

Le, Minh Nghia, Ong, Yew Soon, Menzel, Stefan, Jin, Yaochu, Sendhoff, Bernhard. Evolution by adapting surrogates. *Evolutionary Computation*, uncorrected proof.

---

## Evolution by Adapting Surrogates

**Minh Nghia Le**

mnle@ntu.edu.sg

School of Computer Engineering, Nanyang Technological University,  
639798, Singapore

**Yew Soon Ong**

asysong@ntu.edu.sg

School of Computer Engineering, Nanyang Technological University,  
639798, Singapore

**Stefan Menzel**

stefan.menzel@honda-ri.de

Honda Research Institute Europe GmbH, 63073 Offenbach, Main, Germany

**Yaochu Jin**

yaochu.jin@surrey.ac.uk

Department of Computing, University of Surrey Guildford, Surrey,  
GU27XH United Kingdom

**Bernhard Sendhoff**

bernhard.sendhoff@honda-ri.de

Honda Research Institute Europe GmbH, 63073 Offenbach/Main, Germany

---

### Abstract

To deal with complex optimization problems plagued with computationally expensive fitness functions, the use of surrogates to replace the original functions within the evolutionary framework is becoming a common practice. However, the appropriate datacentric approximation methodology to use for the construction of surrogate model would depend largely on the nature of the problem of interest, which varies from fitness landscape and state of the evolutionary search, to the characteristics of search algorithm used. This has given rise to the plethora of surrogate-assisted evolutionary frameworks proposed in the literature with ad hoc approximation/surrogate modeling methodologies considered. Since prior knowledge on the suitability of the data centric approximation methodology to use in surrogate-assisted evolutionary optimization is typically unavailable beforehand, this paper presents a novel evolutionary framework with the evolvability learning of surrogates (EvoLS) operating on multiple diverse approximation methodologies in the search. Further, in contrast to the common use of fitness prediction error as a criterion for the selection of surrogates, the concept of evolvability to indicate the productivity or suitability of an approximation methodology that brings about fitness improvement in the evolutionary search is introduced as the basis for adaptation. The backbone of the proposed EvoLS is a statistical learning scheme to determine the evolvability of each approximation methodology while the search progresses online. For each individual solution, the most productive approximation methodology is inferred, that is, the method with highest evolvability measure. Fitness improving surrogates are subsequently constructed for use within a trust-region enabled local search strategy, leading to the self-configuration of a surrogate-assisted memetic algorithm for solving computationally expensive problems. A numerical study of EvoLS on commonly used benchmark problems and a real-world computationally expensive aerodynamic car rear design problem highlights the efficacy of the proposed EvoLS in attaining reliable, high quality, and efficient performance under a limited computational budget.

### Keywords

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

## 1 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. Strong engineering design know-how results in lower time to market and better quality at lower cost. In recent years, advancements in science, engineering, and the availability of massive computational power have led to the increasing high-fidelity approaches introduced for precise studies of complex systems in silico. Modern computational structural mechanics, computational electromagnetics, computational fluid dynamics (CFD), and quantum mechanical calculations represent some of the approaches that have been shown to be highly accurate (Jin, 2002; Zienkiewicz and Taylor, 2006; Hirsch, 2007). These techniques play a central role in the modelling, simulation, and design process, since they serve as efficient and convenient alternatives for conducting trials on the original real-world complex systems that are otherwise deemed to be too costly or hazardous to construct.

Typically, when high-fidelity analysis codes are used, it is not uncommon for the single simulation process to take minutes, hours, or days of supercomputer time to compute. A motivating example at Honda Research Institute is aerodynamic car rear design, where one function evaluation involving a CFD simulation to calculate the fitness performance of a potential design can take many hours of wall clock time. Since the design cycle time of a product is directly proportional to the number of calls made to the costly analysis solvers, researchers are now seeking novel stochastic optimization approaches, including evolutionary frameworks, that 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 toward surrogate-assisted or meta-model assisted evolutionary frameworks when handling optimization problems imbued with costly nonlinear objective and constraint functions (Liang et al., 1999; Jin et al., 2002; Song, 2002; Ong et al., 2003; Jin, 2005; Lim et al., 2007, 2010; Tenne, 2009; Shi and Rasheed, 2010; Voutchkov and Keane, 2010; Santana-Quintero et al., 2010).

The general consensus on surrogate-assisted evolutionary frameworks is that the efficiency of the search process can be improved by replacing, as often as possible, calls to the costly high-fidelity analysis solvers with surrogates that are deemed to be less costly to build and compute. In this manner, the overall computational burden of the evolutionary search can be greatly reduced, since the effort required to build the surrogates and to use them is much lower than those in the traditional approach that directly couples the evolutionary algorithm (EA) with the costly solvers (Serafini, 1998; Booker et al., 1999; Jin et al., 2000; Kim and Cho, 2002; Emmerich et al., 2002; Jin and Sendhoff, 2004; Ulmer et al., 2004; Branke and Schmidt, 2005; Zhou et al., 2006). Among many datacentric approximation methodologies used to construct surrogates to date, polynomial regression (PR) or response surface methodology (Lesh, 1959), support vector machines (Cortes and Vapnik, 1995; Vapnik, 1998), artificial neural networks (Zurada, 1992), radial basis functions (Powell, 1987), the Gaussian process (GP) referred to as Kriging, or design and analysis of computer experiment models (Mackay, 1998; Buche et al., 2005) and ensembles of surrogates (Zerpa et al., 2005; Goel et al., 2007; Sanchez et al., 2008; Acar and Rais-Rohani, 2009) are among the most prominently investigated (Jin, 2005; Zhou et al., 2006; Shi and Rasheed, 2010). Early proposed approaches have considered using surrogates that target modeling the entire solution space or fitness landscape of the costly exact objective or fitness function. However, due to the

sparseness of data points collected along the evolutionary search, the construction of accurate global surrogates (Ulmer et al., 2004; Buche et al., 2005) that mimics the entire problem landscape is impeded by the effect of the curse of dimensionality (Donoho, 2000). To enhance the accuracies of the surrogates used, researchers have turned to localized models (Giannakoglou, 2002; Emmerich et al., 2002; Ong et al., 2003; Regis and Shoemaker, 2004) as opposed to globalized models, or their synergies (Zhou et al., 2006; Lim et al., 2010). Others have also considered the use of gradient information (Ong et al., 2004) to enhance the prediction accuracy of the constructed surrogate models or physics-based models that are deemed to be more trustworthy than pure datacentric counterparts (Keane and Petruzzelli, 2000; Lim et al., 2008).

