Multi Co-objective Evolutionary Optimization: Cross Surrogate Augmentation for Computationally Expensive Problems

Minh Nghia Le and Yew Soon Ong
School of Computer Engineering
Nanyang Technological University, Singapore
Email: {mnle, asysong}@ntu.edu.sg

Stefan Menzel and Bernhard Sendhoff
Honda Research Institute Europe GmbH
Offenbach/Main, Germany
Email: {stefan.menzel, bernhard.sendhoff}@honda-ri.de

***Abstract*—In this paper, we present a novel cross-surrogate assisted memetic algorithm (CSAMA) as a manifestation of multi co-objective evolutionary computation to enhance the search on computationally expensive problems by means of transferring, sharing and reusing information across objectives. In particular, the construction of surrogate for one objective is augmented with information from other related objectives to improve the prediction quality. The process is termed as a cross-surrogate modelling methodology, which will be used in lieu with the original expensive functions during the evolutionary search. Analyses on the prediction quality of the cross-surrogate modelling and the search performance of the proposed algorithm are conducted on the benchmark problems with assessments made against several state-of-the-art multiobjective evolutionary algorithms. The results obtained highlight the the efficacy of the proposed CSAMA in attaining high quality Pareto optimal solutions with efficient performance under limited computational budget.**

***Keywords*-Multiobjective Evolutionary Algorithm, Memetic Computing, Co-objective, Computationally Expensive Problems, Meta-modelling, Surrogates**

I. Introduction

Engineering reliable and high quality products is now becoming an important practice of many industries to stay competitive in today's increasingly global economy, which is constantly exposed to high commercial pressures. The rapid advancements in computational power as well as simulation capabilities have helped fuel the evaluations and optimizations of a manifold of design proposals across various environmental scenarios. For instance, the complexity of engineering tasks in the automotive area requires increasingly advanced computational methods for the successful development of novel car designs involving multiple interrelating disciplines or scenarios that need to be considered, including aerodynamic optimization, structure optimization or noise, vibration and harshness (NVH) optimization. Even for a single scenario, usually several objectives have to be treated simultaneously, extending from drag reduction, design constraints to internal car air flow, etc., in the aerodynamic scenario. These realistic scenarios have since led to the formulations of multi-objective optimization (MO) problems which have attracted significant attention for the last few decades.

Very often, the tradeoffs among the objectives exist in MO problems in terms of Pareto optimality, thus any improvement for a Pareto-optimal solution in one objective must lead to deterioration in another. The collection of all the Pareto-optimal solutions forms the Pareto set and the set of all Pareto-optimal objective vectors is defined as the Pareto front (PF). The main task in solving multi-object optimization problems is thus to search for an appropriate number of Pareto-optimal solutions that are well distributed along the Pareto front (PF), providing a good representative of the entire PF [1]. Over the years, evolutionary algorithms (EAs) have been recognized and established as a core optimization technology for handling complex multiobjective optimization problems where significant success has been reported [2], [3], [4], [5], [6], [7]. Among the many evolutionary algorithms that have been designed to solve multi-objective problems to date, Non-dominated Sorting Genetic Algorithm (NSGA-II) [8], Pareto archived evolutionary strategy (PAES) [9] and Strength Pareto Evolutionary Algorithm (SPEA2) [10] are among the most prominently investigated. Typically, when high-fidelity analysis codes are used, it is not uncommon for even single simulation process to take minutes, hours to days of supercomputer time to compute. We label such category of problems as expensive optimization problems in the present study. Since the design cycle time of a product is directly proportional to the number of calls made to the costly analysis solvers, researchers are seeking for novel multi-objective optimization frameworks that can handle these forms of problems elegantly. Besides parallelism, which is an obvious choice to achieving near linear order improvement in evolutionary search, researchers are gearing towards surrogate-assisted or meta-model assisted evolutionary frameworks when handling optimization problems imbued with costly non-linear objective and constraint functions [11], [12], [13], [14], [15], [16].

In the context of multiobjective evolutionary optimization, despite the significant progress made in the last decades, there has been an apparent lack of attempts that exchange and reuse information among the objectives in either a competitive and/or collaborative manner, with the common goal of finding the optimal Pareto solutions efficiently. This paper thus presents a study to showcase such an attempt. To proceed

we first define the notion of *multi co-objective* evolutionary optimization in this paper as a *search methodology where information are transferred, exchanged, shared, or reused across objectives, for the purpose of more effective and efficient multi-objective optimization*. In this scenario, it is assumed that the information or knowledge gained in the process of solving one or more objectives can positively influence the optimization of another objective. It is worth noting that in practice, real-world complex problems often exhibit some degrees of similarity or correlations between the multiple objectives. In engineering, for example, the multiobjective optimization of drag and lift in aerodynamic design could share common computational fluid dynamics simulations [17]. The information that is transferred between objectives could come in the form of fitness sharing, solutions (genotype/phenotype) sharing or in the form of complex structures derived from the objective search space.

In this paper, our interest is to enhance the extremely time consuming evolutionary design optimization process involving multiple objectives with computationally costly simulations. Particularly, in the spirit of the multi co-objective evolutionary search paradigm, here we explore the idea of augmenting information from one objective into the process of building the surrogates for other objectives. This has the effect of alleviating the impact of the curse of dimensionality due to the lack of sufficient data samples for the modeling of accurate approximation models, thus bringing about improvements in the prediction quality of the surrogate models built. We term this process to improve the prediction quality as cross-surrogate modelling, which will be used in lieu with the original expensive objective functions.

