advanced Computer Science (LIACS)
Universiteit Leiden, 2333 CA Leiden, The Netherlands
{emmerich,deutz}@liacs.nl

[3] Automation and Control Institute
Vienna University of Technology, 1040 Vienna, Austria
ponweiser@acin.tuwien.ac.at



**Abstract.** Surrogate models, as used for the Design and Analysis of Computer Experiments (DACE), can significantly reduce the resources necessary in cases of expensive evaluations. They provide a prediction of the objective and of the corresponding uncertainty, which can then be combined to a figure of merit for a sequential optimization. In single-objective optimization, the expected improvement (EI) has proven to provide a combination that balances successfully between local and global search. Thus, it has recently been adapted to evolutionary multi-objective optimization (EMO) in different ways. In this paper, we provide an overview of the existing EI extensions for EMO and propose new formulations of the EI based on the hypervolume. We set up a list of necessary and desirable properties, which is used to reveal the strengths and weaknesses of the criteria by both theoretical and experimental analyses.



**Keywords:** Design and Analysis of Computer Experiments, Expected Improvement, Hypervolume Indicator, Multi-Objective Optimization.


## 1 Introduction

Surrogate modeling has become the method of choice to overcome the problem of expensive evaluations in EMO [1]. Using the evaluations already available, surrogate models of the objectives are created, which can then be used to filter or decide on candidate solutions. To accomplish this, a criterion which scalarizes the predictions of the models is required. This criterion should balance between a local refinement of the Pareto-front (PF) approximation and an improvement of the global model quality.

In this paper, such criteria for multi-objective optimization are presented, analyzed, and discussed. The main definitions are provided, existing criteria are summarized, and enhancements in the calculation of these criteria are proposed in section 2 and 3. For the evaluation of the criteria necessary requirements


R. Schaefer et al. (Eds.): PPSN XI, Part I, LNCS 6238, pp. 718–727, 2010.

and desired properties are formulated in section 4. By means of both, formal and empirical, analyses, we study whether these requirements and properties are met by the various criteria. Concluding, a summary of the results and an outlook on further research topics are provided in section 5.

## 2 Single-Objective Optimization Based on EI

A surrogate model allows the objective function value $y = f(\mathbf{x})$ of a decision vector $\mathbf{x}$ to be predicted without an expensive evaluation. This prediction is denoted as $\hat{y}$.[1] Since an evaluation is particularly worthwhile if it provides an improvement to the current state of the optimization, often the improvement $I(\hat{y}, f_{min}) = \max\{f_{min} - \hat{y}, 0\}$ obtained with respect to the best currently known objective value $f_{min}$ is maximized. Consequently, we consider minimization of the objectives. Many modeling techniques, such as the ones used in DACE [2], predict both the mean $\hat{y}$ and the standard deviation $\hat{s}$ of a normal distribution. Consequently, the probability density function (PDF) $\phi_{(\hat{y},\hat{s})}(y) = \phi_{(0,1)}(\frac{y-\hat{y}}{\hat{s}})$ and the cumulative density function (CDF) $\Phi_{(\hat{y},\hat{s})}(y) = \Phi_{(0,1)}(\frac{y-\hat{y}}{\hat{s}})$ of an objective value $y$ can be computed (cf. Fig. 1). Based on the definition of the improvement $I(y, f_{min})$ and the PDF of $y$, the expected value of the improvement

$$EI(\hat{y}, \hat{s}, f_{min}) = \int_{-\infty}^{\infty} I(y, f_{min}) \underbrace{\phi_{(\hat{y},\hat{s})}(y)}_{\mathrm{PDF}(y)} \, dy = \int_{-\infty}^{f_{min}} (f_{min} - y)\phi_{(\hat{y},\hat{s})}(y) \, dy \quad (1)$$

has been proposed as criterion by the Vilnius school of global optimization, e. g. [3]. Later, the EI has become popular as part of the single-objective Efficient Global Optimization (EGO) [4] approach.[2] EGO makes an extensive use of DACE models by only evaluating one solution in each iteration on the true objective function, refitting the model, and then determining the next candidate solution based on the EI. Since both predictions, $\hat{y}$ and $\hat{s}$, are considered, a balancing between a local search and a reduction of the model uncertainty is achieved. Thereby, the number of function evaluations could be significantly reduced for many global optimization problems – often below one hundred.