In the context of surrogate-assisted optimization (Jin et al., 2001; Shi and Rasheed, 2010), present performance or assessment metrics and schemes used for surrogate model selection and validation involve many prominent approaches that have taken root in the field of statistical and machine learning (Fielding and Bell, 1997; Shi and Rasheed, 2010). Particularly, the focus have been placed on attaining a surrogate model that has minimal apparent error or training error on some optimization data collected during the evolutionary search, as an estimation of the true error when used to replace the original costly high-fidelity analysis solver. maximum/mean absolute error, root mean square error (RMSE), and correlation measure denote some of the performance metrics that are commonly used (Jin et al., 2001). Typical model selection schemes that stem from the field of statistical and machine learning, including the split sample (holdout) approach, cross-validation, and bootstrapping, are subsequently used to choose surrogate models that have a low estimation of apparent and true errors (Queipo et al., 2005; Tenne and Armfield, 2008b). Tenne and Armfield (2008a) used multiple cross-validation schemes for the selection of low-error surrogates that replace the original costly high-fidelity analysis solvers to avoid convergence at false optima of poor accuracy models.

In spite of the significant research effort spent on optimizing computationally expensive problems more efficiently, existing surrogate-assisted evolutionary frameworks may be fundamentally bounded by the use of apparent or training error as the performance metric used to assess and select surrogates. In contrast, it would be more worthwhile to select surrogates that enhance search improvement in the context of optimization, as opposed to the usual practice of choosing a surrogate model with minimal estimated true error. Further, the influence of the datacentric approximation methodology employed is deemed to have a major impact on surrogate-assisted evolutionary search performance. The varied suitability of approximation methodology to different fitness landscapes, state of the search, and characteristics of the search algorithm suggests for the varieties of surrogate-assisted evolutionary frameworks in the literature that have emerged with ad hoc approximation methodology selection. To the best of our knowledge, little work has been done to mitigate this issue, since only limited knowledge of the black-box optimization problem is available before getting started.

Falling back on the basics of Darwin's grand idea of natural selection as the criterion for the choice of surrogates that brings about fitness improvement to the search, this paper describes a novel evolutionary search process that evolves along with fitness improving surrogates. Here, we focus our study on the evolvability learning of surrogates (EvoLS), particularly the adaptive choice of datacentric approximation methodologies that build fitness improving surrogates in place of the original computationally expensive black-box problem, during the evolutionary search. In the spirit of reward-based or improvement-based adaptation for the choice of variation operators (Goldberg, 1990; Thierens, 2005; DaCosta et al., 2008), EvoLS infers the fitness improvement contribution

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

of each approximation methodology toward the search, which is here referred to as the evolvability measure. Hence, for each individual or design solution in the evolutionary search, the evolvability of each approximation methodology is determined statistically, according to the current state of the search, properties of the search operators, and characteristics of the fitness landscape, while the search progresses online. Using the evolvability measures derived, the search adapts by using the most productive approximation methodology inferred for the construction of surrogates embedded within a trust-region enabled local search strategy, leading to the self-configuration of surrogate-assisted memetic algorithm that deals with complex optimization of computationally expensive problems more effectively. It is worth noting that to the best of our knowledge, our current effort has not been previously studied in the literature of surrogate assisted EA.

The paper is outlined as follows. Section 2 introduces the notion of evolvability as a performance or assessment measure that expresses the suitability of an approximation methodology in guiding toward improved evolutionary search and subsequently the essential ingredients of our proposed EvoLS. Section 3 presents a numerical study of EvoLS on commonly used benchmark problems. Detailed analyses on the suitability and cooperation of surrogates in search, as well as the correlation between the estimated fitness prediction error and evolvability in EvoLS of the surrogate models, are also presented in Section 3. Section 4 then proceeds to showcase the real world application of EvoLS on an aerodynamic car rear design that involves highly computationally expensive CFD simulations. Finally, Section 5 summarizes the present study with a brief discussion and concludes the paper.

## 2 Proposed Framework: Evolvability Learning of Surrogates

In this section, we present the essential ingredients of the proposed EvoLS for handling computationally expensive optimization problems. In particular, we concentrate on the general nonlinear programming problem of the following form:

Minimize:  $f(\mathbf{x})$

Subject to:  $\mathbf{x}_l \leq \mathbf{x} \leq \mathbf{x}_u$

where  $\mathbf{x} \in \mathbb{R}^n$  is the vector of design variables, and  $\mathbf{x}_l, \mathbf{x}_u$  are vectors of lower and upper bounds, respectively. In this paper, we are interested in the cases where the evaluation of  $f(\mathbf{x})$  is computationally expensive, and it is desired to obtain a near-optimal solution on a limited computational budget, using a novel evolutionary process that adapts fitness improving surrogates, that is,  $\hat{f}_M(\mathbf{x})$  as a replacement of  $f(\mathbf{x})$ , in the search.

Datacentric surrogates are (statistical) models that are built to approximate computationally expensive simulation codes or the exact fitness evaluations. They are orders of magnitude cheaper to run and can be used in lieu of the exact analysis during evolutionary search. Let  $\{(\mathbf{x}_i, t_i)\}_{i=1}^m$  where  $t_i = f(\mathbf{x}_i)$  denote the training dataset, where  $\mathbf{x} \in \mathbb{R}^n$  is the input vector of scalars or design parameters, and  $f(\mathbf{x}) \in \mathbb{R}$  is the output or exact fitness value. Based on the approximation methodology  $M$ , the surrogate  $\hat{f}_M(\mathbf{x})$  is constructed as an approximation model of the function  $f(\mathbf{x})$ . Further, the surrogate model can also yield insights into the functional relationship between the input  $\mathbf{x}$  and the objective function  $f(\mathbf{x})$ .

In what follows, we begin with a formal introduction on the notion of evolvability as a performance or assessment measure to indicate the productivity and suitability of an approximation methodology for constructing a surrogate that brings about fitness improvement to the evolutionary search (see Section 2.1). The essential backbone of our

proposed EvoLS framework is an EA coupled with a trust-region enabled local search strategy with adaptive surrogates, in the spirit of Lamarckian learning. In contrast to existing works, we adapt the choice of approximation methodology for the construction of fitness improving datacentric surrogates in place of the original computationally expensive black-box problem when conducting the computationally intensive local search in the context of memetic optimization (Hart et al., 2004; Krasnogor and Smith, 2005; Ong et al., 2010; see Section 2.2).

### 2.1 Evolvability of Surrogates

Conventionally, surrogate models are assessed and chosen according to their estimated true error,  $|f(\mathbf{x}) - \hat{f}_M(\mathbf{x})|$ , where  $\hat{f}_M(\mathbf{x})$  denotes the predicted fitness value of input vector  $\mathbf{x}$  by a surrogate constructed using approximation method  $M$ . In contrast to existing surrogate-assisted evolutionary search, the surrogate model employed for each individual design solution in the present study favors fitness improvement as the choice of merit to assess the usefulness of surrogates in enhancing search improvement, as opposed to the estimated true error.

