An Ensemble Surrogate-based Framework for Expensive Multiobjective Evolutionary Optimization

Qiuzhen Lin, *Member, IEEE*, Xunfeng Wu, Lijia Ma, Jianqiang Li, Maoguo Gong, *Senior Member, IEEE*, and Carlos A. Coello Coello, *Fellow, IEEE*

***Abstract*—Surrogate-assisted evolutionary algorithms (SAEAs) have become very popular for tackling computationally expensive multiobjective optimization problems (EMOPs), as the surrogate models in SAEAs can approximate EMOPs well, thereby reducing the time cost of the optimization process. However, with the increased number of decision variables in EMOPs, the prediction accuracy of surrogate models will deteriorate, which inevitably worsens the performance of SAEAs. To deal with this issue, this paper suggests an ensemble surrogate-based framework for tackling EMOPs. In this framework, a global surrogate model is trained under the entire search space to explore the global area, while a number of surrogate sub-models are trained under different search subspaces to exploit the sub-area, so as to enhance the prediction accuracy and reliability. Moreover, a new infill sampling criterion is designed based on a set of reference vectors to select promising samples for training the models. To validate the generality and effectiveness of our framework, three state-of-the-art evolutionary algorithms (nondominated sorting genetic algorithm III (NSGA-III), multiobjective evolutionary algorithm based on decomposition with differential evolution (MOEA/D-DE) and reference vector-guided evolutionary algorithm (RVEA)) are embedded, which significantly improve their performance for solving most of the test EMOPs adopted in this paper. When compared to some competitive SAEAs for solving EMOPs with up to 30 decision variables, the experimental results also validate the advantages of our approach in most cases.**

***Index Terms*—Evolutionary algorithms, multiobjective optimization, ensemble surrogate, model management.**

I. Introduction

In engineering applications, there exist some multiobjective optimization problems (MOPs) that require the simultaneous optimization of multiple (often conflicting) objectives. As multiobjective evolutionary algorithms (MOEAs) can find a set of Pareto-optimal solutions in a single run, they become an effective and popular tool for tackling MOPs [1]. Based on the selection criteria, most MOEAs can be classified into three main categories [2]: Pareto-based MOEAs [3], [4], decomposition-based MOEAs [5]-[6], and indicator-based MOEAs [7], [8]. These traditional MOEAs generally assume a sufficient number of function evaluations, so that the population can converge. However, for some MOPs modeled from practical applications, e.g., finite element analysis [9], computational fluid dynamics [10], and computational electromagnetics [11], the function evaluations require computationally expensive simulations, which consume a considerable amount of time or material resources. These problems are often called expensive MOPs (EMOPs) [12].

In recent years, surrogate-assisted evolutionary algorithms (SAEAs) have been widely used for tackling EMOPs, as they employ a computationally efficient surrogate model to replace the true expensive functions [13]. This way, the computational cost for evolutionary optimization in SAEAs can be significantly reduced. Many machine learning models can be employed as surrogate models, such as Kriging [14]-[17], radial basis functions (RBF) [18]-[19], artificial neural networks [20], support vector regression models [21]-[22], polynomial regression (PR) models [23], or a combination of multiple surrogate models [24]-[32]. Different surrogate models have distinct advantages for tackling specific EMOPs [33]. After selecting the surrogate models, the way to update them, which is referred to as the surrogate management criterion or infill sampling criterion, is also important, e.g., the probability of improvement (PoI) [34], the expected improvement (ExI) [35], and the lower confidence bound (LCB) [36]. These methods select promising solutions to be evaluated by true expensive functions and then use them to retrain the surrogate models.

According to the availability of new data, there are two kinds of SAEAs: (1) offline SAEAs, which depend on the existing date for optimization, and (2) online SAEAs, which can actively generate new data for optimization [37], [38]. This paper focuses only on online SAEAs that have additional samples for managing the surrogate models, which are more flexible for enhancing the optimization performance in practical cases. When designing online SAEAs, we often need to solve several challenges [39]. The first one is to select one appropriate surrogate model. In general, the surrogate model is chosen based on the experience of the users or engineers

Manuscript received XX. XX. 2019; revised XX. XX. 2019; accepted XX. XX. 2019. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 61876110, the Joint Funds of the NSFC under Key Program Grant U1713212, the Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20190808164211203, the Guangdong "Pearl River Talent Recruitment Program" under Grant 2019ZT08X603, and Shenzhen Science and Technology Innovation Commission (R2020A045). Prof. Coello Coello acknowledges support from the CONACyT project no. 1920, a 2018 SEP-Cinvestav grant (application no. 4), and the Basque Government through the BERC 2018-2021 program by the Spanish Ministry of Science.(*Corresponding author: Jianqiang Li*)

Q. Lin, X. Wu, L. Ma and J. Li are with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China (e-mail of Q. Lin: qiuzhlin@szu.edu.cn).

M. Gong is with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China.

Carlos A. Coello Coello is with the Computer Science Department, CINVESTAV-IPN, Mexico City, Mexico, and with Basque Center for Applied Mathematics (BCAM) & Ikerbasque, Spain.

[15]-[23], since there is little theoretical guidance available. The second challenge is how to use the surrogate model. The most conventional method is to use the surrogate model as the approximation of objective functions [40], and other potential methods include the approximation of a scalarizing function by converting an EMOP into a set of subproblems [41]-[43], the estimation of the rankings of solutions [44], the prediction of the hypervolume (HV) [45], and the classification of samples based on their fitness [20]. The third challenge is to select a specific MOEA. As different MOEAs have different advantages and limitations, one MOEA cannot always perform well in all EMOPs. Hence, suitable MOEAs should be selected based on the properties of the target EMOPs. The last challenge is to determine the termination criterion of online SAEAs, as the true function evaluations are very expensive. In spite of the above challenges, online SAEAs are still regarded as very promising for solving EMOPs [13].

In [15]-[23], the capabilities of Kriging models and other surrogate models to effectively approximate the objective functions were validated when tackling EMOPs with a small number of decision variables. However, the performance of these models will significantly deteriorate for tackling EMOPs with more than ten decision variables due to the poor prediction accuracy. Recently, some research studies [24]-[27] have indicated that a number of member models can provide a more accurate approximation performance for tackling EMOPs, but their performance is still not satisfactory enough on EMOPs with a large number of decision variables [46], [47]. Thus, this paper proposes an ensemble surrogate-based framework (ESF) for expensive multiobjective evolutionary optimization, which can achieve superior approximation performance for tackling various EMOPs. In our ESF, one global surrogate model and a number of surrogate sub-models are combined to compose an ensemble surrogate model with improved accuracy. Moreover, a new infill sampling criterion is presented in this paper based on a set of reference vectors, which selects promising solutions to retrain the surrogate models. To summarize, the main contributions of this paper are as follows:

(1) A new surrogate-based framework is proposed for tackling EMOPs, which trains surrogate models under different search spaces, while most of the existing ensemble models train their surrogate models in the entire search space [24]-[27]. One global surrogate model is trained under the entire search space to explore the global area, while a number of surrogate sub-models are trained under different search subspaces to exploit the sub-area. This way, the prediction accuracy and reliability of our framework are strengthened.
(2) A new infill sampling criterion is proposed based on a set of weight vectors to evaluate the quality of solutions. Solutions whose approximate objective values are dominated by truly evaluated samples will be removed to ensure convergence and then the remaining ones closest to a set of randomly selected reference vectors with good diversity are finally used to train the surrogate models.
(3) Any MOEA can be easily embedded into our framework. As examples, three well-known MOEAs (nondominated sorting genetic algorithm III (NSGA-III) [3], multiobjective evolutionary algorithm based on decomposition with differential evolution (MOEA/D-DE) [5], and reference vector-guided evolutionary algorithm (RVEA) [48]) have been embedded into our framework. The experiments show that their performance is significantly improved for tackling most of the test EMOPs adopted in this paper, which confirms the effectiveness of our framework.

The rest of this paper is organized as follows. Section II briefly discusses the related techniques and the motivations of this paper. Section III introduces the details of our framework, and the experimental results of the ESF with different SAEAs are presented in Section IV. Finally, Section V concludes the paper with a summary and a discussion of future work.

II. Related Work and Motivations

In this section, a number of representative SAEAs for tackling EMOPs are first introduced in Section II.A, where their strengths and limitations are summarized. Then, the motivations to design our ESF are presented in Section II.B. Please note that the details of EMOPs and Kriging model are given in Section 1 of the supplementary material due to page limitations.

A. Previous Related Work

In recent years, a number of SAEAs have been designed for solving EMOPs and some of them have shown very competitive performance. For example, a new MOEA/D variant with the Gaussian stochastic process model (MOEA/D-EGO) was proposed in [42] to solve EMOPs. In this approach, a Gaussian model is built for each subproblem based on the data obtained from the evolutionary search, which is then used to produce candidate solutions. A Kriging-assisted RVEA (K-RVEA) was presented in [16] for incorporating the Kriging model into RVEA [48] to approximate the objective values. K-RVEA exploits the uncertainty information from the Kriging model, the distribution of reference vectors, and the location of individuals, so as to balance the diversity and convergence in the model management, which can limit its training cost without sacrificing prediction accuracy. A classification-based SAEA (CSEA) was designed in [20] to solve EMOPs with many objectives, which can predict the dominance relationship between candidate solutions and reference solutions. CSEA can select promising solutions to run the true function evaluations by considering both the uncertainty information given by the surrogate model and their dominance relationship.

On the other hand, there are also some SAEAs suggesting to use multiple surrogates for tackling EMOPs, which are expected to generate more accurate approximation results. For example, an adaptive knowledge reuse framework was designed in [25] for SAEAs to solve EMOPs, which realizes an efficient transfer evolutionary multiobjective optimization with multiproblem surrogates (TEMO-MPS). This approach can acquire and transfer learned models across problems using the idea of multiproblem surrogates, so as to promote global optimization in EMOPs. A Gaussian process (GP)-based co-sub-Pareto front surrogate augmentation strategy

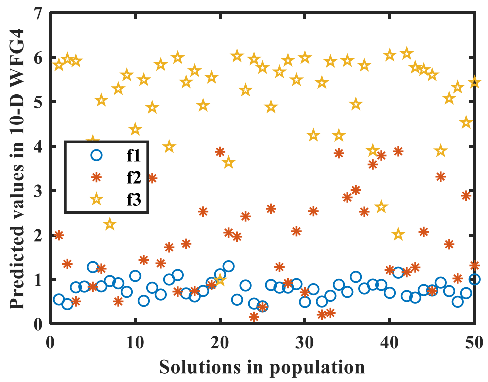


Fig. 1 Predicted objective values for the 10-D WFG4 problem

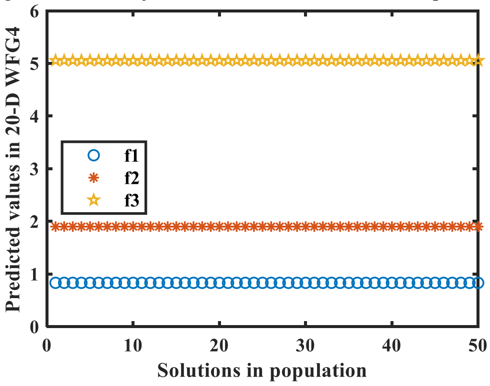


Fig. 2 Predicted objective values for the 20-D WFG4 problem

(GCS-MOE) was suggested in [26] to solve EMOPs, which decomposes an EMOP into a number of subproblems and uses the solution of each subproblem to approximate one sub-PF. Then, a multitask GP model is used to exploit the correlations among the subproblems, such that knowledge learned from subproblems can be transferred across sub-PFs to better solve EMOPs. A heterogeneous ensemble assisted MOEA (HeE-MOEA) was designed in [27] by using ensembles as surrogates and infill criteria for model management in evolutionary optimization. In this approach, a heterogeneous ensemble with a least square support vector machine and two RBFs is constructed to enhance its reliability for uncertainty estimation. Moreover, a selected subset of decision variables and a set of transformed variables are used as inputs of the heterogeneous ensemble to further enhance its diversity. A surrogate-ensemble assisted optimization method (SAEMO) was introduced in [28] to train multiple surrogate models to assist MOEAs. This approach suggests a new model management strategy, which measures the uncertainty by considering the distance to the samples in decision space and the approximation variance in objective space.