By expanding equation 1 and integrating the first factor, the EI can also be written as $f_{min}\Phi_{(\hat{y},\hat{s})}(f_{min}) - \int_{-\infty}^{f_{min}} y\phi_{(\hat{y},\hat{s})}(y) \, dy$, and thus

$$EI(\hat{y}, \hat{s}, f_{min}) = \left( f_{min} - \underbrace{\frac{\int_{-\infty}^{f_{min}} y\phi_{(\hat{y},\hat{s})}(y)dy}{\Phi_{(\hat{y},\hat{s})}(f_{min})}}_{\overline{y}} \right) \Phi_{(\hat{y},\hat{s})}(f_{min}). \quad (2)$$

[1] For notational simplicity, we omit the dependency of the predictions on $\mathbf{x}$.

[2] In the evolutionary computation community, EGO has become popular under the SPO (Sequential Parameter Optimization) acronym [5].

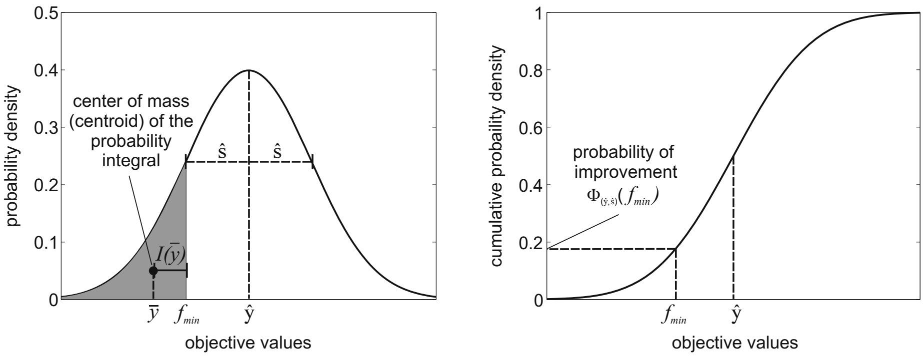


**Fig. 1.** Graphical explanation of the components $I(\overline{y}, f_{min})$ (left) and $\Phi_{(\hat{y},\hat{s})}(f_{min})$ (right) of the EI definition in equation 2

Consequently, the EI can be regarded as the improvement $I(\overline{y}, f_{min})$ obtained by the center of mass (centroid) $\overline{y}$ of the area under $\phi_{(\hat{y},\hat{s})}$ in the interval $]-\infty, f_{min}]$ weighted with the corresponding CDF $\Phi_{(\hat{y},\hat{s})}(f_{min})$ (cf. Fig. 1).

## 3 Multi-objective Optimization Based on EI

Over the last decade, a set-based view on multi-objective optimization has been established [6]. According to equation 1, a true multi-objective formulation $EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF})$ requires the PDF of $\mathbf{y}$ and a definition of the improvement $I(\mathbf{y}, \mathbf{A}_{PF})$ of the PF approximation $\mathbf{A}_{PF}$ obtained by a specific candidate vector $\mathbf{y}$. Despite the usually conflicting objectives in EMO, it is common practice [7,8,9,10,11] to make the independence assumption (correlation coefficient $\rho = 0$). Then, the multivariate PDF of $\mathbf{y}$ as $\prod_{i=1}^{m} \phi_{(\hat{y}_i,\hat{s}_i)}(y_i)$ can be directly computed, and the important aspect is the design of an appropriate improvement function.

An overview of recently proposed multi-objective EI definitions is given in Table 1. The acronyms introduced in this table are used throughout the paper. Unfortunately, we cannot describe the approaches due to space requirements. Detailed explanations will be found in the references given in Table 1.

Most of the presented approaches do not directly define a set-based improvement $I(\mathbf{y}, \mathbf{A}_{PF})$. Emmerich [7] proposed SExI – the expected increment of $\mathbf{y}$ to the hypervolume (HV) or $\mathcal{S}$-metric. The HV is the Lebesgue measure of the hyperspace dominated by $\mathbf{A}_{PF}$ and bounded by a reference point $\mathbf{r}$. A closed-form expression for SExI is based on integration over interval boxes determined by the coordinates of the points in $\mathbf{A}_{PF}$ [10]. Independently, Emmerich [7] and Ponweiser et al. [14] have proposed a EI criterion, whose computation is simpler. This measure is the increment of the hypervolume when $\mathbf{y}_{LCB} = \hat{\mathbf{y}} - \alpha\hat{\mathbf{s}}$ (lower confidence bound) is added to $\mathbf{A}_{PF}$. The gain factor $\alpha$ is computed based on a given probability level $p$ as $\alpha(p) = -\Phi^{-1}(0.5 \sqrt[m]{p})$ (in this study $p = 0.5$ is used).