The paper is outlined as follows. Section II formulates the optimization problem of interest and provides a formal introduction to multi co-objective evolutionary optimization. Subsequently, Section II-A presents the proposed cross-surrogate learning methodology for the multi co-objective evolutionary optimization of computationally expensive problems. Section II-B proceeds to introduce the cross-surrogate assisted memetic algorithm (CSAMA) for multi co-objective optimization used in the study. Section III then presents an empirical study on the prediction quality and search performance of the proposed cross-surrogate algorithm on the set of multi-objective benchmark problems, with assessment made against several state-of-the-art multiobjective evolutionary algorithms (MOEAs). Last but not least, Section IV concludes the present study with a brief discussion of future work.

II. Multi Co-objective Evolutionary Computation

Without loss of generality, we consider here the general multiobjective evolutionary optimization of the following form:

$$\begin{aligned} \text{Minimize}: \quad & \mathbf{F}(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \\ \text{Subject to}: \quad & l_i \leq x_i \leq u_i \end{aligned}$$

where $i = 1, 2, \ldots, n$, n is the dimensionality of the search problem. l_i, u_i are the lower and upper bounds of the i-th dimension, respectively. $\mathbf{x} \in \mathbb{R}^n$ denotes the input vector of scalars $\{x_1, x_2, \ldots, x_n\}$ or design parameters and k is the number of objective functions.

A solution $\mathbf{x}$ is said to be dominated by another solution $\mathbf{y}$ if and only if

$$\begin{aligned} \forall i, f_i(\mathbf{y}) \leq f_i(\mathbf{x}) \\ \exists j, f_j(\mathbf{y}) < f_j(\mathbf{x}) \end{aligned}$$

The set of all non-dominated or Pareto-optimal solutions, denoted as *PS*, is called the Pareto set. The set of the Pareto objective vectors, $PF = \{F(\mathbf{x}) | \mathbf{x} \in PS\}$, is known as the Pareto front. The task of a multiobjective optimization algorithm is to find a *representative* Pareto-optimal solutions that well cover the Pareto front.

Conventional MOEA relies on the populations of solutions $\mathbf{D} = \{(\mathbf{x} | f_1(\mathbf{x}), \ldots, f_k(\mathbf{x}))\}$ to infer on the regions where non-dominated solutions lie. Multi co-objective evolutionary optimization takes a further step in facilitating the search using solution datasets $\{\mathbf{D}_1, \ldots, \mathbf{D}_k\}$ that can come in different distributions, where $\mathbf{D}_j = \{(\mathbf{x} | f_j(\mathbf{x}))\}$ for the objective f_j, or alternatively in the form of complex structures on the objective functions $\{I(f_i), \ldots, I(f_i, f_j, \ldots, f_k)\}$. For objective f_i, $I(f_i)$ can manifest in the form of generalized fitness model, which is the core interest of our study later in Section II-A. For a set of objectives $\{f_i, f_j, \ldots, f_k\}$, the structure $I(f_i, f_j, \ldots, f_k)$ can be revealed as the correlations between functions or complex models that describe how one objective responses to the increment of other objective functions. The learning of $I(f_j)$ relies not only on the solutions in single objective $\mathbf{D}_j$, but also solutions from other objectives $\mathbf{D}_i$ (via solutions sharing), and the complex structures $I(f_i)$ and/or $I(f_j, f_i, \ldots, f_k)$. In this manner, multi co-objective evolutionary optimization thus promotes the transfer, exchange or sharing of information among objectives for the purpose of more effective and efficient search.

A. Cross-Surrogate Modeling

Data-centric surrogate models or metamodels are statistical models built to approximate the computationally expensive simulation codes. They are approximation models that are many orders of magnitude cheaper to run and can be used to replace or in lieu of the original high-fidelity analysis codes during evolutionary search to reduce the overall optimization search time. Here, our interest is on cases where the evaluation of $\mathbf{F}(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\}$ are computationally expensive, and it is desirable to obtain a near optimal Pareto solutions under a limited computational budget. In the spirit of the multi co-objective evolutionary search paradigm, we propose the cross-surrogate modelling methodology that augments information from one objective in the process of building approximation models or surrogates for other objectives. Without loss of generality, in what follows we illustrate the multi co-objective evolutionary optimization process involving two computationally expensive functions f_1 and f_2.

Let $\mathbf{D}_1 = \{(\mathbf{x}_i^{(1)} | f_1(\mathbf{x}_i^{(1)}))\}_{i=1}^{m_1}$ and $\mathbf{D}_2 = \{(\mathbf{x}_i^{(2)} | f_2(\mathbf{x}_i^{(2)}))\}_{i=1}^{m_2}$ denote the training datasets for objective

functions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$, respectively. With the sampling training data points $\mathbf{D}_1$ of the expensive simulation $f_1(\mathbf{x})$, an approximation model $\widehat{f}_1(\mathbf{x})$, or the base surrogate, is first constructed using common approximation methodology M such as polynomial regression (PR) [18], radial basis function (RBF) [19] or Gaussian process or Kriging [20]. Subsequently, the construction of the surrogate model for $f_2(\mathbf{x})$, denoted as the cross surrogate, thus leverages the correlation between objective functions f_1 and f_2 to reinforce the prediction quality. For the sake of illustration and simplicity, the correlation relationship is expressed here using a quadratic model in Equation 1.