In this subsection, we introduce the concept of evolvability of an approximation methodology as the basis for adaptation. Since the term evolvability has been used in different contexts,<sup>1</sup> it is worth highlighting that our concept of evolvability generalizes from that of learnability in machine learning (Valiant, 2009). Here the evolutionary process is regarded as evolvable on a given optimization problem if the progress in search performance is observed for some moderate number of generations. Hence, evolvability of an approximation methodology here refers to the propensity of the method in constructing a surrogate model that guides toward viable, or potentially favorable, individuals with improved fitness quality.

In particular, the evolvability measure of an approximation methodology  $M$  for the construction of a fitness improving datacentric surrogate on individual solution  $\mathbf{x}$  at generation  $t$ , assuming a minimization problem, is denoted here as  $Ev_M(\mathbf{x})$  and derived in the form of

$$\begin{aligned}
 Ev_M(\mathbf{x}) &= \text{Exp}[f(\mathbf{x}) - f(\mathbf{y}^{\text{opt}})|\mathbf{P}^t, \mathbf{x}] \\
 &= f(\mathbf{x}) - \int_{\mathbf{y}} f(\varphi_M(\mathbf{y})) \times P(\mathbf{y}|\mathbf{P}^t, \mathbf{x}) d\mathbf{y}
 \end{aligned} \tag{1}$$

Here  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$  denotes the conditional density function of the stochastic variation operators applied on parent  $\mathbf{x}$  to arrive at solution  $\mathbf{y}$  at generation  $t$ , that is,  $\mathbf{y} \sim P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$ , where  $\mathbf{P}^t$  is the current reproduction pool consisting of individual solutions after undergoing natural selection.  $\varphi_M(\mathbf{y})$  denotes the resultant solution derived by the local search strategy that operates on the surrogate constructed based on approximation method  $M$ . The evolvability measure of an approximation methodology indicates the expectation of fitness improvement which the refined offspring, denoted here as  $\mathbf{y}^{\text{opt}} = \varphi_M(\mathbf{y})$ , has gained over its parent, upon undergoing local search improvement on the respective constructed surrogate.<sup>2</sup> A high evolvability measure encapsulates two core essences of

<sup>1</sup>In Wagner and Altenberg (1996), evolvability is defined as the genome’s ability to produce adaptive variants when acted upon by the genetic system. Others have generally referred the term to the ability of stochastic or random variations to produce improvement for adaptation to happen (Ong, Lim et al. 2006).

<sup>2</sup>The local search strategy serves as a function that takes the starting solution  $\mathbf{x}$  as input and returns  $\mathbf{y}^{\text{opt}}$  as the refined solution. As such, the function is represented as  $\mathbf{y}^{\text{opt}} = \varphi_M(\mathbf{y})$ .

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

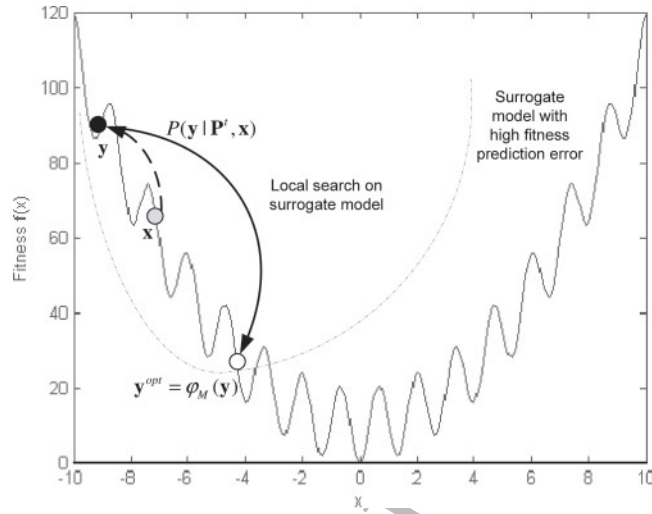


Figure 1: Illustration of evolvability under the effect of the blessing of uncertainty.

the surrogate-assisted evolutionary search: (1) When a surrogate exhibits low true error estimates, fitness improvement on the refined offspring  $\mathbf{y}^{\text{opt}}$  over initial parent  $\mathbf{x}$  can be expected; and (2) When a surrogate exhibits high true error estimates, the discovery of offspring solutions with improved fitness  $\mathbf{y}^{\text{opt}}$  of  $\mathbf{x}$  can be still attained due to the blessing of uncertainty phenomenon (Ong, Zhou, et al., 2006; see Figure 1 for an example illustration).

Taking into account the current state of the evolutionary search, properties of the search operators, and characteristics of the fitness landscape, a statistical learning approach to estimate the evolvability measure  $Ev_M(\mathbf{x})$  of each approximation methodology, as defined in Equation (1), for use on a given individual solution  $\mathbf{x}$  at generation  $t$ , is proposed. Let  $\Phi_M = \{(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))\}_{i=1}^K$  denote the database of distinct samples archived along the search, which represents the historical contribution of the approximation methodology on the problem considered. Through a weighted sampling approach, the weight  $w_i(\mathbf{x})$  that defines the probability of choosing a sample  $(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))$  for the estimation of expected improvement  $Ev_M(\mathbf{x})$ , or the integral  $\int_{\mathcal{Y}} f(\varphi_M(\mathbf{y})) \times P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})d\mathbf{y}$  in particular, is first derived and efficiently estimated here as follows:

$$w_i(\mathbf{x}) = \frac{P(\mathbf{y}_i|\mathbf{P}^t, \mathbf{x})}{\sum_{j=1}^K P(\mathbf{y}_j|\mathbf{P}^t, \mathbf{x})} \quad (2)$$

The weight  $w_i(\mathbf{x})$  essentially reflects the current probability of jumping from solution  $\mathbf{x}$ , via stochastic variation  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$ , to offspring  $\mathbf{y}_i$ , which was archived in the database. In other words, the weight  $w_i(\mathbf{x})$  measures the relevancy of the samples  $(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))$  for the evolvability learning process on solution  $\mathbf{x}$ . Considering  $\{(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))\}_{i=1}^K$  as distinct samples from current distribution  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$ , the weights  $w_i(\mathbf{x})$  associated with samples  $(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))$  satisfy the equations:  $\sum_{i=1}^K w_i(\mathbf{x}) = 1$  and  $w_i(\mathbf{x})$  is proportional to  $P(\mathbf{y}_i|\mathbf{P}^t, \mathbf{x})$ . Note that the conditional density function  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$  is modeled probabilistically based on the properties of the evolutionary variation operators used to reflect the current state of the search. Subsequently, using the archived samples in  $\Phi_M = \{(\mathbf{y}_i, \varphi_M(\mathbf{y}_i))\}_{i=1}^K$  and