**Table 1.** Overview of existing multi-objective EI criteria

| authors (reference) | acronym | definition of improvement | PDF | direct integration |
|---|---|---|---|---|
| Knowles [12] | ParEGO | single-objective EI of an augmented Tchebycheff aggregation | univariate | yes |
| Jeong and Obayashi [13] | EI-EMO | $m$ single-objective EIs | univariate | yes |
| Liu et al. [9] | WS-EI | sum over single-objective EIs of different weighted sums (WS) | multivariate | partially (only subproblems) |
| Zhang et al. [11] | TA-EI | maximum over single-objective EIs of different Tchebycheff aggregations (TA) | multivariate | partially (only subproblems) |
| Keane [8] | Euclid | Euclidean distance to the nearest vector of the PF | multivariate | partially (only PDF) |
| Ponweiser et al. [14] | SMS-EGO | HV increment to the PF | multivariate | no |
| Emmerich et al. [7,10] | SExI | HV increment to the PF | multivariate | yes |

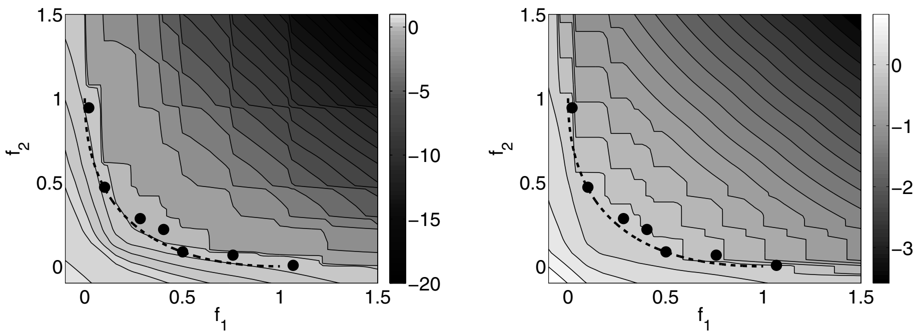


**Fig. 2.** Comparison of the old (left) and new (right) variant of SMS-EGO. The details of the calculation of the figure is described in section 4.

In order to guide search in dominated regions of the objective space, Ponweiser et al. [14] augmented this criterion by a penalty. In this paper, we introduce a new definition of this penalty. Still, a set of penalties for the $\varepsilon$-dominating solutions $\mathbf{y}^{(i)} \in \mathbf{A}_{PF}$ is computed

$$\Psi(\mathbf{y}_{LCB}) = \begin{cases} -1 + \prod_{j=1}^{m} \left(1 + (y_{LCB,j} - y_j^{(i)})\right) & \text{if } \mathbf{y}^{(i)} \preceq_\varepsilon \mathbf{y}_{LCB} \\ 0 & \text{otherwise} \end{cases} .$$

Whereas we computed the sum over all penalties $\sum \Psi$ in the old version, which resulted in discontinuities of the criterion whenever a dominating solution enters or drops out, we take only the maximum component of $\Psi$ in the new one. This modification leads to a continuous global trend toward $\mathbf{A}_{PF}$ (cf. Fig 2).

## 4 Analysis and Evaluation

For a formally sound evaluation of multi-objective EI criteria, we propose the following necessary conditions. Given two different predictions of mean vectors $\hat{\mathbf{y}}$ and $\hat{\mathbf{y}}'$ and corresponding uncertainties $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}'$,