$$\begin{aligned} f_2(\mathbf{x}) &= Q(\mathbf{x}, \widehat{f}_1(\mathbf{x})) \\ &= \gamma + \beta_0 \times \widehat{f}_1(\mathbf{x}) + \sum_{i=1}^{n} \beta_i \times x_i \\ &+ \sum_{i=1}^{n} \alpha_{0,i} \times \widehat{f}_1(\mathbf{x}) \times x_i + \sum_{i,j=1}^{n} \alpha_{i,j} \times x_i \times x_j \\ &= (\beta_0 + \sum_{i=1}^{n} \alpha_{0,i} \times x_i) \times \widehat{f}_1(\mathbf{x}) \\ &+ (\gamma + \sum_{i=1}^{n} \beta_i \times x_i + \sum_{i,j=1}^{n} \alpha_{i,j} \times x_i \times x_j) \end{aligned} \tag{1}$$

Let

$$\alpha(\mathbf{x}) = \beta_0 + \sum_{i=1}^{n} \alpha_{0,i} \times x_i \tag{2}$$

$$P(\mathbf{x}) = \gamma + \sum_{i=1}^{n} \beta_i \times x_i + \sum_{i,j=1}^{n} \alpha_{i,j} \times x_i \times x_j \tag{3}$$

then Equation 1 can be written as

$$f_2(\mathbf{x}) = \alpha(\mathbf{x}) \times \widehat{f}_1(\mathbf{x}) + P(\mathbf{x}) \tag{4}$$

where $P(\mathbf{x})$ denotes a quadratic model of the design (input) variables $\mathbf{x}$. To illustrate the correlation implied in Equation 1, note that f_1 and f_2 exhibit a positive or inverted correlation when $P(\mathbf{x})$ is a constant and $\alpha(\mathbf{x})$ is a positive or negative constant, respectively. On the other hand, if the two functions possess zero correlation, i.e., $\alpha(\mathbf{x}) \approx 0$, $\widehat{f}_2$ is thus a quadratic approximation of the design variables.

From Equation 1, the cross surrogate $\widehat{f}_2(\mathbf{x})$ is constructed as a quadratic model $Q(.)$ of $\{(\mathbf{x}, \widehat{f}_1(\mathbf{x}) | f_2(\mathbf{x}))\}$. The training data for cross surrogate construction of f_2 is denoted as $\mathbf{D}_2^{aug} = \{(\mathbf{x}_i^{(2)}, \widehat{f}_1(\mathbf{x}_i^{(2)}) | f_2(\mathbf{x}_i^{(2)}))\}_{i=1}^{m_2}$, where $\mathbf{D}_2$ is augmented with model $\widehat{f}_1$ (which serves as the surrogate for objective f_1). The proposed cross-surrogate modelling methodology is summarized in Algorithm 1. Subsequently, for a given new input vector $\mathbf{x}^{new}$, the predicted fitness value of f_2 is inferred using both the input vector and the predicted fitness from base surrogate $\widehat{f}_1(\mathbf{x}^{new})$ as $Q(\widehat{f}_1(\mathbf{x}^{new}), \mathbf{x}^{new})$. In comparison to the high cost of computation in original functions f_1 and f_2, the building and evaluation cost of surrogate $\widehat{f}_1(\mathbf{x}^{new})$ are typically considered as negligible.

It is worth noting that the proposed algorithm presents a fast learning technique (i.e., quadratic regression) so that it can be used as local surrogate in the evolutionary optimization framework, where the correlation in the local regions is more likely to happen. Besides, the construction of the cross surrogate $\widehat{f}_2(\mathbf{x})$ only requires a small number of data points for proper estimation of the unknown coefficients, i.e., $m_2 = (n+2) \times (n+3)/2$. It is also worth highlighting that the proposed methodology allows the training datasets $\mathbf{D}_1$ to be sampled independent of $\mathbf{D}_2$ and vice versa. This is an important feature, since it is often the case that the computational costs of the objectives are heterogeneous in many real world complex problems. For instance, the assessment of a potential design solution involving simulations may takes few minutes on one objective function but tens or hundreds of minutes to compute on other objectives.

Algorithm 1 Cross-Surrogate Modeling Methodology

Query training datasets $\mathbf{D}_k = \{(\mathbf{x}_i^{(k)} | f_k(\mathbf{x}_i^{(k)}))\}_{i=1}^{m_k}$, $k = 1, 2$ for objective function f_1 and f_2
Construct base surrogate model $\widehat{f}_1(\mathbf{x})$ for f_1 based on training dataset $\mathbf{D}_1$ using approximation methodology M
Augment training dataset $\mathbf{D}_2$ with $\widehat{f}_1(\mathbf{x})$ to arrive at $\mathbf{D}_2^{aug} = \{(\mathbf{x}_i^{(2)}, \widehat{f}_1(\mathbf{x}_i^{(2)}) | f_2(\mathbf{x}_i^{(2)}))\}_{i=1}^{m_2}$
Construct cross surrogate model $\widehat{f}_2(\mathbf{x})$ for f_2 based on training dataset $\mathbf{D}_2^{aug}$

B. Cross-surrogate Assisted Memetic Algorithm

Next, the proposed cross-surrogate assisted memetic algorithm (CSAMA) for multi co-objective evolutionary optimization using the novel cross-surrogate modelling methodology in Section II-A is described. Here we consider computationally expensive minimization problems under limited computational budget, in which one evaluation of the design solution may take minutes to hours of computer simulation. Using a multi-objective memetic framework [5], [6], [21], [22], the proposed CSAMA algorithm is outlined in Algorithm 2. Here we adopt a strategy of building separate approximation models for each computationally expensive objective function. This allows an independent construction of surrogates which is useful in dealing with the possible heterogeneous nature of the computational cost incurred by the objective functions.