**Algorithm 1** EvoLS framework

---

```

1: Generate and evaluate an initial population
2: while computational budget is not exhausted do
3:   Select individuals for the reproduction pool  $\mathbf{P}^t$ 
4:   if evaluation count < database building phase ( $G_{db}$ ) then
5:     Evolve population by evolutionary operators (crossover, mutation)
6:   else
7:     Evolve population by evolutionary operators (crossover, mutation)
8:     for each individual  $\mathbf{x}$  in the population do
9:       Perform approximation methodology selection on  $\mathbf{x}$  to arrive at  $M_{id}$ 
10:      /** Local Search Phase on Surrogate Model***/
11:      Find  $m$  nearest points to  $\mathbf{y}$  in the database  $\Psi$  as training points for surrogate model
12:      Build surrogate model  $\hat{f}_{M_{id}}(\mathbf{x})$  based on training points
13:      Apply local search strategy  $\varphi_{M_{id}}(\mathbf{y})$  to arrive at  $\mathbf{y}^{\text{opt}}$ 
14:      Replace  $\mathbf{y}$  with  $\mathbf{y}^{\text{opt}}$  (Lamarckian learning)
15:      Archive sample  $(\mathbf{y}, \varphi_{M_{id}}(\mathbf{y}))$  into the database  $\Phi_{M_{id}}$ 
16:     end for
17:   end if
18:   Evaluate new population using exact fitness function  $f(\mathbf{x})$ 
19:   Archive all exact evaluations  $(\mathbf{x}, f(\mathbf{x}))$  into the database  $\Psi$ 
20: end while

```

---

weights  $w_i$  obtained using Equation (2),  $Ev_M(\mathbf{x})$  is estimated as follows:

$$Ev_M(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^K f(\varphi_M(\mathbf{y}_i)) \times w_i(\mathbf{x}) \quad (3)$$

## 2.2 Evolution with Adapting Surrogates

The proposed EvoLS for solving computationally expensive optimization problems is presented and outlined in Algorithm 1. The essential ingredients of our proposed EvoLS framework are composed of multiple datacentric approximation methodologies having diverse characteristics,<sup>3</sup> denoted here as  $\{M_{id}\}_{id=1}^D$ . 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 cost or fitness value of each individuals in the population is then determined using  $f(\mathbf{x})$ . The evaluated population then undergoes natural selection, for instance, via fitness-proportional or tournament selection. Each individual  $\mathbf{x}$  in the reproduction pool  $\mathbf{P}^t$  is evolved to arrive at the offspring  $\mathbf{y}$  using stochastic variation operators including crossover and mutation. Subsequently, with ample design points in the database  $\Psi$  or after some predefined database building phase of exact evaluations  $G_{db}$  (line 4), the trust-region enabled local search with adaptive surrogate kicks in for each nonduplicated design point or individuals in the population. For a given individual solution  $\mathbf{x}$  at generation  $t$ , the evolvability  $Ev_{M_{id}}(\mathbf{x})$  of each datacentric approximation methodology is estimated statistically by taking into account the current state of the search, properties of search operators and characteristics of the fitness landscape via the historical contribution by the respective constructed surrogates, while

<sup>3</sup>Typical approximation techniques are radial basic function (RBF), Kriging or GP, and PR.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

---

**Algorithm 2** Approximation methodology selection process

---

- 1: Construct density distribution  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$  of variation operators
  - 2: **for** each approximation methodology  $M_{id}$  **do**
  - 3:   Query archived data  $\Phi_{M_{id}} = \{(\mathbf{y}_j, \varphi_{M_{id}}(\mathbf{y}_j))\}$  for  $M_{id}$
  - 4:   Calculate weight  $w_i(\mathbf{x}) = P(\mathbf{y}_i|\mathbf{P}^t, \mathbf{x})$  for each sample  $\mathbf{y}_i$
  - 5:   **if**  $\sum_{i=1}^K w_i(\mathbf{x}) < \epsilon$  **then**
  - 6:      $w_i(\mathbf{x}) = 0$  {No relevant data available}
  - 7:      $Ev_{M_{id}}(\mathbf{x}) = -\infty$
  - 8:   **else**
  - 9:      $w_i(\mathbf{x}) = w_i(\mathbf{x}) / \sum_{j=1}^K w_j(\mathbf{x})$  {Normalize  $w_i$ }
  - 10:     $Ev_{M_{id}}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^K f(\varphi_{M_{id}}(\mathbf{y}_i)) \times w_i(\mathbf{x})$  {Equation (3)}
  - 11:   **end if**
  - 12: **end for**
  - 13: **if**  $Ev_{M_{id}}(\mathbf{x}) < 0 \forall M_{id}$  **then**
  - 14:   Select approximation methodology randomly
  - 15: **else**
  - 16:   Select approximation methodology with highest  $Ev_{M_{id}}(\mathbf{x})$  for  $\mathbf{x}$
  - 17: **end if**
- 

the search progresses online.<sup>4</sup> Without loss of generality, in the event of a minimization problem, the most productive datacentric approximation methodology, which is deemed as one that has the highest estimated evolvability measure  $\arg \max Ev_{M_{id}}(\mathbf{x})$ , is then chosen to construct a surrogate that will be used by the local search to bring about fitness improvement on an individual  $\mathbf{x}$ .

The outline of the approximation methodology selection process is detailed in Algorithm 2. It is worth noting on the generality of the proposed framework in the use of density function  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$  to represent and reflect the unique characteristics of the stochastic search operators (line 1), thus not restricting to any specific type of operator. By formulating the search operator with a density function, the framework allows the incorporation of different stochastic variation operators, which depends on suitability to the given problem of interest. It is worth highlighting that the condition  $\sum_{i=1}^K w_i(\mathbf{x}) < \epsilon$  (line 5) caters to the scenario when all samples  $(\mathbf{y}_i, \varphi_{M_{id}}(\mathbf{y}_i))$  are irrelevant for evolvability estimation on solution  $\mathbf{x}$ , that is,  $P(\mathbf{y}_i|\mathbf{P}^t, \mathbf{x})$  is too small. The role of  $\epsilon$  thus specifies the threshold level of irrelevance for archived samples in the evolvability learning. In particular,  $\epsilon$  is configured to a precision of  $1E - 9$  in our present study.