To have an overview for the SAEAs mentioned above, their strengths and limitations are summarized in Table A.II of the Supplementary Material due to page limitations. Although they can be applied to solve high-dimensional EMOPs with up to 30 dimensions, their performance is still not so satisfactory, which can be observed from our experimental studies in Section IV.

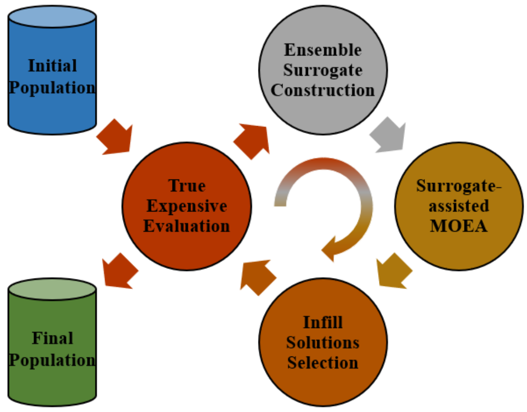


Fig. 3 A flowchart for our framework

B. *Motivations*

As mentioned above, various surrogate models have been proposed for SAEAs [14]-[32]. However, most of them are designed to effectively solve low-dimensional EMOPs, which are not so effective for EMOPs with more than ten decision variables. One important phenomenon we have observed in the experiments is that the predicted objective values of different solutions become extremely similar when the Kriging model is used for approximating high-dimensional objective functions.

In Fig. 1 and Fig. 2, the predicted objective values of 100 individuals are plotted for approximating the 10-dimensional (10-D) and 20-D WFG4 problems [49]. In each figure, the predicted value of each individual is plotted for each objective. From Fig. 1, we can observe that for each objective, all the predicted values are clearly distinguishable. However, in Fig. 2, the predicted values of all the individuals become nearly the same, making it impossible to distinguish the quality of different individuals in SAEAs. The above empirical results can be properly explained by analyzing the Kriging model provided in the Supplementary Material, from which we can find that the prediction of a new unknown individual $\mathbf{x}^{new}$ is significantly correlated with the distance between $\mathbf{x}^{new}$ and the training samples. Thus, when there is only a small number of training samples in a high-dimensional decision space, all the new individuals will be predicted to have extremely similar objective values, as they are all significantly far from these samples.

Based on the above discussion, it is reasonable to conclude that for high-dimensional EMOPs, the predicted objective values of different solutions will become barely distinguishable, resulting in SAEAs being less effective for searching promising solutions. Thus, to solve the above problem, a novel ensemble surrogate model is designed to provide more accurate predictions for new solutions of high-dimensional EMOPs, which includes one global surrogate model constructed in the original decision space and k surrogate sub-models built in different low-dimensional subspaces. Moreover, a new infill sampling criterion is designed based on the reference vectors to select solutions with good convergence, which can better manage our ensemble surrogate model. The details of the above proposal will be introduced in the following section.

Algorithm 1: The General Framework

1. Generate a set of reference vectors *V* and *N* solutions as a population *P*
2. Initialize an empty training database $DB = \varnothing$
3. Initialize the maximum number of infill solutions $N_{is} = 5$
4. **while** $FE \le FE_{max}$ **do**
5. $F \leftarrow$ True-expensive-evaluation (*P*)
6. Set $FE = FE + |P|$ and $DB = DB \cup \{P, F\}$
7. *Ensemble* $\leftarrow$ Ensemble-model-construction (*DB*)
8. $S \leftarrow$ Surrogate-assisted-MOEA (*DB*, *Ensemble*, g_{max})
9. $P \leftarrow$ Infill-solutions-selection (S, V, N_{is})
10. **end while**
11. Output nondominated solutions from *DB*

III. The Details of Our Framework

In this section, the details of our ESF for tackling EMOPs are introduced. To have an overview of the ESF, Fig. 3 illustrates the running of the ESF, which consists of five main components: initialization, true expensive evaluation, ensemble surrogate construction, surrogate-assisted MOEA, and infill solutions selection. Please note that after the initialization, our ESF will enter the evolutionary loop for true expensive evaluation, ensemble surrogate construction, surrogate- assisted MOEA, and infill solutions selection. After running out of the limited true function evaluations, the final nondominated solutions in the training data will be used as the output.

To clarify the running of the ESF, its complete pseudocode is also provided in **Algorithm 1**.

- **Initialization**: The initialization is run in lines 1-3. In line 1, a set of reference vectors *V* is generated by the normal boundary intersection (NBI) method [50], and *N* training sample points are generated by the Latin hypercube sampling method [51] to compose the initial population *P*. Then, an empty training database (*DB*) and the maximum number of infill solutions N_{is} are respectively initialized in line 2 and line 3.
- **Stopping condition**: After initialization, the evolutionary loop is run in lines 4-10 until the stopping condition ($FE \le FE_{max}$) is met in line 4.
- **True expensive evaluation**: At the beginning of the evolutionary loop, all the individuals of the initial population are evaluated by true expensive functions in line 5, which are then added into *DB* in line 6. Please note that these solutions in *DB* will be used to train the surrogate models.
- **Ensemble surrogate construction**: An ensemble surrogate model *Ensemble* will be constructed in line 7, which will train one global surrogate model and *k* surrogate sub-models.
- **Surrogate-assisted MOEA**: The surrogate-assisted MOEA is run in line 8 to generate an offspring population *S* with a prefixed number of generations, and the above ensemble surrogate model is used to replace the true expensive functions for evaluating solutions.
- **Infill solutions selection**: The offspring population *S* will undergo infill solutions selection, which will identify several promising solutions in the search space to update the ensemble surrogate model in line 9, thereby guaranteeing the accuracy of the ensemble surrogate model.

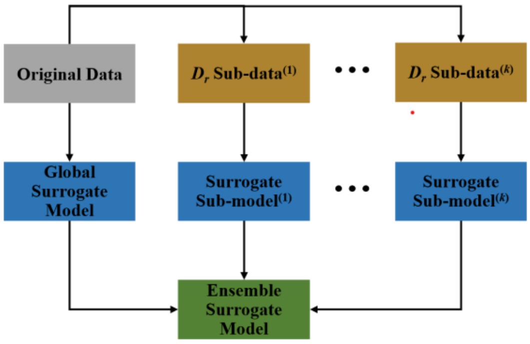


Fig. 4 Structure of the ensemble surrogate in ESF

The details of the above ensemble surrogate construction, surrogate-based evolutionary optimization, and infill solutions selection are respectively described below.

A. Ensemble Surrogate Construction

In this method, the Kriging model is adopted to build our ensemble surrogate model due to its high efficiency and simple working principle. As shown in Section II.B, the performance of the Kriging model will deteriorate for solving EMOPs with more than ten decision variables, as the predicted objective values are extremely similar and cannot be distinguished. Hence, an ESF is suggested here to better approximate the objective values in this case.

Fig. 4 depicts the structure of the ensemble surrogate model used in ESF, which includes a global Kriging model and *k* Kriging sub-models. The global Kriging model has been widely used in most existing SAEAs [15]-[16], which will be trained in the original search space using the solutions in *DB*, while the *k* Kriging sub-models will be trained in different subspaces. To clarify the running of this procedure, its pseudocode is provided in **Algorithm 2** with the input *DB*. In line 1 of **Algorithm 2**, the dimension of the subspaces is denoted by D_{LSM}, which is obtained by multiplying a coefficient *r* with its original dimension *D* as follows:

$$D_{LSM} = \lfloor r \times D \rfloor, \tag{1}$$

where *r* is a randomly generated real number in (0, 1) and the operation $\lfloor r \times D \rfloor$ keeps the integer part of $r \times D$. Then, the *k* Kriging surrogate sub-models (LSM_1, …, LSM_k) are trained in lines 2-6. Specifically, in line 3, D_{LSM} different indexes for decision variables are randomly selected from [1, *D*] in line 3 for the i^{th} surrogate sub-model, and then the i^{th} training data subset (DB_i) is sampled from the original database *DB* in line 4, which only includes D_{LSM} decision variables of each solution in *DB* with the indexes recorded in $Dind_i$. Then, a surrogate sub-model (LSM_i) is trained by DB_i in line 5. In line 7, the global Kriging surrogate model (*GSM*) is trained using *DB*. Here, similar samples will be removed before training these surrogate models. At last, the obtained global surrogate model and *k* surrogate sub-models are used to compose an ensemble surrogate model adopted in the ESF.

This ensemble surrogate model will be retrained when a new sample with a true evaluation is added into the database *DB*. As the ensemble model includes *k*+1 types of Kriging models, each member model should be assigned a weight to

Algorithm 2: Ensemble-model-construction (DB)

Calculate the dimension of surrogate sub-models $D_{LSM} = r \times D$
for i = 1 to k **do**
  Randomly select D_{LSM} decision variables from the variable vector as $Dind_i$
  $DB_i = DB(:, Dind_i)$
  Construct a surrogate sub-model LSM_i using DB_i
end for
Construct a global surrogate model GSM using DB
$Ensemble \leftarrow \{GSM, LSM_1, \ldots, LSM_k\}$
Output the ensemble surrogate model $Ensemble$

Algorithm 3: Surrogate-assisted-MOEA (DB, $Ensemble$, g_{max})

1. $S \leftarrow$ Nondominated solutions in DB
2. Set generation counter $g = 0$
3. **while** $g \le g_{max}$
4. $Q \leftarrow$ Offspring-reproduction(S)
5.   Evaluate solutions in Q by $Ensemble$ from Algorithm 2
6. $S \leftarrow$ Environmental-selection($S \cup Q$)
7. $g = g + 1$
8. **end while**
9. Output S

predict the objective value at the same time. In the initial stage of optimization, since the number of samples in DB is very limited, the Kriging model cannot provide high accuracy for approximating the objective values, especially for EMOPs with more than ten decision variables. With more truly evaluated samples added into DB, the performance of the Kriging model will be gradually improved. Thus, the weights of their member models should be dynamically adjusted to control the final output of our ensemble surrogate model. In this paper, the assigned weights for the k surrogate sub-models should be set large enough in the early stage of optimization, as they are more accurate when the number of trained samples is limited. With the increased number of evaluations, more samples will be added into DB, and the global surrogate model will become more accurate. In this case, the assigned weight for the global surrogate model will be gradually set as a larger value. Based on the above discussion, the output of the ensemble surrogate model can be expressed as follows:

$$y_{ens} = (1-c) \times \frac{1}{k} \sum_{i=1}^{k} y_i + c \times y_{k+1}, \quad (2)$$

$$c = FE / FE_{max}, \quad (3)$$

where y_i denotes the output of the i^{th} surrogate sub-model (i=1, 2,…, k), y_{k+1} is the output of the global surrogate model, FE is the number of true expensive evaluations, and FE_{max} is the maximum number of true expensive evaluations allowed.

B. Surrogate-assisted MOEA

After constructing the above ensemble surrogate model, our framework will move to the surrogate-assisted MOEA. In this procedure, any MOEA can be used and the candidate solutions produced by this MOEA are evaluated by our ensemble surrogate model rather than the true expensive functions. The details of the surrogate-assisted MOEA are presented in **Algorithm 3** with the inputs: DB, $Ensemble$ (the built ensemble surrogate model), g_{max} (the preset maximum number of generations). First, all the nondominated solutions in DB are used as the initial population S for the surrogate-assisted MOEA in line 1. Then, the generation counter is initialized in line 2, and the algorithm will enter into the evolutionary loop in lines 3-8.

Algorithm 4: Infill-solutions-selection (S, V, N_{is})

1. Remove duplicate solutions and solutions dominated by DB from S
2. Initialize an empty set $P = \varnothing$
3. **if** $|S| \le N_{is}$
4. $P = S$
5. **else**
6.   Randomly select N_{is} reference vectors from V as a reference vector set
7.   Calculate the angles of $s_i \in S$ to $w_i \in W$ by (4)
8.   **for** $j = 1 : |W|$ **do**
9.   Select one solution s closest to w_i
10.   **if** s is different from any point in P
11. $P = P \cup s$
12.   **end if**
13.   **end for**
14. **end if**
15. Output P

In line 4, an offspring population Q with the same size as S is generated by reproduction strategies, such as simulated binary crossover [52] or differential evolution [53], and polynomial-based mutation [54]. The fitness values for all the offspring solutions in Q are evaluated in line 5 by our ensemble surrogate model as introduced in Section III.A, environmental selection is run in line 6 on the combination of S and Q to select the next population, and the generation counter is increased by 1 in line 7. When the termination criterion ($g \le g_{max}$) is satisfied in line 3, the above evolutionary loop is ended, and the final population S will be used as the output.