In the first step, a population of N individuals is initialized either randomly or using design of experiment techniques such as Latin hypercube sampling. The fitness values of each individuals in the population are then determined based on the original expensive objective functions $\{f_i(\mathbf{x})\}$. The evaluated population then undergoes natural selection, namely binary tournament based on Pareto ranking and crowding distance to control the diversity of the population [8]. Each individual $\mathbf{x}$ in the reproduction pool is evolved to arrive at the offspring population using multiobjective stochastic variation operators including crossover and mutation. Subsequently, with ample

design points in the databases $\{\mathbf{D}_i\}$ or after some predefined database building phase of exact evaluations G_{db} (line 5), the trust-region enabled local search on the aggregated function $f_{aggr}(\mathbf{x})$ of cross-surrogate kicks in for each non-duplicated design point or individuals in the population. Here the aggregated objective function $g(\mathbf{x})$ is attained based on the Tchebycheff approach [1]. More specifically, for each individual in the population, the local search method L-BFGS-B proceeds on the aggregated surrogate objective function $f_{aggr}(\mathbf{x})$ with a sub-problem of the form:

$$\begin{aligned} \text{Minimize}: \quad & f_{aggr}(\mathbf{x}) = \max_{1\leq i\leq k} \{\lambda_i \times |\widehat{f}(\mathbf{x}) - z_i^*|\} \\ \text{Subject to}: \quad & ||\mathbf{x}|| \leq \Omega \end{aligned}$$

where Ω represents the local search region of interest based on the offspring $\mathbf{x}$, $\{\lambda_i\}$ are random weights such that $\sum_{i=1}^{k} \lambda_i = 1$ and $\mathbf{z}^* = (z_1^*, z_2^*, \ldots, z_k^*)$ denotes the reference point in the objective space. Here $z_i^* = \min f_i(\mathbf{x})$, for each $i = 1, \ldots, k$. For an unknown optimization problem with no prior knowledge on the exact reference point $\mathbf{z}^*$, the value of z_i^* is then updated by applying local search on surrogate $\widehat{f}_i$ and evaluate the output of the local search as the current minimum of f_i at each generation (line 8 in Algorithm 2).

For a given individual solution $\mathbf{x}$ at generation t, the local surrogates $\{\widehat{f}_i(\mathbf{x})\}$ is constructed using the cross-surrogate modelling methodology in Section II-A where f_1 is used as the base surrogate (line 10). Rather than constructing a global surrogate based on the entire archived samples, nearest sampled points to the offspring $\mathbf{x}$ in the database $\mathbf{D}_i$ are selected as training dataset $T = \{(\mathbf{x}_j, f_i(\mathbf{x}_j))\}_{j=1}^{m}$ for building local model $\widehat{f}_i(\mathbf{x})$ which serves as the surrogate for the computationally expensive objective function f_i. Note that the use of local surrogate results in a gain of approximation accuracy in the local region of interest where the correlation relationship between objectives is more likely to exist.

The improved solution found using the aggregated function of surrogates is subsequently evaluated using the original computationally expensive fitness functions f_i and replaces the parent individual in the population, in the spirit of Lamarckian learning. The refined offspring and parent populations are then merged into the next population. Exact evaluations of all newly found individuals $\{(\mathbf{x}, f_i(\mathbf{x})\}$ are then archived into the database $\mathbf{D}_i$ for subsequent use in the search. The entire process repeats until the specified stopping criteria are satisfied. Note that an external population (EP) is used to store all non-dominated solutions found during the search as the output of the algorithm.

Algorithm 2 Cross-surrogate assisted Memetic Algorithm

1: *Initialization*: Generate N initial solutions randomly (by experimental design method) and evaluate $\{f_i(\mathbf{x})\}$ objective functions at N solutions
2: **while** NOT stopping condition **do**
3: *Select* reproduction pool based using binary tournament selection
4: *Generate* offspring population P_o using MO evolutionary operators (crossover, mutation)
5: **if** evaluation count $<$ database building phase G_{db} **then**
6: *Evaluate* offspring population by exact fitness functions
7: **else**
8: Update the reference point $\mathbf{z}^* = (z_1^*, z_2^*, \ldots, z_k^*)$
9: **for** each individual $\mathbf{x}$ in the offspring population **do**
10: *Build local surrogates* $\widehat{f}_i(\mathbf{x})$ using ***Cross-Surrogate Modelling*** (Algorithm 1)
11: *Apply local search* on aggregation of surrogates $f_{aggr}(\mathbf{x})$ to arrive at a refined offspring
12: **end for**
13: *Evaluate* refined offspring population P_o by exact fitness functions
14: **end if**
15: *Merge* to form the next population, $P^{(t+1)} = P^{(t)} \bigcup P_o$, based on Pareto ranking and crowding distance
16: *Archive* evaluated solutions to $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_k$
17: **end while**

III. Empirical Study

This section presents the numerical study on several aspects of the proposed work in multi co-objective evolutionary optimization. In particular, the predictive efficacy of the cross-surrogate modelling methodology is first analyzed in Section III-A on the synthetic objective functions to illustrate the two scenarios of positive correlation and inverted correlation between the objectives. Subsequently, analyses on the search performance of the proposed CSAMA on the benchmark problems with assessment made against state-of-the-art multiobjective evolutionary algorithms are presented in Section III-B.