N1 the dominance relation between $\hat{\mathbf{y}}$ and $\hat{\mathbf{y}}'$ is preserved by the EI for $\hat{\mathbf{s}} = \hat{\mathbf{s}}'$:
$\hat{\mathbf{y}} \prec \hat{\mathbf{y}}' \wedge \hat{\mathbf{s}} = \hat{\mathbf{s}}' \Rightarrow EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF}) > EI(\hat{\mathbf{y}}', \hat{\mathbf{s}}', \mathbf{A}_{PF})$,
N2 the EI monotonically increases with $\hat{\mathbf{s}}$ for $I(\hat{\mathbf{y}}, \mathbf{A}_{PF}) \leq 0$ and $\hat{\mathbf{y}} = \hat{\mathbf{y}}'$:
$I(\hat{\mathbf{y}}, \mathbf{A}_{PF}) \leq 0 \wedge \hat{\mathbf{y}} = \hat{\mathbf{y}}' \wedge \hat{\mathbf{s}} > \hat{\mathbf{s}}' \Rightarrow EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF}) > EI(\hat{\mathbf{y}}', \hat{\mathbf{s}}', \mathbf{A}_{PF})$,
N3 the EI monotonically increases with $I(\hat{\mathbf{y}}, \mathbf{A}_{PF})$ for $\hat{\mathbf{s}} = \hat{\mathbf{s}}' = \mathbf{0}$:
$I(\hat{\mathbf{y}}, \mathbf{A}_{PF}) > I(\hat{\mathbf{y}}', \mathbf{A}_{PF}) \wedge \hat{\mathbf{s}} = \hat{\mathbf{s}}' = \mathbf{0} \Rightarrow EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF}) > EI(\hat{\mathbf{y}}', \hat{\mathbf{s}}', \mathbf{A}_{PF})$.

The necessary conditions can be analytically checked in most cases and should be considered during the design of a multi-objective EI criterion in order to identify conceptual problems. The restriction to solutions with no improvement in condition N2 and vanishing uncertainties in N3 was made since an increase in $\hat{\mathbf{s}}$ is related to a balancing between risk and opportunity for $I(\hat{\mathbf{y}}, \mathbf{A}_{PF}) > 0$.

Moreover, we compiled a second list, which includes properties that are desired with respect to the internal optimization:

D1 For small $\hat{\mathbf{s}}$ in relation to the range of $\mathbf{A}_{PF}$, a solution should be preferred whose $\hat{\mathbf{y}}$ improves the distribution and/or spread of $\mathbf{A}_{PF}$.
D2 Discontinuities and nondifferentiabilities of the criterion should be avoided, particularly if gradient-based methods are used for the internal optimization.
D3 The fitness landscape of the criterion should guide the optimizer to its global optimum, e. g., plateaus should be avoided and basin sizes should grow with the quality of the corresponding local optimum.
D4 The criterion should be easy to implement and efficient to calculate.

Since the importance of these properties depends on the internal optimization approach and on the application domain, their discussion can assist in choosing the right criterion for a given application.

An overview of the results of our analyses is provided in Table 2. Whenever possible, the necessary conditions N1-N3 were checked analytically.[3] In order to also provide a visual impression of the EI criteria and to allow the assessment of the desirable properties D1-D3, contour plots of the criteria were generated in a bi-objective space – omitting ParEGO and EI-EMO because of the a-priori reduction to the single-objective EI. The contour lines represent the evaluation of different $\hat{\mathbf{y}}$ for constant $\hat{\mathbf{s}}$ using MATLAB® implementations of the criteria based on code of the corresponding authors. The reference set $\mathbf{A}_{PF}$ of size $|\mathbf{A}_{PF}| = 7$ was created by the evaluation of a 65-point Latin Hypercube Design in the domain $[-1, 2]^2$ on the bi-objective generalized Schaffer problem [7] with exponent $\gamma = 0.5$ (convex). The true PF is located within the domain $[0, 1]^2$. In order to analyze the influence of $\hat{\mathbf{s}}$, predictions slightly outside the objective space were also considered. Therefore, the evaluation of the possible predictions $\hat{\mathbf{y}}$ were visualized in the domain $[-0.1, 1.5]^2$ using a constant $\hat{\mathbf{s}} = \mathbf{0.2}$. This relatively high value was chosen because the behavior for low $\hat{\mathbf{s}}$ can be derived analytically in most cases. For the calculation of the indicator-based criteria, the ideal point $\mathbf{i} = (-0.1, -0.1)$ and the reference point $\mathbf{r} = (2, 2)$ were chosen. Consequently,

[3] When not explicitly stated, we omit the special case of $\mathbf{s} = \mathbf{0}$, as it holds only for known evaluations which have a negligibly low probability of being evaluated again.