C. Infill Solutions Selection

As one true function evaluation may consume much time in EMOPs, SAEAs aim to find a set of good approximations with very limited function evaluations. Our ESF is designed for online SAEAs, which allows additional function evaluations to be performed during the optimization process. Thus, an infill solutions selection is proposed in the ESF to select individuals to be evaluated by the true expensive functions after running the above surrogate-assisted MOEA. The details of this strategy are introduced in **Algorithm 4** with the inputs: S (the generated population from the surrogate-assisted MOEA), V (a set of evenly distributed reference vectors), N_{is} (the maximum number of infill solutions). This strategy aims to select a number of N_{is} infill solutions in each generation, which will be evaluated by true expensive functions and then used to update the ensemble surrogate model.

As shown in **Algorithm 4**, in lines 1-2, some duplicate solutions and solutions dominated by DB are removed from S, and the population P is initialized as an empty set. If there are fewer than N_{is} solutions in S, they are directly added into P for true function evaluation in line 4. Otherwise, for the sake of strengthening diversity, N_{is} reference vectors are randomly selected from V, which compose a reference vector set W in line 6, and then the individual having the minimal acute angle value to each reference vector in W is selected as the infill solutions, as shown in lines 7-13. Here, the acute angle value is reflected by $\cos\theta_{i,j}$, as follows:

$$\cos\theta_{i,j} = \frac{s_i \cdot w_j}{\|s_i\|}, \quad (4)$$

where $s_i \in S$ to $w_j \in W$. Please note that if the approximate objective values of these infill solutions using our ensemble surrogate model are dominated by the truly evaluated samples in *DB*, they won't be truly evaluated to avoid meaningless function evaluations. Only the remaining solutions with potentially good convergence are evaluated by true expensive function to retrain our ensemble surrogate model. At each generation, the N_{is} reference vectors used in (5) are different, as they are randomly selected from the uniformly distributed set *W*, which also guarantees the diversity of the selected infill solutions from *DB*.

IV. Experimental Studies

A. Test Problems and Performance Indicator

In our numerical experiments, four sets of well-studied test suites are used, namely, DTLZ1 to DTLZ7 [55], WFG1 to WFG9 [49], UF1 to UF9 [56], and MaF1 to MaF7 [57], where UF1 to UF7 are two-objective problems and other test cases are three-objective problems. Due to page limitations, the main features of the DTLZ, WFG, UF, and MaF test suites are summarized in Table A.III of the Supplementary Material.

This paper uses inverted generational distance (IGD) [58] to give an overall assessment of the compared algorithms. Let $\mathbf{P}^*$ be a set of evenly distributed points along the PF in the objective space. Let $\mathbf{S}$ be an approximate set obtained by the MOEAs to the PF; then, IGD is defined as the average distance from $\mathbf{P}^*$ to $\mathbf{S}$, as follows:

$$IGD(\mathbf{S},\mathbf{P}^*) = \frac{\sum_{\mathbf{x}\in\mathbf{P}^*} dist(\mathbf{x},\mathbf{S})}{|\mathbf{P}^*|}, \quad (5)$$

where $dist(\mathbf{x}, \mathbf{S})$ represents the minimum Euclidean distance between $\mathbf{x}$ and the points in $\mathbf{S}$, and $|\mathbf{P}^*|$ indicates the size of $\mathbf{P}^*$. If $\mathbf{P}^*$ is large enough to represent the PF very well, $IGD(\mathbf{S}, \mathbf{P}^*)$ can measure both the diversity and convergence of S in a sense. To have a low $IGD(\mathbf{S}, \mathbf{P}^*)$ value, the set S must be very close to the PF and cannot miss any part of the PF.

B. Results of Embedding Three MOEAs into the ESF

Here, to evaluate the effectiveness of the ESF, three competitive MOEAs (NSGA-III [3], MOEA/D-DE [5], and RVEA [48]) are embedded into the ESF, forming three new SAEAs (ESF-NSGA-III, ESF-MOEA/D-DE, and ESF-RVEA) for solving the adopted test EMOPs. Thus, three comparisons of ESF-NSGA-III versus NSGA-III, ESF-MOEA/D-DE versus MOEA/D-DE, and ESF-RVEA versus RVEA are conducted for solving DTLZ1-DTLZ7, WFG1-WFG9, UF1-UF9, and MaF1-MaF7 with 10, 20, and 30 decision variables. In all our experimental studies, the initial population size and the maximum number of true expensive evaluations are set as $11\times D$-1 and $11\times D$+120, respectively (D is the dimensionality of the problems), which are suggested in [27]. The number of surrogate sub-models k in all variants with ESF is set to 2. Other parameters are set the same as suggested in their original references [3], [5], [48]. Some parameters in our framework, e.g., the number of surrogate sub-models and the dimension of surrogate sub-models in **Algorithm 2**, are tuned by using the trial-and-error method. All the experiments are run on a personal computer equipped with an Intel(R) Core (TM) i7-7700 CPU, 3.60 GHz (processor), and 24.0 GB (RAM). In order to reproduce our experiments, the source code of ESF can be downloaded in https://github.com/Xunfeng-Wu/ESF-EMOO.

TABLE I
Summary of the Significance Test of three comparisons

Problem		ESF-NSGA-III vs. NSGA-III	ESF-MOEA/D-DE vs. MOEA/D-DE	ESF-RVEA vs. RVEA
DTLZ and WFG (D = 10, 20, 30)	Better	44	40	45
	Worse	0	2	0
	Similar	4	6	3
UF and MaF (D = 10, 20, 30)	Better	47	39	44
	Worse	0	5	1
	Similar	1	4	3

TABLE II
Summary of Comparison of ESF-RVEA with different k values

ESF-RVEA (k = 2) vs.		k = 0	k = 1	k = 3	k = 4	k = 5
DTLZ, WFG, UF, and MaF (D = 20)	Better	20	20	5	8	10
	Worse	2	2	7	3	2
	Similar	10	10	20	21	20
Best	6	2	2	11	6	5

The detailed IGD results of all three comparisons are presented in Table A.IV and Table A.V of the Supplementary Material. To ensure statistically sound conclusions, Wilcoxon rank-sum tests with a significance level of 0.05 are run, which indicate statistically significant differences on the IGD results. Here, Table I provides a summary of the significance test on the IGD results for all the test problems with D = 10, 20, 30, where "Better", "Worse" and "Similar" respectively indicate the number of test problems in which the performance of the variants using ESF is better than, worse than, and similar to that of the original algorithms.

As observed from the summary on all the DTLZ and WFG test problems, 44, 40, and 45 out of 48 cases show the superiority of ESF-NSGA-III over NSGA-III, ESF-MOEA/D-DE over MOEA/D-DE, and ESF-RVEA over RVEA, respectively. The superiority of MOEA/D-DE over ESF-MOEA/D-DE is only found in 2 cases. Both NSGA-III and RVEA cannot perform better than ESF-NSGA-III and ESF-RVEA in all cases, respectively. For the summary on all UF and MaF test problems, ESF-NSGA-III, ESF-MOEA/D-DE, and ESF-RVEA outperform their original algorithms in 47, 39, and 44 out of 48 cases, respectively, while they are only outperformed by their original algorithms in 0, 5, and 1 cases, respectively. Therefore, it is reasonable to conclude that the embedding of these MOEAs into the ESF can bring significant improvements to their ability to tackle EMOPs.

In order to provide more details, as observed from the IGD results in Table A.IV and Table A.V for all test problems with different numbers of decision variables, the performance of ESF-NSGA-III can be obviously enhanced in most of the test problems with all the considered dimensions, except for DTLZ1, DTLZ3, WFG3, and MaF3. ESF-MOEA/D-DE is significantly better than MOEA/D-DE on DTLZ6, DTLZ7, MaF1, MaF2, MaF7, and all the WFG and UF test problems. ESF-RVEA achieves remarkable improvement over RVEA on all the adopted test problems except for MaF3 and MaF5.

C. Impact of Parameter k in ESF

To study the impact of parameter k (the number of surrogate sub-models in the ensemble surrogate model), some experiments are run in this subsection. Using the parameters settings in Section IV.B, ESF-RVEA with different k values from {0, 1, 2, 3, 4, 5} are experimentally compared. Due to page limita-