Rather than constructing a global surrogate based on the entire archived samples, the nearest sampled points to  $\mathbf{y}$  in the database  $\Psi$  are selected as the training dataset  $T = \{(\mathbf{x}_i, f(\mathbf{x}_i))\}_{i=1}^m$  for building local fitness improving surrogate  $\hat{f}_{M_{id}}$ . The improved solution found using the respective constructed surrogate, denoted here as  $\mathbf{y}^{\text{opt}} = \varphi_{M_{id}}(\mathbf{y})$ , is subsequently evaluated using the original computationally expensive fitness function  $f(\mathbf{x})$  and replaces the parent individual in the population, in the spirit of Lamarckian learning. Exact evaluations of all newly found individuals  $\{(\mathbf{y}^{\text{opt}}, f(\mathbf{y}^{\text{opt}}))\}$ , together with  $\{(\mathbf{y}, \varphi_{M_{id}}(\mathbf{y}))\}$ , are then archived into the database  $\Psi$  and  $\Phi_{M_{id}}$ , respectively, for

---

<sup>4</sup>Note that a uniform selection of approximation methodologies is employed in EvoLS instead of Algorithm 2 in the first generation right after the database building phase. In this way, all approximation methodologies are ensured to be given equal opportunities to work on each individual in the population. Thus, sufficient data points are expected to be made available for the evolvability learning of each approximation methodology in the subsequent generations.



subsequent use in the search. The entire process repeats until the specified stopping criteria are satisfied.

### 3 Empirical Study

In this section, we present the numerical results obtained by the proposed EvoLS using three commonly used approximation methodologies, namely: (1) interpolating linear spline RBF, (2) second order PR, and (3) interpolating Kriging/GP. For the details on GP, PR, and RBF, the reader is referred to Section 5. A representative group of 30-dimensional benchmark problems considered in the present study are summarized in Table 1 while the algorithmic parameters of EvoLS are summarized in Table 2.

The stochastic variation operators considered in the present study are uniform crossover and mutation, which have been widely used in real-coded genetic evolution. For the sake of brevity, we consider the use of simple variation operators in our illustrating example to showcase the mechanism and generality of the proposed EvoLS. It is worth noting that other advanced real-parameter search operators, such as that discussed in (Herrera et al., 1998), can also be considered in the framework through density distribution  $P(\mathbf{y}|\mathbf{P}', \mathbf{x})$ .

In particular, the crossover procedure considered here is outlined as follows: (1) randomly select two solutions,  $\mathbf{x}$  and  $\mathbf{x}'$ , from the population; (2) perform uniform crossover on  $\mathbf{x}$  and  $\mathbf{x}'$  to create two offspring  $\mathbf{y}_1$  and  $\mathbf{y}_2$  where the locus  $i$  of offspring  $\mathbf{y}_1^{(i)}/\mathbf{y}_2^{(i)}$  has a value of  $\mathbf{x}_1^{(i)}/\mathbf{x}_2^{(i)}$  with a crossover probability of  $p_{\text{cross}}$ , and (3) select  $\mathbf{y} = \mathbf{y}_1$  or  $\mathbf{y}_2$  as the offspring of  $\mathbf{x}$ . Uniform mutation is conducted on  $\mathbf{y}$  such that each locus  $\mathbf{y}^{(i)}$  is assigned to a random value bounded by  $[\min_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}, \max_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}]$  with a mutation probability of  $p_{\text{mut}}$ . Note that  $N$  denotes the population size.

The task of deriving exact density distribution function is nontrivial due to the complexity of the search dynamics. Thus, a simplified assumption of uniformity in the offspring distribution is considered for practical purposes in our empirical study. The density distribution  $P(\mathbf{y}|\mathbf{P}', \mathbf{x})$  is then modeled as a uniform distribution with

$$P(\mathbf{y}|\mathbf{P}', \mathbf{x}) = \text{UniformDist}(\mathbf{R}) = \begin{cases} \frac{1}{\text{Vol}(\mathbf{R})} & \text{if } \mathbf{y} \in \mathbf{R} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where  $\text{Vol}(\mathbf{R})$  denotes the hypervolume of a hyper-rectangle with bounds  $\mathbf{R}$  defined as

$$\mathbf{R} = [\min_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}, \max_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}]_{i=1\dots n}$$

Note that the stochastic variations considered impose the resultant offspring  $\mathbf{y}$  to be bounded by  $\min_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}$  and  $\max_{j=1\dots N} \{\mathbf{x}_j^{(i)}\}$  for each dimension,<sup>5</sup> that is,  $\forall i = 1 \dots n$ . Since the hyper-rectangle  $\mathbf{R}$  reduces as the search progresses, the probabilistic model of the variation operators reflects well on the refinement of the search space throughout the evolution.

On the other hand, the local search strategy of EvoLS involves a trust-region framework (Ong, Zhou, et al., 2006) with the Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B)

<sup>5</sup>If  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  and  $\mathbf{y}$  denote the parents and offspring, then each locus of the offspring  $\mathbf{y}$  satisfies the inequality

$$\min \{\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i)}\} \leq \mathbf{y}^{(i)} \leq \max \{\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i)}\}, \quad \forall i = 1 \dots n \quad (5)$$

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

Table 1: Benchmark problems considered in the empirical study. On shifted rotated problems, note that  $\mathbf{z} = \mathbf{M} \times (\mathbf{x} - \mathbf{o})$  where  $\mathbf{M}$  is the rotation matrix and  $\mathbf{o}$  is the shifted global optimum. Otherwise,  $\mathbf{z} = \mathbf{x}$ .

Function	Benchmark test functions	Range of $\mathbf{x}$	Multi*	Non-sep*
Ackley (F1)	$F_{\text{Ackley}} = 20 + e - 20e^{-0.2\sqrt{\frac{1}{n}\sum_{i=1}^n z_i^2} - e^{\frac{1}{n}\sum_{i=1}^n \cos(2\pi z_i)}}$	$[-32, 32]^n$	Yes	Yes
Griewank (F2)	$F_{\text{Griewank}} = 1 + \sum_{i=1}^n \frac{z_i^2}{4000} - \prod_{i=1}^n \cos(z_i/\sqrt{i})$	$[-600, 600]^n$	Yes	Yes
Rosenbrock (F3)	$F_{\text{Rosenbrock}} = \sum_{i=1}^{n-1} (100 \times (z_{i+1} - z_i^2)^2 + (1 - z_i)^2)$	$[-2.048, 2.048]^n$	Yes	Yes
Shifted rotated Rastrigin (F4)	$F_{\text{Rastrigin-SR}} = 10n + \sum_{i=1}^n (z_i^2 - 10 \cos(2\pi z_i))$	$[-5, 5]^n$	Yes	Yes
Shifted rotated Weierstrass (F5)	$F_{\text{Weierstrass-SR}} = \sum_{i=1}^n (\sum_{k=0}^{k_{\text{max}}} (a^k \cos(2\pi b^k (z_i + 0.5)))) - n \sum_{k=0}^{k_{\text{max}}} (a^k \cos(\pi b^k))$	$[-0.5, 0.5]^n$	Yes	Yes
Expanded Griewank plus Rosenbrock (F6)	$F_{\text{Grie+Rosen}} = \sum_{i=1}^D F_{\text{Griewank}}(F_{\text{Rosenbrock}}(z_i, z_{i+1}))$ $z_{D+1} = z_1$	$[-3, 1]^n$	Yes	Yes

Table 2: Algorithm configuration.