**Table 2.** Overview of the compliance of the multi-objective EI criteria with the defined conditions and properties

| | ParEGO | EI-EMO | WS-EI | TA-EI | Euclid | SMS-EGO (old) | SMS-EGO (new) | SExI |
|---|---|---|---|---|---|---|---|---|
| N1 | $\surd^*$ | $\surd$ | $\surd$ | $\surd^*$ | $-$ | $\surd$ | $\surd$ | $\surd$ |
| N2 | $\surd$ | $\surd$ | $\surd$ | $\surd$ | $-$ | $\surd$ | $\surd$ | $\surd^{**}$ |
| N3 | $\surd$ | $(\surd)$ | $-$ | $\surd$ | $\surd$ | $\surd$ | $\surd$ | $\surd$ |
| D1 | ○ | $-$ | $--$ | $+$ | $++$ | $++$ | $++$ | $+$ |
| D2 | $+$ | $++$ | $++$ | $-$ | $-$ | $--$ | $-$ | $++$ |
| D3 | ○ | $+$ | $-$ | $+$ | $+$ | $-$ | ○ | $+$ |
| D4 | $++$ | $+$ | $+$ | ○ | $--$ | $++$ | $++$ | $--$ |

$^*$Only for weight vectors with strictly positive components.
$^{**}$Empirical evidence, no formal proof could be provided until now.

it is ensured that all evaluated vectors are dominated by $\mathbf{i}$ and dominate $\mathbf{r}$. In practice, this can be accomplished by determining $\mathbf{i}$ and $\mathbf{r}$ by minimizing and maximizing the surrogate model of each objective. Thus, it is also assumed in the proofs of this section. If required, $N = 501$ uniformly distributed weight vectors including $(0, 1)$ and $(1, 0)$ were used. Due to space limitations, only a few of the contour plots can be shown in the paper. All figures computed for this study (also for $\gamma = 1$, $\gamma = 2$, and $\hat{\mathbf{s}} = \mathbf{0.01}$) can be found online.[4]

**N1:** For Euclid and SExI, $EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF}) = \int_{-\infty}^{\mathbf{r}} I(\mathbf{y}, \mathbf{A}_{PF})\phi_{\hat{\mathbf{y}},\hat{\mathbf{s}}}(\mathbf{y})\, dy$ holds. By centering the PDF, we get $EI(\hat{\mathbf{y}}, \hat{\mathbf{s}}, \mathbf{A}_{PF}) = \int_{-\infty}^{\mathbf{r}} I(\mathbf{y}+\hat{\mathbf{y}}, \mathbf{A}_{PF})\phi_{\mathbf{0},\hat{\mathbf{s}}}(\mathbf{y})\, dy$. Thus, N1, assuming equal $\hat{\mathbf{s}}$, is directly related to the compliance of $I(\mathbf{y}, \mathbf{A}_{PF})$ with the dominance relation. This relation also holds for $I(\mathbf{y}_{LCB}, \mathbf{A}_{PF})$ because the constant displacement $\alpha\hat{\mathbf{s}}$ can be neglected. Whereas the HV used in SExI and SMS-EGO is Pareto-compliant [15], the Euclidean distance is not (cf. Fig. 3).

All other approaches directly use $I(y, f_{min})$ of the single-objective EI. Thus, their compliance with N1 is related to the preprocessing before the EI computation. Both, the TA and the WS, are compliant with the dominance relation as long as no component of the weight vector is zero [16]. In this case, an improvement in the objective with the zero component is not reflected in the scalarization (cf. Fig. 4 (left) for $f_1 = 0$ or $f_2 = 0$). Consequently, ParEGO and TA-EI are only compliant with N1 if no such weight vectors are used. WS-EI takes the sum over the EI of all weight vectors. Thus, at least one weight vector with a positive component for each objective is required, which is very likely to be fulfilled. Despite the a-posteriori selection of the extreme solutions, all single-objective EIs are considered during EI-EMO. Thus, N1 holds for this criterion.

**N2:** It is has been shown by Jones et al. [4, pp. 172f.] that a higher $\hat{s}$ monotonically improves the single-objective EI, even when $I(\hat{y}, f_{min}) > 0$. Based on this result, N2 is fulfilled for ParEGO, EI-EMO, and all subproblems of TA-EI and WS-EI, which directly transfers to the final aggregation. Since the $\mathbf{y}_{LCB}$ is linearly improved by $\hat{\mathbf{s}}$, N2 also holds for both variants of SMS-EGO.