TABLE III
COMPARISON OF ESF-RVEA WITH THREE COMPETITIVE ALGORITHMS ON THE DTLZ AND WFG TEST PROBLEMS

Problem	M	D	MOEADEGO	CSEA	KRVEA	ESF-RVEA
WFG1	3	10	2.2096e+0 (8.12e-2) -	1.9012e+0 (1.03e-1) =	**1.8715e+0 (1.52e-1) =**	1.8969e+0 (1.14e-1)
	3	20	2.2595e+0 (5.13e-2) -	**1.8653e+0 (1.11e-1) =**	1.9154e+0 (1.41e-1) =	1.9180e+0 (1.34e-1)
	3	30	2.2102e+0 (5.37e-2) -	**1.8357e+0 (9.20e-2) =**	2.0391e+0 (1.54e-1) -	1.8918e+0 (1.31e-1)
WFG2	3	10	7.3253e-1 (5.11e-2) -	6.3425e-1 (8.13e-2) -	4.7208e-1 (4.77e-2) -	**4.1092e-1 (6.92e-2)**
	3	20	8.2115e-1 (7.07e-2) -	6.9103e-1 (4.33e-2) -	6.4857e-1 (4.59e-2) -	**5.5683e-1 (5.30e-2)**
	3	30	7.9521e-1 (5.53e-2) -	6.7734e-1 (3.63e-2) =	7.2626e-1 (3.06e-2) -	**6.6353e-1 (3.26e-2)**
WFG3	3	10	6.4219e-1 (3.23e-2) -	5.3710e-1 (4.81e-2) -	4.8030e-1 (5.68e-2) -	**4.1545e-1 (6.93e-2)**
	3	20	6.7566e-1 (2.69e-2) -	6.7603e-1 (2.18e-2) -	6.6656e-1 (3.06e-2) -	**5.8340e-1 (5.15e-2)**
	3	30	**6.5069e-1 (2.22e-2) +**	7.1985e-1 (2.02e-2) -	7.0974e-1 (2.28e-2) -	6.8703e-1 (2.36e-2)
WFG4	3	10	5.9668e-1 (3.77e-2) -	4.8781e-1 (3.53e-2) -	4.9023e-1 (2.28e-2) -	**4.6838e-1 (3.22e-2)**
	3	20	5.7035e-1 (3.49e-2) -	5.1747e-1 (2.56e-2) -	5.1092e-1 (1.40e-2) -	**4.9165e-1 (1.73e-2)**
	3	30	5.8112e-1 (4.58e-2) -	5.2168e-1 (2.04e-2) -	5.2011e-1 (1.92e-2) -	**5.0290e-1 (1.45e-2)**
WFG5	3	10	6.1371e-1 (4.35e-2) -	6.0581e-1 (3.03e-2) -	4.8252e-1 (5.28e-2) =	**4.7142e-1 (4.96e-2)**
	3	20	6.5444e-1 (3.95e-2) -	6.6042e-1 (2.47e-2) -	5.9721e-1 (2.36e-2) -	**5.8536e-1 (4.61e-2)**
	3	30	7.1948e-1 (2.31e-2) -	6.7227e-1 (1.89e-2) =	6.6755e-1 (1.92e-2) =	**6.6428e-1 (2.18e-2)**
WFG6	3	10	8.3326e-1 (5.28e-2) -	7.4787e-1 (3.95e-2) -	7.7005e-1 (3.50e-2) -	**7.0174e-1 (5.40e-2)**
	3	20	9.0592e-1 (5.82e-2) -	8.2740e-1 (2.00e-2) -	8.3107e-1 (1.88e-2) -	**7.9811e-1 (2.80e-2)**
	3	30	8.8022e-1 (3.85e-2) -	8.4863e-1 (2.55e-2) =	8.5984e-1 (1.08e-2) -	**8.3507e-1 (1.49e-2)**
WFG7	3	10	6.8092e-1 (2.60e-2) -	6.2651e-1 (4.10e-2) -	6.3650e-1 (2.97e-2) -	**5.9057e-1 (3.91e-2)**
	3	20	6.7410e-1 (1.78e-2) -	6.6076e-1 (2.14e-2) -	6.5610e-1 (2.05e-2) -	**6.2804e-1 (2.50e-2)**
	3	30	6.6880e-1 (1.25e-2) -	6.6410e-1 (1.85e-2) -	6.6097e-1 (9.44e-3) -	**6.4408e-1 (1.31e-2)**
WFG8	3	10	8.7123e-1 (2.55e-2) -	7.8145e-1 (4.18e-2) -	7.5785e-1 (3.32e-2) -	**6.4641e-1 (4.02e-2)**
	3	20	8.1724e-1 (1.44e-2) -	7.7783e-1 (3.12e-2) -	7.1260e-1 (2.06e-2) -	**6.5669e-1 (2.60e-2)**
	3	30	8.0869e-1 (2.05e-2) -	7.5291e-1 (1.97e-2) -	7.1620e-1 (1.86e-2) -	**6.8163e-1 (2.28e-2)**
WFG9	3	10	8.0904e-1 (5.80e-2) -	**7.0577e-1 (6.65e-2) =**	7.2568e-1 (6.38e-2) =	7.2375e-1 (9.32e-2)
	3	20	8.8078e-1 (4.68e-2) -	8.2519e-1 (4.77e-2) -	8.5691e-1 (3.87e-2) -	**7.9755e-1 (5.12e-2)**
	3	30	9.1645e-1 (4.28e-2) -	8.7109e-1 (4.19e-2) =	8.8436e-1 (3.50e-2) -	**8.5267e-1 (3.65e-2)**
DTLZ1	3	10	9.3945e+1 (2.08e+1) =	**7.2993e+1 (2.01e+1) +**	9.2572e+1 (2.50e+1) =	9.5635e+1 (2.57e+1)
	3	20	2.8346e+2 (7.10e+1) +	**2.7031e+2 (3.56e+1) +**	3.2056e+2 (3.28e+1) =	3.0916e+2 (5.10e+1)
	3	30	5.9812e+2 (9.87e+1) -	**5.1746e+2 (6.13e+1) +**	5.7724e+2 (5.49e+1) =	5.6578e+2 (5.28e+1)
DTLZ2	3	10	3.4234e-1 (2.82e-2) -	2.5911e-1 (2.94e-2) -	1.5747e-1 (2.32e-2) -	**1.2522e-1 (2.01e-2)**
	3	20	6.3969e-1 (6.39e-2) -	6.1130e-1 (7.10e-2) -	5.8391e-1 (7.29e-2) -	**3.7523e-1 (5.43e-2)**
	3	30	**7.9886e-1 (9.67e-2) =**	1.0436e+0 (1.39e-1) -	1.1155e+0 (1.42e-1) -	8.0800e-1 (1.26e-1)
DTLZ3	3	10	2.3071e+2 (4.48e+1) =	**1.9245e+2 (5.42e+1) +**	2.4851e+2 (6.07e+1) =	2.4419e+2 (5.51e+1)
	3	20	**6.9018e+2 (2.12e+2) +**	7.9731e+2 (1.08e+2) +	9.9337e+2 (1.18e+2) -	8.6241e+2 (1.02e+2)
	3	30	**1.3314e+3 (4.18e+2) +**	1.5342e+3 (1.14e+2) +	1.7653e+3 (1.43e+2) -	1.6467e+3 (1.61e+2)
DTLZ4	3	10	6.2963e-1 (4.98e-2) -	5.2100e-1 (1.28e-1) =	**3.9444e-1 (9.58e-2) +**	4.5916e-1 (1.25e-1)
	3	20	1.1629e+0 (9.81e-2) -	**7.2133e-1 (1.10e-1) +**	9.5826e-1 (1.44e-1) -	8.2261e-1 (1.65e-1)
	3	30	1.6645e+0 (2.06e-1) -	**1.1112e+0 (1.63e-1) +**	1.6841e+0 (1.59e-1) -	1.3225e+0 (2.04e-1)
DTLZ5	3	10	2.6198e-1 (3.01e-2) -	1.4470e-1 (3.85e-2) -	1.0889e-1 (2.75e-2) -	**4.6751e-2 (1.35e-2)**
	3	20	5.4535e-1 (8.11e-2) -	5.7556e-1 (9.85e-2) -	4.8721e-1 (7.51e-2) -	**2.6170e-1 (7.81e-2)**
	3	30	7.1932e-1 (9.29e-2) =	1.0525e+0 (1.02e-1) -	1.0150e+0 (1.50e-1) -	**6.8732e-1 (1.36e-1)**
DTLZ6	3	10	**2.2506e+0 (6.26e-1) +**	5.5350e+0 (4.36e-1) -	3.3660e+0 (3.51e-1) +	3.9463e+0 (5.77e-1)
	3	20	**7.6919e+0 (1.78e+0) +**	1.4648e+1 (6.61e-1) -	1.1324e+1 (8.65e-1) =	1.1258e+1 (8.60e-1)
	3	30	**1.6182e+1 (3.02e+0) +**	2.3248e+1 (6.49e-1) -	2.0946e+1 (7.86e-1) -	2.0286e+1 (1.05e+0)
DTLZ7	3	10	3.3034e-1 (1.13e-1) =	2.1261e+0 (7.37e-1) -	**2.2164e-1 (7.42e-2) +**	3.6177e-1 (2.00e-1)
	3	20	4.3352e+0 (2.14e+0) -	5.0067e+0 (1.00e+0) -	5.9988e-1 (4.24e-1) =	**4.6737e-1 (1.98e-1)**
	3	30	7.3391e+0 (1.46e+0) -	6.0387e+0 (8.67e-1) -	4.3108e+0 (3.52e+0) -	**8.5037e-1 (3.30e-1)**
+/-/=			7/36/5	8/31/9	3/33/12	--

tions, the IGD results on all the DTLZ, WFG, UF, and MaF test problems with 20 decision variables are given in Table A.VI of the Supplementary Material. Table II summarizes the final comparison of results of ESF-RVEA with different k values based on their IGD results, where "Better", "Worse" and "Similar" respectively indicate the number of test problems in which the performance of ESF-RVEA using k = 2 is better than, worse than, and similar to that of ESF-RVEA with other k values, while "Best" indicates the number of test problems in which the corresponding algorithm performs best.

As observed from Table II, ESF-RVEA using k = 2 has statistically similar IGD results with respect to those using k = 3, k = 4 and k = 5 on 20, 21, and 20 out of 32 cases, respectively. It seems that ESF-RVEA is not too sensitive to the setting of k when changing from 2 to 5. However, ESF-RVEA with k = 2 obtains significantly better or statistically similar IGD results than those with k = 0 and k = 1 on both 27 out of 32 cases. Setting k = 0 means that no sub-model is employed and only a global Kriging model is used; thus, the performance of ESF-RVEA with k = 0 will significantly deteriorate. To visually explain this phenomenon, the predicted objective values obtained by one global Kriging and our ensemble surrogate are plotted in Fig. A.1 due to page limitations, when solving the 20-D WFG4, DTLZ6, UF2, and MaF1 problems. The predicted objective values by one global Kriging are nearly the same as analyzed in Section II.B, while our ensemble surrogate can predict more accurate and diversified objective values, which can well explain the better performance of our ensemble surrogate for solving these 20-D EMOPs. After exploiting information on the low-dimensional subspaces that are extracted from the original decision space, Kriging sub-models constructed in different low-dimensional subspaces can be used to compose our ensemble surrogate model. The pairwise comparison results in Table II also indicate that using only one sub-model in the ensemble surrogate is still not enough to effectively capture information in low-dimensional subspaces.

Moreover, their average running times (in seconds: s) from 30 runs are provided in Table A.VII of the Supplementary Material due to page limitations. The running time of ESF-RVEA will be highly increased with the number of

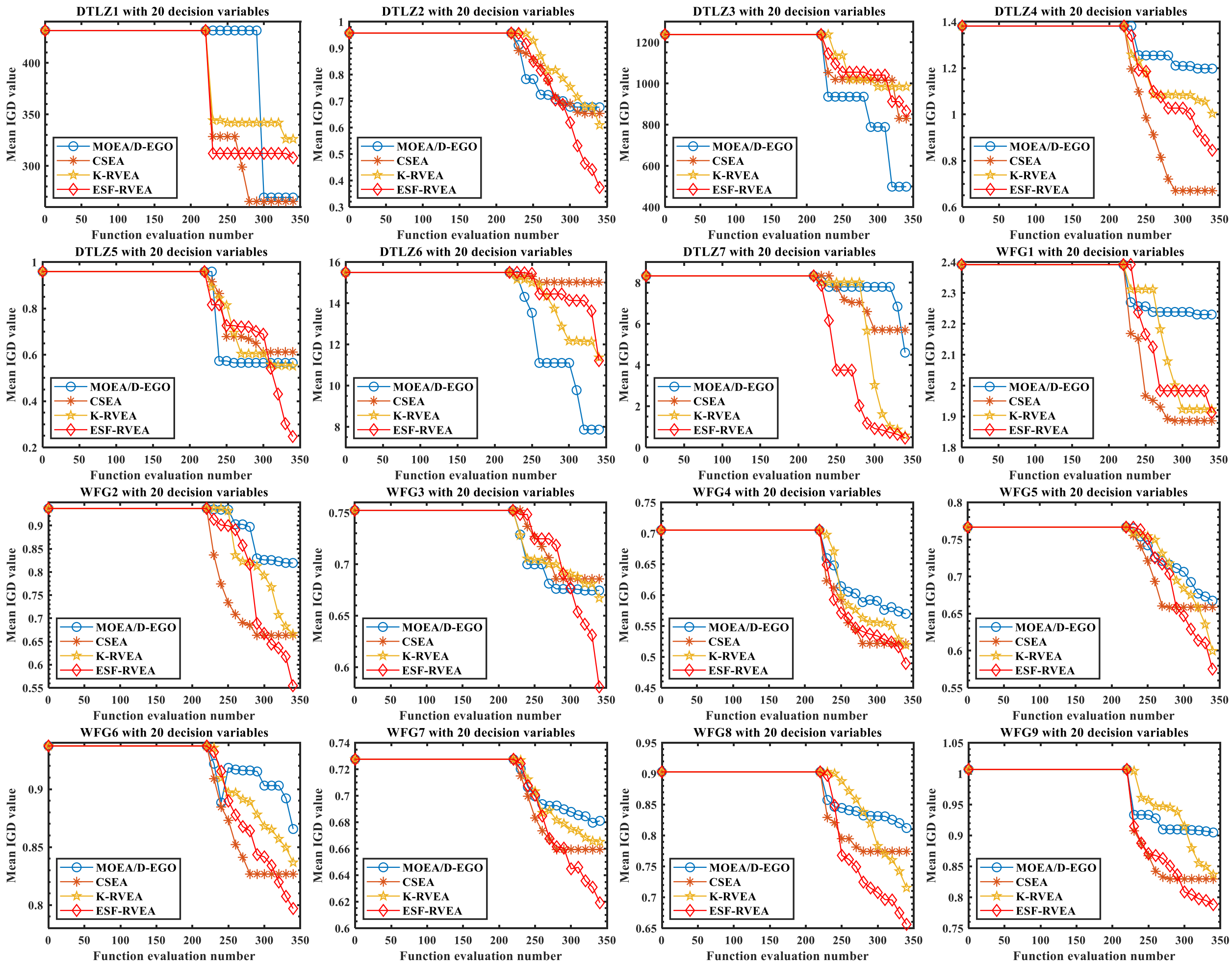


Fig. 5 Convergence profiles of all the compared algorithms on the DTLZ and WFG test problems with $D = 20$

surrogate sub-models. This is because training one Kriging model is very time consuming. Thus, $k = 2$ is suggested for the proposed ESF when considering both the performance and time cost for solving all the given EMOPs.