Parameters	
Population size	100
Selection scheme	Roulette wheel
Stopping criteria	8,000 evaluations
Local search method	L-BFGS-B
Number of trust region iteration	3
Crossover probability ( $p_{\text{cross}}$ )	1
Mutation probability ( $p_{\text{mut}}$ )	0.01
Variation operator	Uniform crossover and mutation
Database building phase ( $G_{db}$ )	2,000 evaluations
Precision accuracy	1E-8

method (Zhu et al., 1997). Note that under mild assumptions, the trust-region framework ensures theoretical convergence to the local optimum or stationary point of the exact objective function, despite the use of surrogate models (Alexandrov et al., 1998; Ong et al., 2003). For each individual  $\mathbf{y}_i$  in the population, the local search method L-BFGS-B proceeds on the inferred most productive surrogate model  $\hat{f}_M(\mathbf{x})$  to perform a sequence of trust-region subproblems of the form:

$$\begin{aligned} &\text{Minimize: } \hat{f}_M(\mathbf{y} + \mathbf{y}_i^k) \\ &\text{Subject to: } \|\mathbf{y}\| \leq \Omega^k \end{aligned}$$

where  $k = 0, 1, 2, \dots, k_{\max}$ ,  $\hat{f}_M(\mathbf{y})$  denotes the approximate function corresponding to the original fitness function  $f(\mathbf{y})$ , and  $\mathbf{y}_i^k$  and  $\Omega^k$  denote the starting point and trust-region radius at iteration  $k$ , respectively. For each individual  $\mathbf{y}_i$ , the surrogate model  $\hat{f}_M(\mathbf{y})$  of the original fitness function is created dynamically using training data from the archived database  $\Psi = \{(\mathbf{x}_i, f(\mathbf{x}_i))\}_{i=1}^Q$  to estimate the fitness during local search. Note that  $\mathbf{y}_i^{k+1} = \arg \min \hat{f}_M(\mathbf{y} + \mathbf{y}_i^k)$  denotes the local optimum of the trust-region subproblem at iteration  $k$ . At each  $k$ th iteration,  $\mathbf{y}_i^{k+1}$  and the trust-region radius  $\Omega^k$  are updated accordingly. In the present study, the resultant individual denoted here as  $\mathbf{y}_i^{k_{\max}} = \varphi_M(\mathbf{y}_i)$ , is the improved solution attained by L-BFGS-B over surrogate model  $\hat{f}_M(\mathbf{x})$ .

### 3.1 Search Quality and Efficiency of EvoLS

To see how adapting the choice of most productive approximation methodology affects the performance and efficiency of the search as compared to the use of the single approximation method, in this subsection, the performance of symbiotic evolution is compared with the canonical surrogate-assisted evolutionary algorithms (SAEAs) with the single approximation methodology, that is, EA-RBF, EA-PR, and EA-GP. Note that the procedure of canonical SAEA is consistent with the EvoLS outlined in Algorithm 1, except that the former lacks any approximation methodology selection mechanism (line 11). In addition, EA-perfect, which refers to a canonical SAEA that employs an imaginary approximation method that generates error-free surrogates<sup>6</sup> is also considered

<sup>6</sup>An error-free surrogate model is realized by using an exact fitness function for evaluation inside the local search strategy where a surrogate model should be used. Note that the incurred fitness evaluation here is not counted as part of the computational budget.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