A. Prediction Quality Analysis

The prediction quality of the proposed cross-surrogate modelling is pitted against the conventional approximation approach on two objective functions f_1 and f_2. Note that in the present study, standard quadratic regression or polynomial regression is also employed for the construction of the base surrogate. With the goal of building approximation models for two objective functions f_1 and f_2, the cross-surrogate modelling is denoted as $\{\widehat{f}_1(\mathbf{x}) \leftarrow Q_1(\mathbf{x}), \widehat{f}_2(\mathbf{x}) \leftarrow Q_2(\mathbf{x}, \widehat{f}_1(\mathbf{x}))\}$ while the conventional approach is denoted as $\{\widehat{f}_1(\mathbf{x}) \leftarrow Q_1(\mathbf{x}), \widehat{f}_2(\mathbf{x}) \leftarrow Q_2(\mathbf{x})\}$.

Consider three 2-dimensional synthetic objective functions, i.e., $n = 2$:

$$F_{Sphere} = \sum_{i=1}^{n} x_i^2$$

$$F_{Rastrigin} = \sum_{i=1}^{n} x_i^2 - 10 \times \sum_{i=1}^{n} \cos(2\pi x_i) + 10n$$

$$F_{Griewank} = \sum_{i=1}^{n} x_i^2/4000 - \prod_{i=1}^{n} \cos(x_i/\sqrt{i}) + 1$$

$$-10 \leq x_i \leq 10, i = 1 \ldots n$$

1) Positive correlation: When $f_1(\mathbf{x}) = F_{Sphere}$ and $f_2(\mathbf{x}) = F_{Rastrigin}$, it is worth noting that Equation 4 becomes

$$f_2(\mathbf{x}) = f_1(\mathbf{x}) - 10 \times \sum_{i=1}^{n} \cos(2\pi x_i) + 10n \quad (5)$$

If $f_1(\mathbf{x}) = F_{Sphere}$ and $f_2(\mathbf{x}) = F_{Griewank}$, we then have

$$f_2(\mathbf{x}) = 1/4000 \times f_1(\mathbf{x}) - \prod_{i=1}^{n} \cos(x_i/\sqrt{i}) + 1 \quad (6)$$

The positive coefficients of f_1 in both Equations 5 and 6 thus indicate positive correlations between f_1 and f_2 in the range of $[-10, 10]^2$.

2) Negative correlation: If $f_1(\mathbf{x}) = -F_{Sphere}$ and $f_2(\mathbf{x}) = F_{Rastrigin}$, it is worth noting that Equation 4 becomes

$$f_2(\mathbf{x}) = -f_1(\mathbf{x}) - 10 \times \sum_{i=1}^{n} \cos(2\pi x_i) + 10n \quad (7)$$

Similarly, if $f_1(\mathbf{x}) = -F_{Sphere}$ and $f_2(\mathbf{x}) = F_{Griewank}$, we then have

$$f_2(\mathbf{x}) = -1/4000 \times f_1(\mathbf{x}) - \prod_{i=1}^{n} \cos(x_i/\sqrt{i}) + 1 \quad (8)$$

The negative coefficients of f_1 in both Equations 7 and 8 instead indicate inverted correlation between f_1 and f_2 in the range of $[-10, 10]^2$.

For both cases, the training dataset which consists of 8000 points is used for f_1 and f_2, while the testing dataset consists of 2000 points randomly sampled from $[-10, 10]^2$. For each point in the testing dataset, the prediction errors on f_2 using cross-surrogate modelling and conventional surrogate modelling based on quadratic regression are evaluated and normalized. The normalized prediction errors in the cases of positive and inverted correlation are then plotted in Figures 1 and 2, respectively. It can be observed from the figures that the cross surrogate $\hat{f}_2(\mathbf{x})$ with the augmented information from $f_1(\mathbf{x})$ exhibits lower prediction error than the conventional approach in both cases, thus indicating the improvements in the prediction quality of cross-surrogate modelling when correlations among objectives exist.

B. Performance Results and Analysis

Next, the proposed CSAMA for multi co-objective evolutionary optimization is assessed by comparing its search performance to other state-of-the-art multiobjective evolutionary algorithms, including NSGA-II [8], SPEA2 [10] and MOEA/D [23], on the representative *ZDT* benchmark multiobjective problem set [24]. As discussed in [24], this benchmark problem set contains problems with features of convexity/nonconvexity, discreteness and nonuniformity, thus posing difficulties to many evolutionary algorithms in converging to the true Pareto-optimal front. The algorithmic parameters of CSAMA are summarized in Table I while the details of the benchmark problems used in the present study are summarized in Table II. To facilitate a fair comparison study, the parametric configurations of NSGA-II, SPEA2 and MOEA/D are set to the default values as provided and implemented in *jMetal* [25] with the same population size of 100. It is worth highlighting that here we consider a limited computational budget of only 8000 evaluations for all the algorithms to discover a good representative of the true Pareto front.