D. *Comparisons with Three Competitive SAEAs*

To validate the advantage of ESF, three competitive SAEAs (MOEA/D-EGO [42], CSEA [20], and K-RVEA [16]) are also included for comparison in this subsection. Here, ESF-RVEA is selected as our representative algorithm due to its superior performance in Section IV.B. All the compared algorithms are implemented in PlatEMO [59] and their parameters are set as suggested in their corresponding references [42], [20], [16]. The initial population size and the maximum number of true expensive evaluations are set the same in Section IV.B, while the settings of other parameters for these compared algorithms are also introduced in Section 2 of the Supplementary Material due to page limitations.

Please note that in the following Tables III-IV, the best result is highlighted for each test case. Moreover, to obtain a statistically sound conclusion, Wilcoxon rank-sum tests are run with a significance level $\alpha = 0.05$, showing the statistically significant difference on the results of ESF-RVEA and other algorithms. The symbols '+', '-', and '=' in Tables III-IV indicate that the results of other algorithms are significantly better than, worse than, and similar to those of ESF-RVEA, respectively, using this statistical test.

(1) Comparison Results on the DTLZ and WFG Problems

Here, the performances of all the algorithms are compared on all the DTLZ and WFG test problems. Their IGD values under 30 independent runs on the DTLZ and WFG test problems are given in Table III. As observed from Table III, ESF-RVEA performs significantly better than MOEA/D-EGO, CSEA, and K-RVEA in 36/48 ≈ 75%, 31/48 ≈ 65%, and 33/48 ≈ 69% of the DTLZ and WFG test problems, respectively. The smallest and largest percentages are 65% and 75%, respectively. On the other hand, the largest percentage of the test problems in which ESF-RVEA is significantly worse than the peer algorithms is found in the pairwise comparison of ESF-RVEA and MOEA/D-EGO (only 8/48 ≈ 17%). On these IGD results, ESF-RVEA achieves the best results on 29 cases, followed by CSEA (performing best on 9 cases). Therefore, it is reasonable to conclude that our ESF-RVEA performs best in most cases of the DTLZ and WFG test problems.

For DTLZ1 and DTLZ3, due to their multifrontal landscapes, all the compared algorithms are unable to find a set of well-converged final solutions using a small number of func-

TABLE IV
COMPARISON OF ESF-RVEA WITH THREE COMPETITIVE ALGORITHMS ON THE UF AND MAF TEST SUITES

Problem	M	D	MOEA/D-EGO	CSEA	K-RVEA	ESF-RVEA
UF1	2	10	2.7278e-1 (7.48e-2) -	3.4553e-1 (1.32e-1) -	1.5031e-1 (5.50e-2) =	**1.4202e-1 (5.50e-2)**
	2	20	7.4861e-1 (1.89e-1) -	7.8035e-1 (1.72e-1) -	2.9151e-1 (9.55e-2) -	**2.0461e-1 (5.04e-2)**
	2	30	1.0508e+0 (1.66e-1) -	9.9889e-1 (1.12e-1) -	5.7897e-1 (2.22e-1) -	**2.7109e-1 (6.59e-2)**
UF2	2	10	2.1482e-1 (3.50e-2) -	2.3913e-1 (5.41e-2) -	9.5903e-2 (3.33e-2) -	**7.6867e-2 (1.28e-2)**
	2	20	3.4161e-1 (5.31e-2) -	3.7459e-1 (5.40e-2) -	1.4438e-1 (1.85e-2) =	**1.3606e-1 (3.49e-2)**
	2	30	4.4161e-1 (6.13e-2) -	4.5033e-1 (5.13e-2) -	2.0171e-1 (2.45e-2) -	**1.8464e-1 (3.74e-2)**
UF3	2	10	1.0510e+0 (1.08e-1) -	1.2396e+0 (1.69e-1) -	9.7852e-1 (7.77e-2) -	**7.8337e-1 (2.25e-1)**
	2	20	8.5519e-1 (7.04e-2) -	9.5419e-1 (9.89e-2) -	**6.3997e-1 (3.37e-2) =**	6.4006e-1 (2.90e-2)
	2	30	7.4706e-1 (4.92e-2) -	8.9654e-1 (8.99e-2) -	5.4058e-1 (3.01e-2) -	**4.8766e-1 (1.97e-2)**
UF4	2	10	1.1615e-1 (8.89e-3) -	1.3747e-1 (8.51e-3) -	1.1363e-1 (7.97e-3) -	**1.0678e-1 (1.29e-2)**
	2	20	**1.3703e-1 (8.00e-3) +**	1.6318e-1 (5.62e-3) -	1.5547e-1 (8.38e-3) -	1.4254e-1 (1.05e-2)
	2	30	**1.5434e-1 (6.40e-3) +**	1.7338e-1 (5.41e-3) -	1.7386e-1 (2.30e-3) -	1.6633e-1 (4.90e-3)
UF5	2	10	2.8402e+0 (5.66e-1) -	2.5795e+0 (5.96e-1) -	1.7329e+0 (7.82e-1) =	**1.8017e+0 (6.52e-1)**
	2	20	4.1771e+0 (4.02e-1) -	3.6890e+0 (5.04e-1) -	3.0166e+0 (6.68e-1) =	**2.6316e+0 (7.60e-1)**
	2	30	4.9329e+0 (2.59e-1) -	4.3849e+0 (3.91e-1) -	3.7921e+0 (7.50e-1) -	**3.3433e+0 (5.16e-1)**
UF6	2	10	1.9530e+0 (3.93e-1) -	2.5298e+0 (7.33e-1) -	**1.2546e+0 (2.11e-1) =**	1.2708e+0 (2.24e-1)
	2	20	3.2389e+0 (7.89e-1) -	3.3705e+0 (7.38e-1) -	1.1435e+0 (5.17e-1) -	**8.3002e-1 (1.46e-1)**
	2	30	4.3114e+0 (8.35e-1) -	4.2604e+0 (1.02e+0) -	1.5144e+0 (3.87e-1) -	**9.8571e-1 (4.55e-1)**
UF7	2	10	3.6686e-1 (8.77e-2) -	4.3823e-1 (1.06e-1) -	2.9954e-1 (1.12e-1) =	**2.5111e-1 (1.18e-1)**
	2	20	7.8611e-1 (1.79e-1) -	7.0659e-1 (1.74e-1) -	4.4981e-1 (1.85e-1) =	**3.8462e-1 (1.41e-1)**
	2	30	1.0375e+0 (1.91e-1) -	9.1260e-1 (1.87e-1) -	5.7598e-1 (1.36e-1) -	**4.5890e-1 (1.00e-1)**
UF8	3	10	5.2245e-1 (1.12e-1) -	7.9704e-1 (2.84e-1) -	3.2743e-1 (4.03e-2) =	**3.1099e-1 (4.34e-2)**
	3	20	9.9878e-1 (1.88e-1) -	1.4455e+0 (2.91e-1) -	4.5222e-1 (6.53e-2) -	**3.7932e-1 (6.59e-2)**
	3	30	1.4706e+0 (3.36e-1) -	1.7279e+0 (2.48e-1) -	6.7635e-1 (4.39e-1) -	**4.5022e-1 (1.04e-1)**
UF9	3	10	5.3379e-1 (9.64e-2) -	8.8003e-1 (2.68e-1) -	**2.9674e-1 (6.01e-2) +**	3.8474e-1 (6.86e-2)
	3	20	1.0213e+0 (1.76e-1) -	1.3969e+0 (2.27e-1) -	4.3197e-1 (6.93e-2) =	**4.3117e-1 (6.96e-2)**
	3	30	1.7001e+0 (4.14e-1) -	1.9671e+0 (2.80e-1) -	5.4669e-1 (7.93e-2) -	**4.5595e-1 (4.74e-2)**
MaF1	3	10	3.3513e-1 (4.13e-2) -	2.0524e-1 (3.41e-2) -	**7.4349e-2 (5.08e-3) +**	8.5563e-2 (8.72e-3)
	3	20	6.4422e-1 (8.19e-2) -	6.1646e-1 (8.91e-2) -	2.6381e-1 (7.49e-2) -	**1.8684e-1 (3.72e-2)**
	3	30	9.1125e-1 (1.28e-1) -	1.1184e+0 (1.24e-1) -	8.2419e-1 (1.38e-1) -	**5.0775e-1 (8.80e-2)**
MaF2	3	10	5.6373e-2 (1.69e-3) -	5.4802e-2 (1.91e-3) -	**4.0740e-2 (1.17e-3) +**	4.2768e-2 (2.44e-3)
	3	20	8.7882e-2 (3.46e-3) -	8.6074e-2 (4.94e-3) -	7.1628e-2 (5.09e-3) -	**6.1732e-2 (4.95e-3)**
	3	30	1.1982e-1 (5.52e-3) -	1.2893e-1 (6.04e-3) -	1.2046e-1 (4.43e-3) -	**1.0605e-1 (9.89e-3)**
MaF3	3	10	3.2616e+5 (1.69e+5) =	3.4237e+5 (3.23e+5) =	4.3343e+5 (2.39e+5) =	**3.1825e+5 (1.68e+5)**
	3	20	2.4977e+6 (7.93e+5) =	**2.3430e+6 (9.53e+5) =**	2.7370e+6 (6.81e+5) -	2.4029e+6 (8.64e+5)
	3	30	7.1189e+6 (1.97e+6) =	6.4106e+6 (1.25e+6) =	**6.3241e+6 (1.22e+6) =**	6.4728e+6 (1.63e+6)
MaF4	3	10	9.6107e+2 (2.80e+2) =	**7.7849e+2 (1.88e+2) =**	8.9087e+2 (2.96e+2) =	9.1464e+2 (2.72e+2)
	3	20	3.5512e+3 (5.08e+2) -	**2.7228e+3 (4.90e+2) =**	3.0626e+3 (5.06e+2) =	2.9352e+3 (5.12e+2)
	3	30	6.2049e+3 (7.44e+2) -	**5.0310e+3 (6.07e+2) +**	5.8491e+3 (3.98e+2) -	5.3658e+3 (4.72e+2)
MaF5	3	10	2.5552e+0 (3.01e-1) -	2.1602e+0 (4.90e-1) =	**1.8068e+0 (3.66e-1) =**	1.8608e+0 (5.27e-1)
	3	20	4.6053e+0 (5.98e-1) -	**3.0733e+0 (5.92e-1) =**	3.8649e+0 (4.94e-1) -	3.2897e+0 (7.05e-1)
	3	30	6.3719e+0 (8.38e-1) -	**4.5042e+0 (5.97e-1) =**	6.2925e+0 (8.12e-1) -	4.8174e+0 (7.40e-1)
MaF6	3	10	6.0626e+0 (2.13e+0) -	8.6929e+0 (2.95e+0) -	1.1787e+0 (5.10e-1) -	**6.2066e-1 (3.60e-1)**
	3	20	2.4714e+1 (8.11e+0) -	4.0783e+1 (8.63e+0) -	2.0515e+1 (7.16e+0) -	**1.4100e+1 (4.43e+0)**
	3	30	4.3649e+1 (1.17e+1) =	8.2726e+1 (1.04e+1) -	6.8562e+1 (1.20e+1) -	**3.8403e+1 (8.72e+0)**
MaF7	3	10	3.0561e-1 (8.23e-2) =	2.4410e+0 (8.86e-1) -	**1.9475e-1 (6.75e-2) +**	4.5442e-1 (3.09e-1)
	3	20	4.3622e+0 (1.63e+0) -	4.6561e+0 (1.05e+0) -	7.6811e-1 (1.47e+0) =	**3.7486e-1 (1.43e-1)**
	3	30	7.7243e+0 (1.38e+0) -	5.7454e+0 (9.30e-1) -	4.2723e+0 (3.45e+0) -	**8.3953e-1 (2.36e-1)**
+/-/=			2/40/6	1/39/8	4/28/16	--

tion evaluations. ESF-RVEA outperforms only K-RVEA on the 20-D and 30-D DTLZ3 and MOEA/D-EGO on the 30-D DTLZ1. CSEA is the best algorithm for solving DTLZ1 and DTLZ3, as it achieves the best results on all the cases of DTLZ1 and the 10-D DTLZ3. On DTLZ2 and DTLZ4, ESF-RVEA performs better than or similarly to MOEA/D-EGO, CSEA, and K-RVEA on DTLZ2 but it is outperformed by CSEA and K-RVEA on the 20-D and 30-D DTLZ4 and the 10-D DTLZ4, respectively. ESF-RVEA achieves the best results on the 10-D and 20-D DTLZ2, while CSEA is the best at handling the 20-D and 30-D DTLZ4. Regarding DTLZ5 and DTLZ6, due to their degenerate PFs, all the compared algorithms have difficulties to approximate the true PF. As observed from the IGD results in Table III, ESF-RVEA obtains the best results on all the cases of DTLZ5, while MOEA/D-EGO performs best on all the cases of DTLZ6. As DTLZ7 is a discontinuous problem, all the algorithms encounter the challenge of maintaining diversity. ESF-RVEA shows some advantages on this problem, as it is only defeated by K-RVEA on the 10-D DTLZ7.