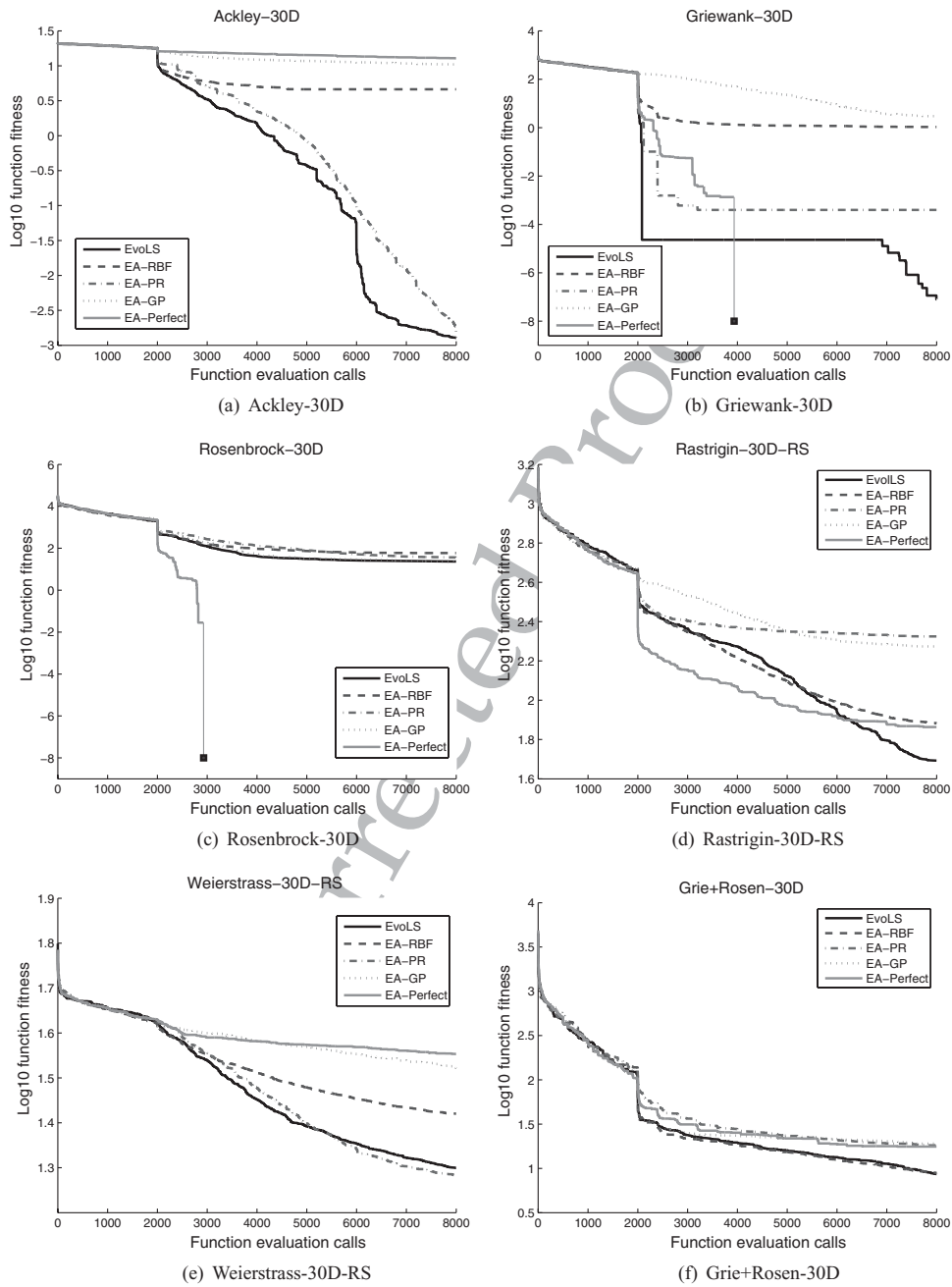


Figure 2: Performance of EvoLS and SAEAs on the benchmark problems. Note that the average fitness values are plotted in logarithmic scale to highlight the differences in the search performance.

Table 3: Results of Wilcoxon test at 95% confidence level, for EvoLS and other SAEAs in solving the 30D benchmark problems. Note that s+, s-, or  $\approx$  indicates that EvoLS is significantly statistically better, worse, or indifferent, respectively.

Algorithm	EA-GP	EA-PR	EA-RBF	EA-Perfect
$F_{\text{Ackley}}$	s+	$\approx$	s+	s+
$F_{\text{Griewank}}$	s+	s-	s+	s-
$F_{\text{Rosenbrock}}$	s-	s+	s+	s-
$F_{\text{Rastrigin-SR}}$	s+	s+	s+	s+
$F_{\text{Weierstrass-SR}}$	s+	$\approx$	s+	s+
$F_{\text{Grie+Rosen}}$	s+	s+	$\approx$	s+

here to assess the benefits of using evolvability measure versus approximation errors as the criterion for model selection.

The averaged convergence trends obtained by EvoLS on the benchmark problems as a function of the total number of exact fitness function evaluations are summarized in Figures 2(a)–(f). The results presented here are the average of 20 independent runs for each test problem. Also shown in the figures are the averaged convergence trends obtained using the canonical SAEAs, that is, EA-RBF, EA-PR, EA-GP, and EA-perfect. Using the statistical nonparametric Wilcoxon test at a 95% confidence level, the search performances of EvoLS when pitted against each SAEA considered on solving the set of benchmark problems are then tabulated in Table 3. To facilitate a fair comparison study, the parametric configurations of EvoLS and SAEAs are maintained consistent, as defined in Table 2. The detailed numerical results of the algorithms considered in the study are summarized in Table 4.

Statistical analysis shows that EvoLS fares competitively or significantly outperforms most of the methods considered, at a 95% confidence level on the 30-dimensional benchmark problems. In the ideal case of the error-free surrogate model, denoted as EA-perfect, the results indicated that the search on surrogate model with low approximation error does not always lead to better performance over the others. As one would expect, the EA-perfect exhibits superior performance on the low-modality Rosenbrock (with two local optima; Shang and Qiu, 2006) and Griewank (which has a relatively smooth fitness landscape at high dimensionalities; Le et al., 2009), as observed in Figures 2(b)–(c), due to the high efficacy of the local search strategy used on a perfect or error-free model. However, on problems with highly multimodal and rugged landscapes including Ackley, Rastrigin-SR Weierstrass-SR, and expanded Griewank plus Rosenbrock (see Figures 2(a), and 2(d–f)), EvoLS, which operates on inferring the most suitable approximation method according to their evolvability measure, is observed to outperform EA-perfect significantly.

Focusing on the Griewank function, it is worth noting that EvoLS exhibits more robust search than EA-PR, as indicated by the relatively lower standard deviation in solution quality (as shown in Table 4). Similar robustness in the EvoLS over EA-GP can also be observed on the Rosenbrock function. On the other hand, the competitive performance displayed by EA-PR and EvoLS on three out of the six benchmark problems suggests both benefit from the landscape smoothing effect of low-order PR. Such a phenomenon is established as the blessing of uncertainty in SAEA (Lim et al., 2010). Overall, the results from the statistical tests confirmed the robustness and superiority of the EvoLS over those that assumed a single fixed approximation methodology throughout the search (EA-GP, EA-PR, EA-RBF, and EA-perfect).

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

Table 4: Statistics of the converged solution quality at the end of 8,000 exact function evaluations for EA-GP, EA-PR, EA-RBF, and EvoLS on benchmark problems. Note that the value will be rounded to zero for performance quality lower than a precision accuracy of  $1E - 8$ .