[4] http://www.pbase.com/emmerich/expected_improvement

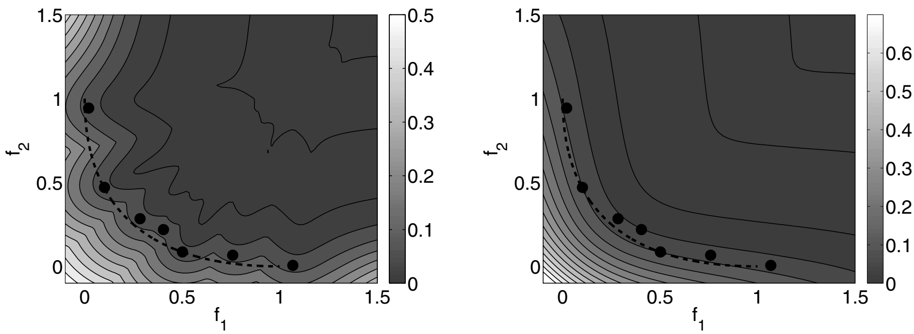


**Fig. 3.** Comparison of Euclid (left) and SExI (right)

For Euclid and SExI, we conducted an experiment, in which $\hat{s} = 0, 0.1, \ldots, 1$ were evaluated for each $\hat{y}$ of Fig. 2-4. For SExI no counterexample was found, but Euclid violated N2 in 797 of 4925 cases. This violation is often caused by a reduced minimum distance due to a movement of the centroid from the dominated to the nondominated area. The results of SExI provide empirical evidence for a compliance with N2, but no formal proof could be provided until now.

**N3:** SMS-EGO, SExI, and Euclid fulfill N3 by definition. If $\hat{\mathbf{s}} = \mathbf{0}$, no displacement of $\mathbf{y}_{LCB}$ occurs or the PDF becomes singular. This results in a direct evaluation of $I(\hat{\mathbf{y}}, \mathbf{A}_{PF})$. In ParEGO and EI-EMO, the single-objective EI is evaluated. Therefore, N3 also holds for the considered subproblems. The center solution of EI-EMO, however, is not related to a clearly formulated improvement, which does not allow a complete evaluation of this approach.

Given the final decision making, applied in TA-EI and WS-EI, both aim for a maximum improvement, either of a single subproblem (TA-EI) or of the sum over all subproblems (WS-EI). However, the separated computation of EIs and the subsequent aggregation is only straightforward for maximizing the improvement on a single subproblem. In WS-EI, the sum of the EIs substitutes the EI of the sum. Since the EI nonlinearly depends on $\hat{y}$ and $\hat{s}$, N3 is violated. In order to calculate the actual EI, the mean and the standard deviation of the sum of scalarizations have to be computed. To accomplish this, the equations for calculating each $\hat{y}_{sc}$ and $\hat{s}_{sc}$ can be applied again.

**D1:** It has been shown that the maximization of the HV increment produces well-distributed sets [7]. Given that $\mathbf{r}$ is sufficiently far away from $\mathbf{A}_{PF}$, the spread is also improved [17]. Therefore, all criteria based on the HV cope with D1 for sufficiently small $\hat{\mathbf{s}}$ (cf. N3). However, a comparison of Fig. 2 and Fig. 3 (right) reveals that the gap-filling property of the SExI fades away with increasing $\hat{\mathbf{s}}$ whereas it is conserved for SMS-EGO. This is caused by the fact that samples from $\mathcal{N}(\hat{\mathbf{y}}, \hat{\mathbf{s}})$ can improve the distribution or spread of $\mathbf{A}_{PF}$, even if $\hat{\mathbf{y}}$ does not improve it. Moreover, this property enhances the guidance to the most promising local optima, as discussed for D3. Since the maximization of the Euclidean distance to the neighboring solution is an established diversity-measure,

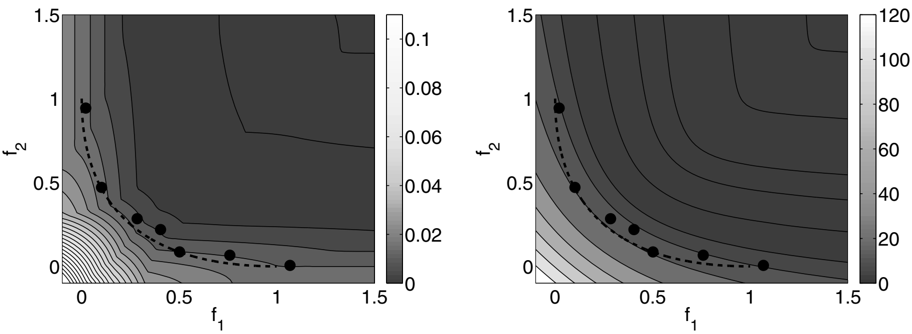


**Fig. 4.** Comparison of TA-EI (left) and WS-EI (right)

Euclid also copes with D1. Contrary to the direct approach, this also holds for high $\hat{\mathbf{s}}$ as shown in Fig. 3 (left).