From the IGD results of Table III, we can observe that for WFG1 with a convex and biased PF, although ESF-RVEA fails to perform best in any case, it achieves better or similar performance in most of the pairwise comparisons. Regarding WFG2 with a mixed and disconnected PF, ESF-RVEA shows the best performance for all the numbers of dimensions, while only CSEA can show a similar performance in the 30-D case of WFG2. To solve WFG3 with a linear and unimodal PF, ESF-RVEA performs best in most cases and is outperformed only by MOEA/D-EGO in its 30-D case. Concerning WFG4 to WFG8 with concave PFs, ESF-RVEA shows superior performance over other compared algorithms, as it performs best in most cases. For WFG9 with a deceptive and multimodal PF, ESF-RVEA also shows advantages and only CSEA is able to obtain a similar IGD value to ESF-RVEA.

To examine the convergence speed of the four algorithms, their mean IGD values versus the function evaluation numbers over 30 independent runs are plotted in Fig. 5, where the DTLZ and WFG problems with 20 decision variables are used as representative cases. It can be observed from Fig. 5 that the

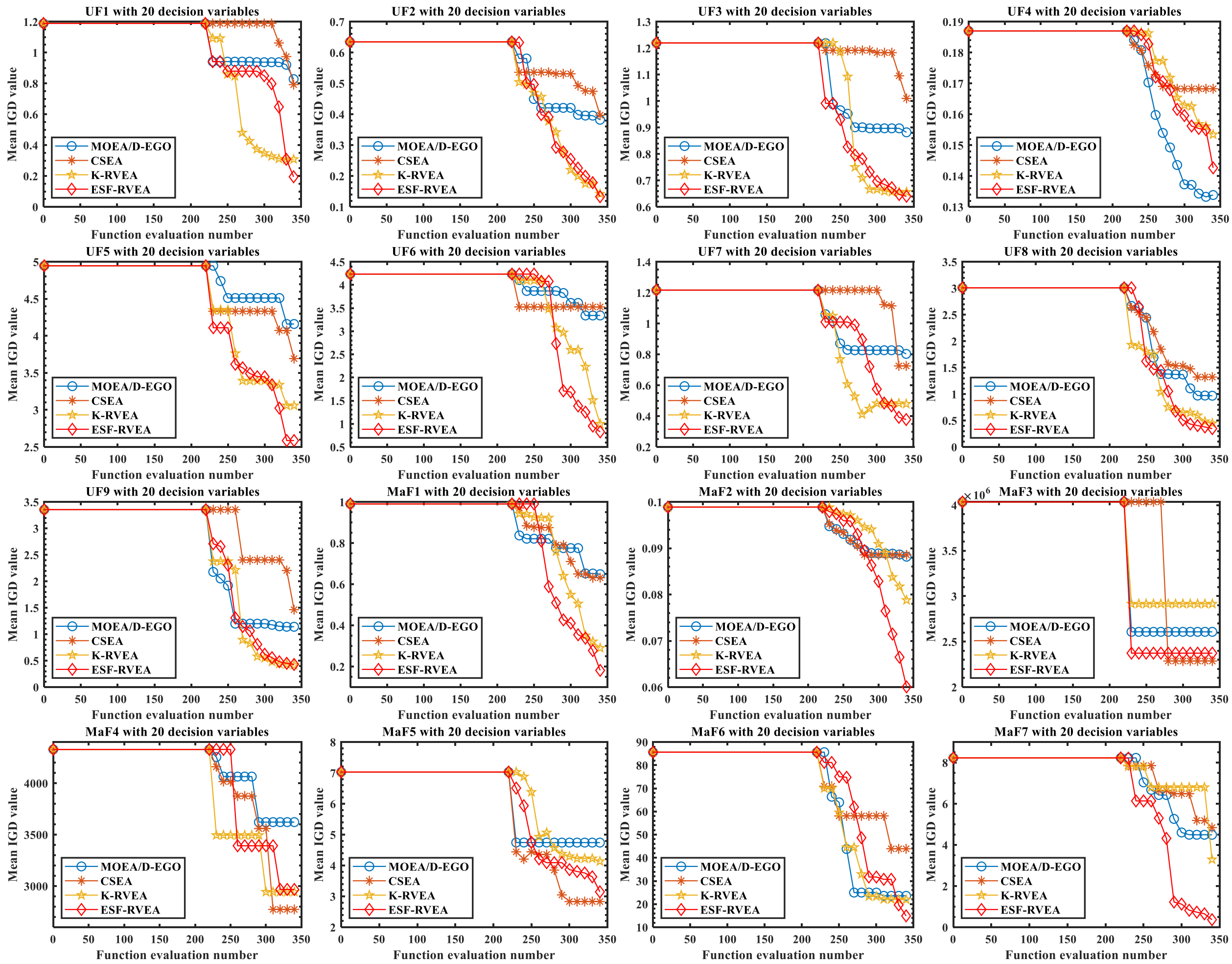


Fig. 6 Convergence profiles of all the compared algorithms on the UF and MaF test suites with $D = 20$

convergence curve of the mean IGD values achieved by ESF-RVEA shows the fastest convergence speed in most test cases, including DTLZ2, DTLZ5, DTLZ7, and WFG1-WFG9. Even for DTLZ4 and DTLZ6, ESF-RVEA still obtains the second lowest mean IGD values. The promising convergence speed of ESF-RVEA is mainly attributed to the use of ESF to accurately approximate objective functions.

To visually show the optimization performance, Fig. A.2 plots the final nondominated solutions obtained by each algorithm with the median IGD values over 30 independent runs on several representative 20-D DTLZ and WFG problems in the Supplementary Material due to page limitations. Fig. A.2 shows that the final solutions of DTLZ7 obtained by all the compared algorithms do not have good convergence to the true PF, while ESF-RVEA is able to achieve a better approximation than the other algorithms. For WFG2, which is plotted in Fig. A.2, the final solutions obtained by all the algorithms have good convergence but cannot evenly spread over the whole PF. However, the final solutions obtained by ESF-RVEA show a better diversity than the other algorithms. When solving WFG4 in Fig. A.2, ESF-RVEA also outperforms MOEA/D-EGO, CSEA, and K-RVEA in terms of both the convergence and diversity according to their final solutions.

(2) Comparison Results on the UF and MaF Problems

The performances of all the algorithms are also compared on all the UF and MaF test problems. The IGD mean values under 30 independent runs on the UF and MaF test problems with 10 to 30 decision variables are provided in Table IV.

As observed from Table IV, ESF-RVEA also performs best in most cases (32 out of 48 cases), while MOEA/D-EGO, CSEA, and K-RVEA perform best in 2, 6, and 8 cases, respectively. Considering the pairwise comparison, ESF-RVEA significantly performs better than MOEA/D-EGO, CSEA, and K-RVEA in 40, 39, and 28 out of 48 cases, respectively, whereas it is worse than MOEA/D-EGO, CSEA, and K-RVEA in 2, 1, and 4 cases, respectively. K-RVEA is the most competitive algorithm to ESF-RVEA, as it can obtain statistically similar results in 16 cases. To be more specific, for the UF test problems, ESF-RVEA is outperformed by K-RVEA on the 10-D UF6 and UF9, and by MOEA/D-EGO on the 30-D UF4, while K-RVEA is slightly better than ESF-RVEA on the 20-D UF3. CSEA performs worse than ESF-RVEA on all the UF test cases. Considering the MaF test suite, ESF-RVEA achieves superior performance in most cases, including the 20-D and 30-D MaF1, the 10-D MaF3, and all the cases of MaF2, MaF6, MaF7. MOEA/D-EGO obtains statistically similar IGD results with ESF-RVEA on only the 10-D and 20-D MaF3, the 10-D MaF4, the 30-D MaF6, and the 10-D MaF7. K-RVEA only achieves significantly better IGD result on the 10-D MaF7.

To further study the performance of ESF-RVEA, the convergence profiles of all the compared algorithms are depicted in Fig. 6 for the UF and MaF cases with 20 decision variables. It should be noted that all the IGD values in Fig. 6 are calculated using all the current solutions that have been truly evaluated. From these profiles, it can be observed that ESF-RVEA outperforms MOEA/D-EGO, CSEA, and K-RVEA in most cases. More specifically, except for UF2, UF3, UF9, and MaF4, ESF-RVEA is the best algorithm, as it is able to achieve the smallest mean IGD results. ESF-RVEA achieves a significantly better performance on UF1, UF4, UF7-UF8, MaF1-MaF3, and MaF5-MaF6, while it shows a competitive performance on UF2-UF3, UF6, UF9, MaF4, and MaF7 according to its convergence trends in Fig. 6.

In addition, the final nondominated solutions achieved by each compared algorithm are also plotted in Fig. A.3 of the Supplementary Material due to page limitations, when solving UF1, UF4, UF7, MaF1, and MaF6 with 20 decision variables in the run associated with the median IGD value. As observed from Fig. A.3, only K-RVEA and ESF-RVEA are capable of performing well in approximating the true PF, while the final solutions yielded by MOEA/D-EGO and CSEA are far from the true PFs of UF1 and UF7. For UF4, ESF-RVEA obtains the final solutions with better convergence, while other compared algorithms show a poor approximation performance. On MaF1, the final solutions of ESF-RVEA show the best diversity. With respect to MaF6, which is a hard problem to tackle, all the algorithms can only search poorly converged solutions with very large values in the three-objective objective space. Nevertheless, ESF-RVEA still performs better in terms of the convergence towards the PF.

E. More Discussions

1) Comparisons with Ensemble-based SAEAs

In recent years, there have been several ensemble-based SAEAs designed for tackling EMOPs [25]-[28], but our ensemble model is composed by one global surrogate model and a number of k surrogate sub-models, which is different from these previous ensemble strategies. In this subsection, four ensemble-based SAEAs, i.e., TEMO-MPS [25], GCS-MOE [26], HeE-MOEA [27], and SAEMO [28], are included for performance comparison with ESF-RVEA for solving all the adopted test problems. Due to page limitations, their IGD results are provided in Table A.VIII of the Supplementary Material. Table IV summarizes their pairwise comparison results based on the IGD results in Table A.VIII.

As observed from Table IV, ESF-RVEA shows advantages for tackling the 20-D UF, MaF, DTLZ, and WFG test problems. To be specific, the performance of ESF-RVEA is better than that of TEMO-MPS, GCS-MOE, HeE-MOEA and SAEMO in 20, 31, 24, and 21 out of 32 cases, respectively, while TEMO-MPS, GCS-MOE, HeE-MOEA, and SAEMO significantly outperform ESF-RVEA in only 4, 1, 2, and 3 cases, respectively. Therefore, the advantages of our ensemble model over other existing ensemble models are validated.