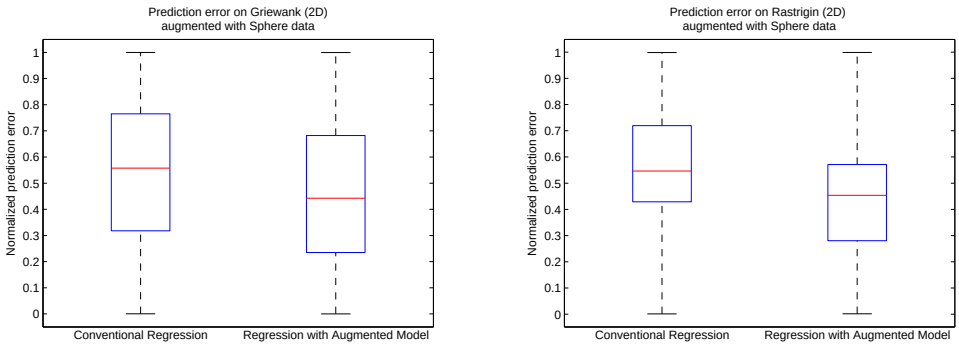
A number of performance metrics for multiobjective optimization (MO) have been proposed and discussed in the literature, which aim to evaluate the closeness to the Pareto-optimal front and the diversity of the obtained solution set, or both criterion [8], [27], [28], [29]. Most of the existing metrics require the obtained set to be compared against a specified set of Pareto-optimal reference solutions. In this study, the inverted generational distance (IGD) is used as the performance metric since it has been shown to reflect both the diversity and convergence of the obtained non-dominated solutions [29].

Let P^* denote a set of uniformly distributed solutions along the Pareto front in the objective space. Let A be an approximate set to the PF found by the algorithm, the distance metric from P^* to A is defined as:

$$IGD(A, P^*) = \frac{\sum_{v \in P^*} d(v, A)}{|P^*|}$$

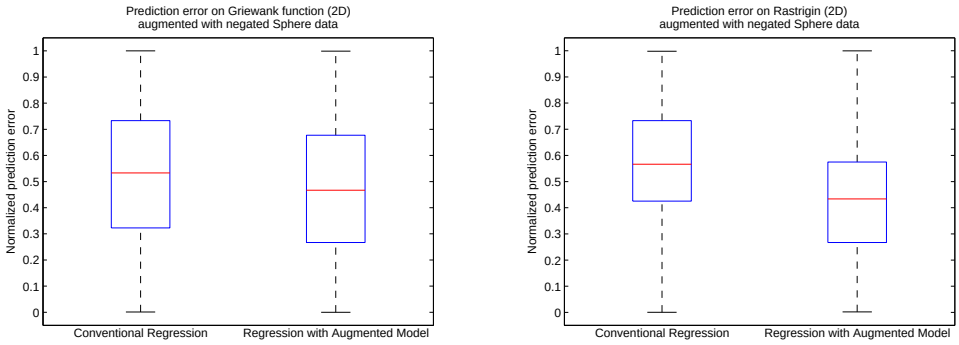
where $d(v, A)$ is the minimum Euclidean distance between v and the points in A, i.e., $d(v, A) = \min_{u \in A} d(v, u)$. Note that the distance is calculated in the objective space, i.e., $\{f_i(\mathbf{x})\}$. As P^* represents the Pareto front, the metric $IGD(A, P^*)$ thus measures both the diversity and convergence of A. A low value of $IGD(A, P^*)$ implies that the set A must be very close to as well as cover well the entire Pareto front.

The performance metric IGD of the non-dominated solutions set obtained from each of 30 independent runs by CSAMA and the other multiobjective evolutionary algorithms (MOEA) considered on each benchmark problems are depicted as box-plots in Figure 3. As seen from the figures, CSAMA exhibits significantly lower IGD by an order of magnitude than others in all of the problems, thus highlighting the efficacy of the proposed algorithm in finding a representative set of the Pareto front under the limited computational budget. Further, the plots of the true Pareto front of each benchmark problem, which is represented by the red line, against the representative sets found by CSAMA and other MOEAs are also presented in Figure 4. The figures clearly illustrate better search performance in both convergence and distribution of the non-dominated solutions found by the proposed algorithm on the true Pareto front of each benchmark problems. These results in a way strengthen our aspiration towards multi co-objective evolutionary optimization and hence the significance of CSAMA in facilitating the knowledge transfer or cross-surrogate modelling for more effective and efficient multiobjective search.



(a) **Sphere - Griewank** (b) **Sphere - Rastrigin**

Fig. 1. Comparison of Prediction Quality for Positive Correlation



(a) **Negated Sphere - Griewank** (b) **Negated Sphere - Rastrigin**

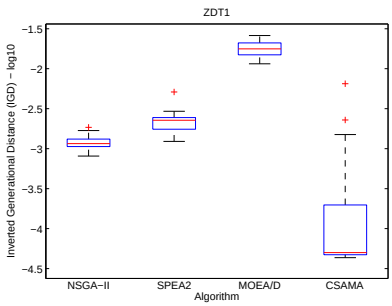
Fig. 2. Comparison of Prediction Quality for Inverted Correlation

TABLE I
CSAMA ALGORITHMIC CONFIGURATION

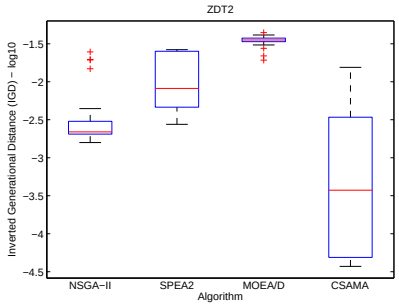
Parameters	
Population size	100
Selection scheme	Binary tournament
Stopping criteria	8000 evaluations
Local search method	L-BFGS-B
Crossover probability (p_{cross})	0.9
Crossover operator	Simulated Binary Crossover (SBX) [26]
Mutation probability (p_{mut})	$1/n$
Mutation operator	Polynomial Mutation
Database building phase (G_{db})	1600 evaluations