For the scalarization-based approaches, only TA-EI copes with D1. As shown in Fig. 4 (left), the contour lines indicate the improvement by filling the gaps in the upper left part of $\mathbf{A}_{PF}$. Compared to TA-EI, which evaluates all weight vectors in each iteration, ParEGO randomly chooses a weight vector which may target toward an already crowded region of $\mathbf{A}_{PF}$, deteriorating its compliance with D1. The WS-EI is generally biased to the knee (convex) or to the extremes (concave) of $\mathbf{A}_{PF}$. The maximization of the sum of EIs produces an additional bias toward the center of the targets defined by the weight vectors (cf. Fig. 4, right). The a-posteriori selection of EI-EMO exclusively focuses on the extremes and the center of the EI Pareto front. Thus, only the spread of $\mathbf{A}_{PF}$ will be improved. A good distribution between the extremes cannot be accomplished.

**D2:** Only SExI, WS-EI and EI-EMO are continuous and differentiable over the whole domain. TA-EI and Euclid use maximum or minimum operations which lead to nondifferentiabilities of the corresponding criterion (cf. the left plots of Fig. 3 and Fig. 4). In ParEGO, the nondifferentiabilities are smoothed out by the surrogate model, making the actual EI criterion continuous and differentiable.

The problem of discontinuities in the old SMS-EGO approach and the answer of the new one has already been described in section 3. However, the nondifferentiabilities at the corners of the attainment surface of $\mathbf{A}_{PF}$ could not be resolved. This is shown in Fig. 2 by the contour lines in the proximity of $\mathbf{A}_{PF}$.

**D3:** All approaches relying on an EI formulation without penalties can show plateaus of zero EI based on the limited machine accuracy. This problem is overcome by the penalty functions used in SMS-EGO. Nevertheless, the fitness landscapes of SExI and TA-EI are evaluated best since these approaches show strong gradients to their local optima. This is shown in Fig 3 (right) and Fig. 4 (left). In contrast, the gradients in the landscape of SMS-EGO are very local, making the search for the global optimum difficult (cf. Fig. 2). The approach of Keane shows the most complex fitness landscape with many local optima. Nevertheless, the gradients to each local optimum are clearly defined even far from those, and the size of each basin grows with quality of the optimum (cf. Fig. 3, left). WS-EI fails to indicate the direction toward the optimum in the knee of $\mathbf{A}_{PF}$ by providing gradients that are normal to the true PF. In ParEGO, the original EI (equation 1) has to be optimized, which is known for multimodality and plateaus. For EI-EMO these problems are relaxed because the multi-objective optimization of the different EIs enhances diversity and avoids the premature convergence to one of the local optima of the single-objective EI.

**D4:** Besides the approaches based on a piecewise integration over the nondominated region which require a tedious partitioning of the objective space, all EI criteria are easy to implement. For ParEGO, only one model has to be computed making it the fastest of all approaches. The multi-objective optimization in EI-EMO slightly increases the runtime compared to the single-objective optimizations performed in all other approaches.

Regarding the empirical runtime for computing the figures in bi-objective space, SMS-EGO and Keane are the fastest approaches ($\approx 4$ $s$ for 6561 evaluations). The scalarization-based EI criteria show a surprisingly high runtime for the recommended number of $N = 501$ weight vectors (Tchebycheff: 978 $s$, weighted sum: 398 $s$), deteriorating their rating in D4. The direct integration takes about 140 $s$ for the bi-objective computations. However, the runtime of the SExI and Euclid may increase exponentially with the number of the objectives $m$.