TABLE V
SUMMARY OF COMPARISON OF ESF-RVEA AND FIVE SAEAS

ESF-RVEA vs.		TEMO-MPS	GCS-MOE	HeE-MOEA	SAEMO	ESF-RVEA-I
DTLZ, WFG, UF, and MaF (D = 20)	Better	20	31	24	21	21
	Worse	4	1	2	3	1
	Similar	8	0	6	8	10
Best	18	4	0	3	4	3

TABLE VI
SUMMARY OF COMPARISON OF ESF-RVEA WITH DIFFERENT c VALUES

ESF-RVEA vs.		c = 0.2	c = 0.4	c = 0.6	c = 0.8	c = rand
DTLZ, WFG, UF, and MaF (D = 20)	Better	23	22	20	17	23
	Worse	1	1	1	2	1
	Similar	8	9	11	13	8
Best	18	2	3	4	4	1

2) Effectiveness of the Global Surrogate Model

In order to study the contribution of the global surrogate model, more experiments are run to compare ESF-RVEA with its variant without using the global surrogate model (called ESF-RVEA-I). Due to page limitations, their IGD results are provided in Table A.VIII of the Supplementary Material. Table V summarizes their pairwise comparison results. It can be observed from Table V that ESF-RVEA shows distinct advantages when solving the 20-D cases of all the UF, MaF, DTLZ, and WFG test problems. ESF-RVEA performs significantly better than or similarly to ESF-RVEA-I in 31 out of 32 instances, while ESF-RVEA-I only performs better in the 20-D MaF3. Therefore, the global surrogate model is validated to have a significant contribution to the performance of the ESF.

3) Effectiveness of the Adaptive Approach

To further verify the effectiveness of our adaptive approach applied to change the weights of each member model in (2), four variants of ESF-RVEA with different fixed values for the weight factor c are realized to replace the adaptive approach in ESF-RVEA. For the sake of simplicity, these ESF-RVEA variants are defined by the fixed values of c, i.e., c = 0.2, 0.4, 0.6, and 0.8. Additionally, a variant by setting c to a random variable from 0 to 1, called c = rand, is also included for this comparison. To ensure a fair comparison, other components of these variants are kept the same as in the original ESF-RVEA.

Due to page limitations, Table A.IX of the Supplementary Material provides the IGD results yielded by ESF-RVEA and the five variants with different c values from {0.2, 0.4, 0.6, 0.8} and a random value of c over 30 independent runs for solving all the 20-D DTLZ, WFG, UF and MaF test problems. Table VI summarizes the pairwise comparison results of ESF-RVEA with the five variants. As observed from Table VI, our adaptive approach used in ESF-RVEA is reasonable and more effective than the variants, as it achieves the best IGD results in more than half of the cases, i.e., in 18 out of 32 cases. The five variants using a constant c value obtain the best IGD results on the remaining 14 cases. Specifically, the five variants with different c values (0.2, 0.4, 0.6, 0.8, and c = rand) obtain the best results on 2, 3, 4, 4, and 1 cases, respectively. As observed from Table VI, for all 32 cases, ESF-RVEA performs better than or similarly to the c = 0.2, c = 0.4, c = 0.6 and c = rand variants on 31 cases and the c = 0.8 variant on 30 cases. Thus, the effectiveness of the adaptive approach for

TABLE VII
SUMMARY OF COMPARISON OF ESF-RVEA WITH DIFFERENT INFILL SOLUTIONS SELECTION STRATEGIES AND SURROGATE MODELS

ESF-RVEA vs.		ExI	LCB	PoI	RBF	PR
DTLZ, WFG, UF, and MaF (D = 20)	Better	22	23	25	18	29
	Worse	1	1	1	7	1
	Similar	9	8	6	7	2
Best	19	1	2	0	9	1

TABLE VIII
SUMMARY OF COMPARISON OF ESF-RVEA WITH DIFFERENT N_{is} VALUES

ESF-RVEA (N_{is} = 5) vs.		N_{is} = 1	N_{is} = 3	N_{is} = 7	N_{is} = 9
DTLZ, WFG, UF, and MaF (D = 20)	Better	0	1	3	8
	Worse	6	3	0	0
	Similar	26	28	29	24
Best	2	24	3	0	3

updating the weights in the proposed ESF is validated by these experiments.

4) Effectiveness of the Infill Solutions Selection

To further verify the effectiveness of our infill solutions selection introduced in Section III.C, three popular infill solutions selection strategies (ExI [35], LCB [36] and PoI [34]) are embedded into ESF-RVEA, giving its three variants identified as ExI, LCB and PoI in this paper for the sake of simplicity.

Due to page limitations, their IGD results are provided in Table A.X of the Supplementary Material for solving all the 20-D DTLZ, WFG, UF and MaF problems. Table VII summarizes the pairwise comparison results of ESF-RVEA with ExI, LCB and PoI. As observed from Table VII, our infill solutions selection in ESF-RVEA shows superior performance in most cases, as it is better than ExI, LCB, and PoI in 22, 23, and 25 out of 32 cases, respectively, while ExI, LCB, and PoI are better than our method in only 1 case. As ExI is originally designed for solving expensive single-objective optimization problems, while ESF-RVEA does not use any strategy to decompose EMOPs, it performs not so well in ESF-RVEA. Thus, the proposed infill solutions selection is more effective as validated by these experiments.

5) Advantages of Using Kriging in ESF

To further study the advantages of using Kriging in ESF, two popular surrogate models used in SAESs (RBF [18] and PR [23]) are embedded into ESF for performance comparison, giving its two variants marked as RBF and PR in this paper for the sake of simplicity.

Due to page limitations, their IGD results are provided in Table A.X of the Supplementary Material for solving all the 20-D DTLZ, WFG, UF and MaF problems. Table VII summarizes the pairwise comparison results of ESF-RVEA with RBF and PR. As observed from Table VII, ESF-RVEA with Kriging performs better than PBF and PR in 18 and 29 out of 32 cases, respectively, while it is underperformed by RBF and PR in 7 and 1 cases, respectively. Therefore, the advantages of using Kriging in ESF are confirmed by these experiments.

6) Sensitivity Analysis of Parameter N_{is} in ESF

To further analyze the impact of the maximum number (N_{is}) of selected infill solutions in ESF, ESF-RVEA using different values of N_{is} (i.e., N_{is} = 1, 3, 5, 7, and 9) are compared when solving 20-D UF, MaF, WFG, and DTLZ test suites. Due to page limitations, their IGD comparison results are provided in Table A.XI of the Supplementary Material and Table VIII summarizes their pairwise comparison results. From Table VIII, the overall performance of ESF-RVEA is not so sensitive to the setting of N_{is} on 20-D UF, MaF, WFG, and DTLZ test suites, as ESF-RVEA with N_{is} = 5 performs similarly to that with N_{is} =1, 3, 7, and 9, on 26, 28, 29, and 24 out of total 32 problems, respectively. Moreover, it is observed that a smaller value of N_{is} will generally have a slightly better overall performance. This is reasonable as a small number of infill solutions used in each generation of **Algorithm 1** will lead to more running of surrogate-assisted-MOEA, which will find a more promising approximation to the PF. However, a smaller value of N_{is} will also result in a significantly increased computational cost, as the times of training surrogates will be increased significantly. Thus, in this paper, the value of N_{is} is fixed as 5 to balance the performance and computational cost.

V. CONCLUSIONS AND FUTURE WORK

The selection of surrogate models is important for enhancing the performance of an SAEA. Although existing surrogate models have been shown to be particularly competitive on low-dimensional EMOPs, they will face challenges for solving EMOPs with more than ten dimensions. To solve this problem, this paper suggests a novel ESF for expensive multiobjective evolutionary optimization. In this framework, a novel and effective ensemble surrogate is designed to approximate the expensive objective values, which include one global Kriging model trained under the entire search space and k Kriging sub-models trained under different low-dimensional subspaces. Moreover, a new infill selection criterion is proposed based on a set of reference vectors to select promising solutions for updating this ensemble surrogate model.

The performance of our framework has been verified by solving four scalable test suites (DTLZ, WFG, UF, and MaF) when three representative MOEAs (NSGA-III, MOEA/D-DE, RVEA) are embedded into our framework. The experimental results have validated the effectiveness of our framework. Moreover, ESF-RVEA performs significantly better than three competitive SAEAs (MOEA/D-EGO, CSEA, and K-RVEA) and four ensemble-based SAEAs (TEMO-MPS, GCS-MOE, SAEMO, and HeE-MOEA) in most cases. Also, the parameter sensitivity analysis on the number of surrogate sub-models and the maximum number of selected infill solutions are also experimentally studied in this paper.

Our ensemble surrogate based framework works well on the used test EMOPs with up to 30 dimensions, as it can provide more accurate and diversified predicted objective values. Its performance on the EMOPs with more than 30 dimensions will be further studied in the future work. Moreover, as benchmarks may not sufficiently represent the features of practical applications, the performance of ESF for tackling some real-world EMOPS will be also studied in the future work.

REFERENCES

[1] K. Deb, "Multi-objective optimization using evolutionary algorithms [M]," John Wiley & Sons, 2001.

[2] H. R. Maier, S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson, "Introductory overview: Optimization using evolutionary algo-

rithms and other metaheuristics," Environmental Modelling & Software, vol. 114, pp. 195-213, 2019.

[3] 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 Transactions on Evolutionary Computation, vol. 18, no. 4, pp. 577-601, 2013.

[4] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," TIK-report, vol. 103, pp. 95–100, 2001.

[5] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," IEEE Transactions on Evolutionary Computation, vol. 13, no. 2, pp. 284-302, 2008.

[6] H. L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," IEEE Transactions on Evolutionary Computation, vol. 18, no. 3, pp. 450–455, Jun. 2014.

[7] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in International Conference on Parallel Problem Solving from Nature, Birmingham, U.K., pp. 832–842, 2004.

[8] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," European Journal of Operational Research, vol. 181, no. 3, pp. 1653–1669, 2007.

[9] S. Koziel and A. Bekasiewicz, "Rapid simulation-driven multiobjective design optimization of decomposable compact microwave passives," IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 8, pp. 2454–2461, Aug. 2016.

[10]D. Guo, T. Chai, J. Ding, and Y. Jin, "Small data driven evolutionary multi-objective optimization of fused magnesium furnaces," in 2016 IEEE Symposium Series on Computational Intelligence, Athens, Greece, Dec. 2016, pp. 1–8.

[11]H. Wang, J. Doherty, and Y. Jin, "Hierarchical surrogate-assisted evolutionary multi-scenario airfoil shape optimization," in 2016 IEEE Symposium Series on Computational Intelligence, 2018, pp. 1–8.

[12]Y. Jin, "Surrogate-assisted evolutionary computation: Recent advances and future challenges," Swarm and Evolutionary Computation, vol. 1, no. 2, pp. 61-70, 2011.

[13]Y. Jin, H. Wang, T. Chugh, D. Guo, and K. Miettinen, "Data-driven evolutionary optimization: An overview and case studies," IEEE Transactions on Evolutionary Computation, vol. 23, no. 3, pp. 442–458, Jun. 2019.

[14]Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, and K. Y. Lum, "Combining global and local surrogate models to accelerate evolutionary optimization," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, no. 1, pp. 66-76, 2006.

[15]D. Lim, Y. Jin, Y.-S. Ong, and B. Sendhoff, "Generalizing surrogateassisted evolutionary computation," IEEE Transactions on Evolutionary Computation, vol. 14, no. 3, pp. 329–355, Jun. 2010.

[16]T. Chugh, Y. Jin, K. Miettinen, J. Hakanen, and K. Sindhya, "A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization," IEEE Transactions on Evolutionary Computation, vol. 22, no. 1, pp. 129–142, Feb. 2018.

[17]T. Chugh, C. Sun, H. Wang, and Y. Jin, "Surrogate-assisted evolutionary optimization of large problems," In High-Performance Simulation-Based Optimization, pp. 165-187, 2020.

[18]A. Isaacs, T. Ray, and W. Smith, "An evolutionary algorithm with spatially distributed surrogates for multiobjective optimization," in Australian Conference on Artificial Life. Springer, pp. 257–268, 2007.

[19]S. Z. Martínez and C. A. C. Coello, "MOEA/D assisted by RBF networks for expensive multi-objective optimization problems," in Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation, Amsterdam, The Netherlands, pp. 1405–1412, 2013.

[20]L. Pan, C. He, Y. Tian, H. Wang, X. Zhang, and Y. Jin, "A classification-based surrogate-assisted evolutionary algorithm for expensive many-objective optimization," IEEE Transactions on Evolutionary Computation, vol. 23 no. 1, pp. 74-88, 2018.

[21]W. Kong, T. Chai, S. Yang, and J. Ding, "A hybrid evolutionary multiobjective optimization strategy for the dynamic power supply problem in magnesia grain manufacturing," Applied Soft Computing, vol. 13, no. 5, pp. 2960–2969, 2013.