IV. CONCLUSIONS

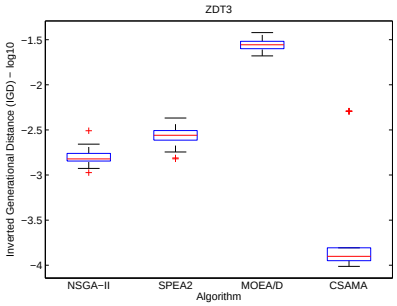
In this paper, a cross-surrogate assisted memetic algorithm (CSAMA) is proposed as an instantiation of the multi co-objective evolutionary computation to enhance the search on computationally expensive problems by means of transferring, sharing and reusing information across objectives. In particular, information from one objective is augmented into the process of constructing the surrogates for other related objectives to bring about improvement in prediction quality of the cross surrogates, as demonstrated on the synthetic objective functions. Subsequently, numerical study of CSAMA with assessment made against other state-of-the-art multiobjective evolutionary algorithms on the set of benchmark problems demonstrated the efficacy of the proposed algorithm and confirmed our motivation for multi co-objective search. CSAMA thus serves as a novel effort towards the design of multi co-objective evolutionary algorithms that exchange and reuse information across objectives in a competition and cooperation manner. Future research will continue to focus on the challenges and open issues in multi co-objective search in real-world scenarios, especially on problems with heterogenous computational costs.



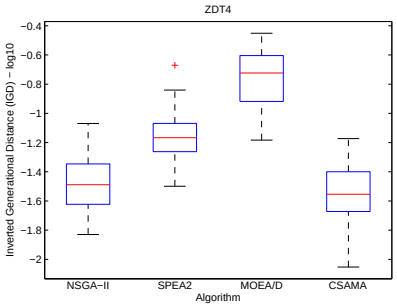
(a) **ZDT1**



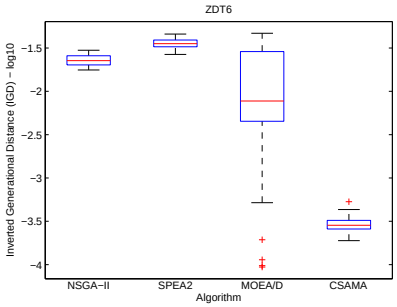
(b) **ZDT2**



(c) **ZDT3**

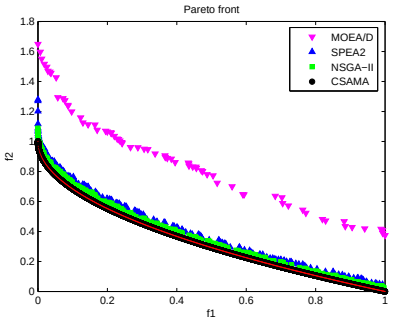


(d) **ZDT4**

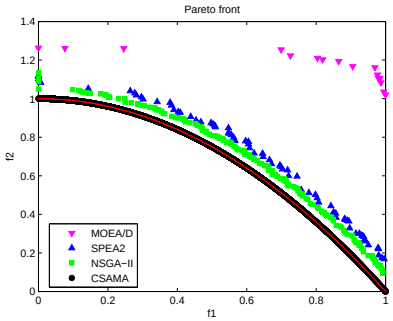


(e) **ZDT6**

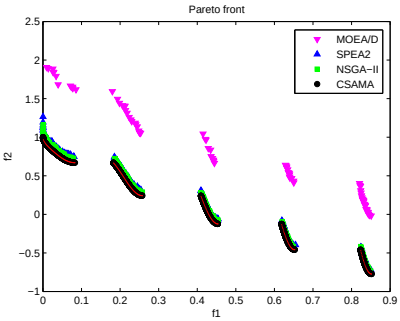
Fig. 3. Inverted Generational Distance (IGD) performance metrics for benchmark problems



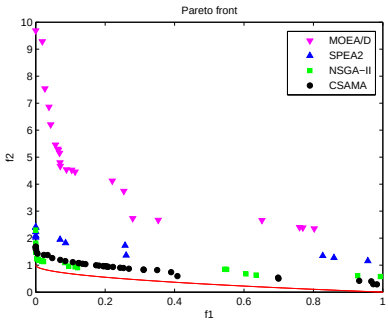
(a) **ZDT1**



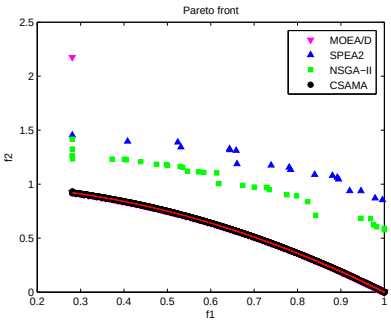
(b) **ZDT2**



(c) **ZDT3**



(d) **ZDT4**



(e) **ZDT6**

Fig. 4. Pareto front of benchmark problems found by NSGA-II, SPEA2, MOEA/D and CSAMA