## 5 Conclusions and Outlook

In this paper, we summarized, compared, and analyzed existing EI criteria for multi-objective optimization. For one of the criteria, an improved variant has been introduced. Moreover, we proposed necessary conditions and desired properties for a formal evaluation. Based on theoretical and empirical analyses, we showed that Euclid and WS-EI are not compliant with the dominance relation or provide no clear formulation of the desired improvement. Thus, these approaches should no longer be used. All other approaches considered in this study fulfill the necessary conditions. Depending on the application, the appropriate criterion can be chosen based on the performance on the desired properties (cf. Table 2).

For the scalarization-based approaches, improvements in the formulation and requirements for the weight vectors could be stated. However, a trade-off between accuracy and runtime still exists. To overcome this problem, an adaptive calculation of the corresponding optimal weight vector [16] during the integration seems promising. The problem of plateaus in the EI landscapes may be solved by combining the EI with penalty functions for values below the machine accuracy.

## Acknowledgments

This paper is based on investigations of the collaborative research center SFB/TR TRR 30, which is kindly supported by the Deutsche Forschungsgemeinschaft (DFG).

## References


1. Knowles, J., Nakayama, H.: Meta-modeling in multiobjective optimization. In: Branke, J., et al. (eds.) Multiobjective Optimization – Interactive and Evolutionary Approaches, pp. 461–478. Springer, Berlin (2008)
2. Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P.: Design and analysis of computer experiments. Stat. Sci. 4(4), 409–435 (1989)
3. Mockus, J.B., Tiesis, V., Zilinskas, A.: The application of bayesian methods for seeking the extremum. In: Dixon, L.C.W., Szegö, G.P. (eds.) Towards Global Optimization, vol. 2, pp. 117–129. Amsterdam, New York (1978)
4. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. J. Glob. Optim. 13(4), 455–492 (1998)
5. Bartz-Beielstein, T., Lasarczyk, C., Preuss, M.: Sequential parameter optimization. In: McKay, B., et al. (eds.) Proc. CEC, pp. 773–780. IEEE, Los Alamitos (2005)
6. Zitzler, E., Thiele, L., Bader, J.: On set-based multiobjective optimization. IEEE Trans. Evol. Comput. 14(1), 58–79 (2010)
7. Emmerich, M.: Single- and Multi-objective Evolutionary Design Optimization Assisted by Gaussian Random Field Metamodels. PhD thesis, Universität Dortmund (2005)
8. Keane, A.J.: Statistical improvement criteria for use in multiobjective design optimization. AIAA J. 44(4), 879–891 (2006)
9. Liu, W., Zhang, Q., Tsang, E., Liu, C., Virginas, B.: On the performance of metamodel assisted MOEA/D. In: Kang, L., Liu, Y., Zeng, S., et al. (eds.) ISICA 2007. LNCS, vol. 4683, pp. 547–557. Springer, Heidelberg (2007)
10. Emmerich, M., Deutz, A.H., Klinkenberg, J.W.: The computation of the expected improvement in dominated hypervolume of pareto front approximations. Technical Report 4-2008 Leiden Institute of Advanced Computer Science, LIACS (2008), `http://www.liacs.nl/~emmerich/TR-ExI.pdf`
11. Zhang, Q., Liu, W., Tsang, E., Virginas, B.: Expensive multiobjective optimization by MOEA/D with gaussian process model. IEEE Trans. Evol. Comput. (2010); Early Access (will be published)
12. Knowles, J.: ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Trans. Evol. Comput. 10(1), 50–66 (2006)
13. Jeong, S., Obayashi, S.: Efficient global optimization (EGO) for multi-objective problem and data mining. In: Corne, D., et al. (eds.) Proc. CEC, pp. 2138–2145. IEEE, Los Alamitos (2005)
14. Ponweiser, W., Wagner, T., Biermann, D., Vincze, M.: Multiobjective optimization on a limited amount of evaluations using model-assisted $\mathcal{S}$-metric selection. In: Rudolph, G., Jansen, T., Lucas, S., Poloni, C., Beume, N. (eds.) PPSN 2008. LNCS, vol. 5199, pp. 784–794. Springer, Heidelberg (2008)
15. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G.: Performance assessment of multiobjective optimizers: An analysis and review. IEEE Trans. Evol. Comput. 7(2), 117–132 (2003)
16. Steuer, R.E.: Multiple Criteria Optimization: Theory, Computation, and Application. Wiley, New York (1986)
17. Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: Optimal $\mu$-distributions and the choice of the reference point. In: Garibay, I., et al. (eds.) Proc. FOGA, pp. 87–102. ACM, New York (2009)