[22]M. Herrera, A. Guglielmetti, M. Xiao, and R. F. Coelho, "Metamodelassisted optimization based on multiple kernel regression for mixed variables," Structural and Multidisciplinary Optimization, vol. 49, no. 6, pp. 979–991, 2014.

[23]Y. Lian and M.-S. Liou, "Multiobjective optimization using coupled response surface model and evolutionary algorithm," AIAA Journal, vol. 43, no. 6, pp. 1316–1325, 2005.

[24]A. Rosales-Pérez, C. A. C. Coello, J. A. Gonzalez, C. A. Reyes-Garcia, and H. J. Escalante, "A hybrid surrogate-based approach for evolutionary multi-objective optimization," in IEEE Congress on Evolutionary Computation, Cancún, Mexico, pp. 2548–2555, 2013.

[25]A. T. W. Min, Y. S. Ong, A. Gupta, and C. K. Goh, "Multiproblem surrogates: Transfer evolutionary multiobjective optimization of computationally expensive problems," IEEE Transactions on Evolutionary Computation, vol. 23, no. 1, pp. 15-28, 2017.

[26]J. Luo, A. Gupta, Y. S. Ong, and Z. Wang, "Evolutionary optimization of expensive multiobjective problems with co-sub-Pareto front Gaussian process surrogates," IEEE Transactions on Cybernetics, vol. 49, no. 5, pp. 1708-1721, 2018.

[27]D, Guo, Y. Jin, J. Ding, and T. Chai, "Heterogeneous ensemble-based infill criterion for evolutionary multiobjective optimization of expensive problems," IEEE Transactions on Cybernetics, vol. 49, no. 3, pp. 1012-1025, 2018.

[28]Y. Zhao, C. Sun, J. Zeng, Y. Tan, and G. Zhang, "A surrogate-ensemble assisted expensive many-objective optimization," Knowledge-Based Systems, 2021.

[29]N. Liu, J. S. Pan, C. Sun, and S. C. Chu, "An efficient surrogate-assisted quasi-affine transformation evolutionary algorithm for expensive optimization problems," Knowledge-Based Systems, no. 209, 2020.

[30]K. S. Bhattacharjee, A. Isaacs, and T. Ray, "Multi-objective optimization using an evolutionary algorithm embedded with multiple spatially distributed surrogates," In MULTI-OBJECTIVE OPTIMIZATION: Techniques and Application in Chemical Engineering, pp. 135-155, 2017.

[31]Y. Wang, D. Q. Yin, S. Yang, and G. Sun, "Global and local surrogate-assisted differential evolution for expensive constrained optimization problems with inequality constraints," IEEE Transactions on Cybernetics, vol. 49, no. 5, pp. 1642-1656, 2018.

[32]X. Lu, T. Sun, and K. Tang, "Evolutionary optimization with hierarchical surrogates," Swarm and Evolutionary Computation, vol. 47, pp. 21-32, 2019.

[33]T. Goel, R. T. Haftka, W. Shyy, and N. V. Queipo, "Ensemble of surrogates," Structural and Multidisciplinary Optimization, vol. 33, no. 3, pp. 199–216, 2006.

[34]M. T. M. Emmerich, K. C. Giannakoglou, and B. Naujoks, "Single-and multiobjective evolutionary optimization assisted by Gaussian random field metamodels," IEEE Transactions on Evolutionary Computation, vol. 10, no. 4, pp. 421–439, Aug. 2006.

[35]N. Namura, K. Shimoyama, and S. Obayashi, "Expected improvement of penalty-based boundary intersection for expensive multiobjective optimization," IEEE Transactions on Evolutionary Computation, vol. 21, no. 6, pp. 898-913, 2017.

[36]X. Wang, Y. Jin, S. Schmitt, and M. Olhofer, "An adaptive Bayesian approach to surrogate-assisted evolutionary multi-objective optimization," Information Sciences, vol. 519, pp. 317-331, 2020.

[37]H. Wang, Y. Jin, and J. O. Janson, "Data-driven surrogate-assisted multi-objective evolutionary optimization of a trauma system," IEEE Transactions on Evolutionary Computation, vol. 20, no. 6, pp. 939–952, Dec. 2016.

[38]C. Yang, J. Ding, Y. Jin, and T. Chai, "Offline Data-Driven Multiobjective Optimization: Knowledge Transfer Between Surrogates and Generation of Final Solutions," IEEE Transactions on Evolutionary Computation, vol. 24, no. 3, pp. 409-423, 2019.

[39]J. Stork, M. Friese, M. Zaefferer, T. Bartz-Beielstein, A. Fischbach, B. Breiderhoff, and T. Tušar, T, "Open issues in surrogate-assisted optimization," In High-Performance Simulation-Based Optimization, pp. 225-244, 2020.

[40]T. Chugh, K. Sindhya, J. Hakanen, and K. Miettinen, "A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms," Soft Computing, vol. 23, no. 9, pp. 3137-3166, 2019.

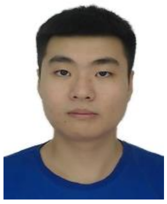
[41]J. Knowles, "ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems," IEEE Transactions on Evolutionary Computation, vol. 10, no. 1, pp. 50–66, Feb. 2006.

[42]Q. Zhang, W. Liu, E. Tsang, and B. Virginas, "Expensive multiobjective optimization by MOEA/D with Gaussian process model," IEEE Trans-

actions on Evolutionary Computation, vol. 14, no. 3, pp. 456–474, Jun. 2010.
[43]M. Tabatabaei, M. Hartikainen, K. Sindhya, J. Hakanen, and K. Miettinen, "An interactive surrogate-based method for computationally expensive multiobjective optimisation," Journal of the Operational Research Society, vol. 70 no. 6, pp. 898-914, 2019.
[44]I. Loshchilov, M. Schoenauer, and M. Sebag, "Dominance-based Pareto-surrogate for multi-objective optimization," in Asia-Pacific Conference on Simulated Evolution and Learning, Kanpur, India, pp. 230–239, 2010.
[45]A. A. M. Rahat, R. M. Everson, and J. E. Fieldsend, "Alternative infill strategies for expensive multi-objective optimisation," in Proceedings of the Genetic and Evolutionary Computation Conference, Berlin, Germany, pp. 873–880, 2017.
[46]T. Chugh, K. Sindhya, J. Hakanen, and K. Miettinen, "A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms," Soft Computing, vol. 23, no. 9, pp. 3137-3166, 2019.
[47]K. Deb, P. C. Roy, and R. Hussein, "Surrogate Modeling Approaches for Multiobjective Optimization: Methods, Taxonomy, and Results," Mathematical and Computational Applications, vol. 26, no. 1, pp. 5, 2021.
[48]R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff, "A reference vector guided evolutionary algorithm for many-objective optimization," IEEE Transactions on Evolutionary Computation, vol. 20, no. 5, pp. 773–791, Oct. 2016.
[49]S. Huband, P. Hingston, L. Barone, and L. While, "A review of multi-objective test problems and a scalable test problem toolkit," IEEE Transactions on Evolutionary Computation, vol. 10, no. 5, pp. 477–506, Oct. 2006.
[50]I. Das, and J. E. Dennis, "Normal Boundary Intersection: A New Method for Generating Pareto Optimal Points in Nonlinear Multicriteria Optimization Problems," SIAM Journal on Optimization, vol. 8, no. 3, pp. 631-657, 1998.
[51]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. 42, no. 1, pp. 55–61, 2000.
[52]K. Deb, and R. B. Agrawal, "Simulated binary crossover for continuous search space," Complex Systems, vol. 9, no. 2, pp. 115-148, 1995.
[53]R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, vol. 11, no. 4, pp. 341–359, 1997.
[54]K. Deb and M. Goyal, "A combined genetic adaptive search (geneas) for engineering design," Computer Science and Informatics, vol. 26, pp. 30–45, 1996.s
[55]K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable multiobjective optimization test problems," in IEEE Congress on Evolutionary Computation, vol. 1. Honolulu, HI, USA, pp. 825–830, 2002.
[56]Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, and S. Tiwari, "Multiobjective optimization test instances for the CEC 2009 special session and competition," 2008.
[57]R. Cheng, M. Li, Y. Tian, X. Zhang, S. Yang, Y. Jin, X. Yao, "A benchmark test suite for evolutionary many-objective optimization," Complex and Intelligent Systems, vol. 3, no. 1, pp. 67-81, 2017.
[58]C. A. C. Coello and M. R. Sierra, "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm," in Mexican International Conference on Artificial Intelligence. Springer, 2004, pp. 688–697.
[59]Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization," IEEE Computational Intelligence Magazine, vol. 12, no. 4, pp. 73–87, Nov. 2017.

Qiuzhen Lin (Member, IEEE) received the B.S. degree from Zhaoqing University, Zhaoqing, China, in 2007, the M.S. degree from Shenzhen University, Shenzhen, China, in 2010, and the Ph.D. degree from the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, in 2014.

He is currently an Associate Professor with the College of Computer Science and Software Engineering, Shenzhen University. He has published over 60 research papers since 2008. His current research interests include artificial immune system, multiobjective optimization, and dynamic system.



Xunfeng Wu received the B.S. degree from Shaoguan University, Shaoguan, China, in 2018 and the M.S. degree from Shenzhen University, Shenzhen, China, in 2021.

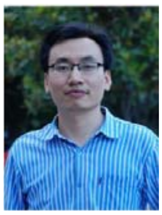
He is currently pursuing the Ph.D. degree in the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. His current research interests are in evolutionary computation, multi-objective optimization, and data-driven optimization.

Lijia Ma received the B.S. degree in communication engineering from Hunan Normal University, Changsha, China, and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2010 and 2015, respectively.

From 2015 to 2016, he was a Postdoctoral Fellow with Hong Kong Baptist University, Hong Kong, and with Nanyang Technological University, Singapore, from 2016 to 2017. He is an assistant professor at the College of Computer and Software Engineering of Shenzhen University. His research interests mainly include evolutionary computation, machine learning and complex networks.

Jianqiang Li (Member, IEEE) received the B.S. and Ph.D. degrees in automation from the South China University of Technology, Guangzhou, China, in 2003 and 2008, respectively.

He is a Professor with the College of Computer and Software Engineering, Shenzhen University, Shenzhen, China. He led a project of the National Natural Science Foundation, and a project of the Natural Science Foundation of Guangdong Province, China. His major research interests include embedded systems and Internet of Things.



Maoguo Gong (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively.

Since 2006, he has been a Teacher with Xidian University. In 2008 and 2010, he was promoted as an Associate Professor and a Full Professor, respectively, both with exceptive admission. His current research interests include computational intelligence with applications to optimization, learning, data mining, and image understanding.

Dr. Gong was a recipient of the Prestigious National Program for the support of Top-Notch Young Professionals from the Central Organization Department of China, the Excellent Young Scientist Foundation from the National Natural Science Foundation of China, and the New Century Excellent Talent in University from the Ministry of Education of China. He is the Vice Chair of the IEEE Computational Intelligence Society Task Force on Memetic Computing, an Executive Committee Member of the Chinese Association for Artificial Intelligence, and a Senior Member of the Chinese Computer Federation. He is also an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION.

Carlos A. Coello Coello (Fellow, IEEE) received the Ph.D. degree in computer science from Tulane University, New Orleans, LA, USA, in 1996.

He is a Professor (CINVESTAV-3F Researcher) with the Department of Computer Science of CINVESTAV-IPN, Mexico City, Mexico. He has authored and coauthored over 450 technical papers and book chapters. He has also coauthored the book *Evolutionary Algorithms for Solving Multi-Objective Problems* (Second Edition, Springer, 2007). His publications currently report over 57 600 citations in Google Scholar (his H-index is 95). His research interests include evolutionary multiobjective optimization and constraint-handling techniques for evolutionary algorithms.

Dr. Coello Coello was a recipient of the 2007 National Research Award from the Mexican Academy of Sciences in the area of *Exact Sciences*, the 2013 IEEE Kiyo Tomiyasu Award, and the 2012 National Medal of Science and Arts in the area of Physical, Mathematical and Natural Sciences. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION. He is a member of the Association for Computing Machinery and the Mexican Academy of Science.