TABLE II
BENCHMARK FUNCTIONS

Benchmark Function	Formulation	Pareto front
ZDT1 ($n = 30$, $x_i \in [0, 1]$)	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times [1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}]$ $g(\mathbf{x}) = 1 + 9 \times \sum_{i=2}^{n} x_i/(n-1)$	2-objective, high dimensionality, convex Pareto front
ZDT2 ($n = 30$, $x_i \in [0, 1]$)	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times [1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2]$ $g(\mathbf{x}) = 1 + 9 \times \sum_{i=2}^{n} x_i/(n-1)$	2-objective, high dimensionality, concave Pareto front
ZDT3 ($n = 30$, $x_i \in [0, 1]$)	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times [1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})} - (f_1(\mathbf{x})/g(\mathbf{x})) \sin(10\pi f_i(\mathbf{x}))]$ $g(\mathbf{x}) = 1 + 9 \times \sum_{i=2}^{n} x_i/(n-1)$	2-objective, high dimensionality, Pareto front of disconnected convex parts
ZDT4 ($n = 10$)	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times [1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}]$ $g(\mathbf{x}) = 1 + 10(n-1) + \sum_{i=2}^{n} (x_i^2 - 10\cos(4\pi x_i))$ $x_1 \in [0, 1]$ and $x_2, \ldots, x_n \in [-5, 5]$	2-objective, multimodality with 21^9 local Pareto fronts (not all distinguishable in the objective space)
ZDT6 ($n = 10$, $x_i \in [0, 1]$)	$f_1(\mathbf{x}) = 1 - \exp(-4x_i)\sin^6(6\pi x_1)$ $f_2(\mathbf{x}) = g(\mathbf{x}) \times [1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2]$ $g(\mathbf{x}) = 1 + 9 \times ((\sum_{i=2}^{n} x_i)/(n-1))^{0.25}$	2-objective, non-uniform distribution of the Pareto-optimal solutions, concave Pareto-front

ACKNOWLEDGMENT

M.N. Le is grateful to the financial support of Honda Research Institute Europe.

REFERENCES

[1] K. Miettinen, *Nonlinear Multiobjective Optimization*. Springer, 1999, vol. 12.

[2] H. Ishibuchi and T. Murata, "A multi-objective genetic local search algorithm and its application to flowshop scheduling," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 28, no. 3, pp. 392–403, 1998.

[3] K. Deb, *Multi-objective Optimization using Evolutionary Algorithms*. Wiley, 2001, vol. 16.

[4] C. A. C. Coello, G. B. Lamont, and D. A. V. Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed. Springer, September 2007, vol. 5.

[5] C. K. Goh, Y. S. Ong, and K. C. Tan, *Multi-Objective Memetic Algorithms*. Springer Verlag, 2009, vol. 171.

[6] A. Jaszkiewicz, H. Ishibuchi, and Q. Zhang, "Multiobjective Memetic Algorithms," *Handbook of Memetic Algorithms*, pp. 201–217, 2011.

[7] 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, 2009.

[8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[9] J. Knowles and D. Corne, "The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Multiobjective Optimization," in *Proceedings of the 1999 Congress on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 1999, pp. 9–105.

[10] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm," Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, Tech. Rep. 103, 2001.

[11] Y. Jin, M. Olhofer, and B. Sendhoff, "A Framework for Evolutionary Optimization with Approximate Fitness Functions," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 5, pp. 481–494, Oct 2002.

[12] Y. S. Ong, P. B. Nair, and A. J. Keane, "Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modelling," *American Institute of Aeronautics and Astronautics Journal*, vol. 41, no. 4, pp. 687–696, 2003.

[13] Y. Jin, "A Comprehensive Survey of Fitness Approximation in Evolutionary Computation," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 9, no. 1, pp. 3–12, 2005.

[14] D. Lim, Y. Jin, Y. S. Ong, and B. Sendhoff, "Generalizing Surrogate-assisted Evolutionary Computation," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 3, pp. 329–355, June 2010.

[15] L. Shi and K. Rasheed, "A Survey of Fitness Approximation Methods Applied in Evolutionary Algorithms," *Computational Intelligence in Expensive Optimization Problems*, pp. 3–28, 2010.

[16] I. Voutchkov and A. Keane, "Multi-objective optimization using surrogates," *Computational Intelligence in Optimization*, pp. 155–175, 2010.

[17] Y. S. Ong, P. B. Nair, and A. J. Keane, "Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modelling," *American Institute of Aeronautics and Astronautics Journal*, vol. 41, no. 4, pp. 687–696, 2003.

[18] F. H. Lesh, "Multi-dimensional least-squares polynomial curve fitting," *Communications of ACM*, vol. 2, no. 9, pp. 29–30, 1959.

[19] M. J. D. Powell, "Radial basis functions for multivariable interpolation: a review," pp. 143–167, 1987.

[20] D. J. Mackay, "Introduction to Gaussian processes," *Neural Networks and Machine Learning*, vol. 168, pp. 133–165, 1998.

[21] J. Knowles and D. Corne, "Memetic Algorithms for Multiobjective Optimization: Issues, Methods and Prospects," *Recent advances in memetic algorithms*, pp. 313–352, 2005.

[22] K. Tang, K. Tan, and H. Ishibuchi, "Guest editorial: Memetic algorithms for evolutionary multi-objective optimization," *Memetic Computing*, vol. 2, no. 1, pp. 1–1, 2010.

[23] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.

[24] E. Zitzler, K. Deb, and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results," *Evolutionary Computation*, vol. 8, pp. 173–195, June 2000.

[25] J. J. Durillo and A. J. Nebro, "jMetal: A Java framework for multi-objective optimization," *Advances in Engineering Software*, vol. 42, pp. 760–771, 2011.

[26] K. Deb and R. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 2, pp. 115–148, 1995.

[27] D. Van Veldhuizen and G. Lamont, "Multiobjective Evolutionary Algorithm Research: A History and Analysis," Air Force Institute of Technology, Tech. Rep., 1998.

[28] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.

[29] E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. da Fonseca, "Performance Assessment of Multiobjective Optimizers: An Analysis and Review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, 2003.