Optimization algorithm	Mean	SD	Median	Best	Worst
Ackley (F1)					
EA-GP	1.05E+01	4.42E+00	1.23E+01	1.17E-03	1.53E+01
EA-PR	1.54E-03	8.34E-04	1.30E-03	4.93E-04	3.39E-03
EA-RBF	4.62E+00	1.67E+00	4.47E+00	2.66E+00	6.40E+00
EA-Perfect	1.28E+01	1.17E+00	1.30E+01	9.73E+00	1.42E+01
EvoLS	1.28E-03	9.84E-04	1.13E-03	1.32E-04	3.40E-03
Griewank (F2)					
EA-GP	2.67E+00	1.12E+01	1.88E-02	4.02E-05	5.02E+01
EA-PR	3.96E-04	1.27E-03	0.00E+00	0.00E+00	5.04E-03
EA-RBF	1.07E+00	3.03E-01	1.10E+00	2.43E-01	1.47E+00
EA-Perfect	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
EvoLS	7.89E-08	2.80E-07	0.00E+00	0.00E+00	1.17E-06
Rosenbrock (F3)					
EA-GP	2.29E+01	1.76E+01	1.92E+01	1.47E+01	9.70E+01
EA-PR	3.63E+01	2.29E+01	2.81E+01	2.66E+01	1.18E+02
EA-RBF	5.90E+01	2.15E+01	5.57E+01	3.07E+01	1.02E+02
EA-Perfect	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
EvoLS	2.32E+01	1.66E+00	2.29E+01	2.11E+01	2.66E+01
Shifted Rotated Rastrigin (F4)					
EA-GP	1.88E+02	7.42E+01	1.79E+02	6.67E+01	3.44E+02
EA-PR	2.11E+02	1.36E+01	2.12E+02	1.86E+02	2.39E+02
EA-RBF	7.63E+01	2.86E+01	8.02E+01	3.43E+01	1.38E+02
EA-Perfect	7.30E+01	1.54E+01	6.99E+01	4.97E+01	1.09E+02
EvoLS	4.93E+01	1.66E+01	4.73E+01	2.36E+01	8.16E+01
Shifted Rotated Weierstrass (F5)					
EA-GP	3.33E+01	4.42E+00	3.42E+01	2.44E+01	3.99E+01
EA-PR	1.91E+01	2.61E+00	1.93E+01	1.38E+01	2.37E+01
EA-RBF	2.63E+01	2.97E+00	2.63E+01	2.08E+01	3.29E+01
EA-Perfect	3.57E+01	2.64E+00	3.58E+01	3.15E+01	4.12E+01
EvoLS	1.99E+01	2.69E+00	1.97E+01	1.52E+01	2.52E+01
Expanded Griewank plus Rosenbrock (F6)					
EA-GP	1.90E+01	4.58E+00	1.87E+01	1.20E+01	2.81E+01
EA-PR	1.81E+01	1.04E+00	1.81E+01	1.61E+01	2.04E+01
EA-RBF	8.85E+00	2.03E+00	9.25E+00	5.84E+00	1.17E+01
EA-Perfect	1.76E+01	6.26E+00	1.66E+01	9.27E+00	3.67E+01
EvoLS	8.60E+00	1.78E+00	8.77E+00	6.27E+00	1.15E+01

In Table 5, the percentage of savings in computational budget by EvoLS (in terms of the number of calls made to the original computational expensive function) to arrive at equivalent solution quality to the respective algorithms on the 30D benchmark problems, are tabulated. Note that only those algorithms that are significantly outperformed by EvoLS on the respective benchmark problem are discussed here. The superior efficiency of the EvoLS over the counterpart algorithms can be observed from the table, where significant savings of 25.32% to 74.91% are noted.

Table 5: Percentage of savings in computational budget by EvoLS on the benchmark problems, in terms of the number of calls made to the original computational expensive function to arrive at equivalent converged solution quality of the respective algorithms.

Algorithm	EA-GP	EA-PR	EA-RBF	EA-perfect
$F_{\text{Ackley}}$	74.60	—	66.91	74.91
$F_{\text{Griewank}}$	74.76	71.75	74.76	—
$F_{\text{Rosenbrock}}$	—	50.27	55.20	—
$F_{\text{Rastrigin-SR}}$	49.64	55.31	26.33	25.32
$F_{\text{Weiertrass-SR}}$	58.57	—	42.28	63.32
$F_{\text{Grie+Rosen}}$	47.67	45.05	—	44.13

### 3.2 Suitability of Surrogates

In this section, the suitability of surrogates in evolutionary search with respect to the benchmark problem of interest, prediction quality, and states of the evolution are investigated and discussed.

#### 3.2.1 Fitness Landscapes

The summarized search performances of the algorithms with single approximation methodology (i.e., RBF, PR, or GP), as tabulated in Table 4, confirm our hypothesis that the suitability of surrogates in an evolutionary search largely depends on problem fitness landscape. For the purpose of our discussion here, we focus the attention on the results for EA-PR and EA-RBF. EA-PR is observed to outperform other SAEAs with single approximation method (GP or RBF) on Ackley and Griewank, while EA-RBF emerged as superior on Rastrigin-SR and expanded Griewank plus Rosenbrock. Through adapting the choice of suitable approximation methodology by means of evolvability learning while the search progresses online, EvoLS attains search performance that is better than or at least competitive with the best performing SAEA (with single approximation method) for the respective problems considered.

#### 3.2.2 Prediction Quality and Fitness Improving Surrogate

In this section, we discuss the correlations between prediction quality and surrogate suitability in bringing about the search improvement in EvoLS. The correlations can be assessed by analyzing the average frequency of usage and the normalized root mean square fitness prediction errors (N-RMSE) of the surrogates (or the associated approximation methodology used to construct the surrogate), across the search,<sup>7</sup> such as that depicted in Figures 3(a) and (b), respectively. From Figure 3, it is notable that the low-error fitness prediction surrogates constructed using RBF have been beneficial in enhancing the search toward the refinement of solutions with improved fitness. The presence of fitness improving surrogates thus naturally led to the high frequency of usage of RBF in EvoLS. Conversely, negative correlations between fitness prediction quality and surrogate suitability may also occur to benefit search. Here, we illustrate one such scenario observed in our study. In spite of the high-error fitness predictions

<sup>7</sup>For each generation  $t$ , the root mean square (RMS) prediction errors of all methodologies (GP, PR, and RBF) are first computed across the independent runs and subsequently normalized, thus bringing the prediction errors to the common scale of  $[0, 1]$  throughout the different stages of the search. Subsequently, for each problem, the normalized RMSEs of GP, PR, and RBF are averaged across all generations of the entire search.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

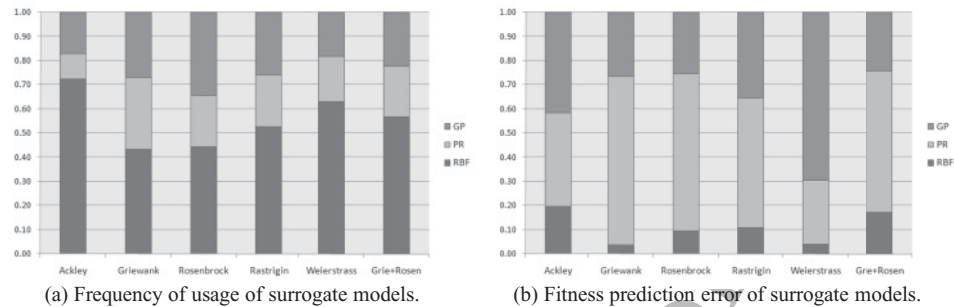


Figure 3: Correlation analyses of EvoLS on the benchmark problems.

exhibited by the second-order PR on Griewank (as can be observed in Figure 3), the persistently high frequency of usage for PR in EvoLS clearly depicts the inference toward productive fitness improving over fitness error-free surrogates.

### 3.2.3 State of Evolution

To assess the suitability of the surrogates at different stages of the evolution, next we analyze the fitness prediction errors of the surrogates and their frequencies of usage in the EvoLS search, which are summarized in Figures 4(a–f). The upper subplot of each figure depicts how often an approximation methodology is inferred as most productive and hence chosen for use within the local search strategy of the EvoLS. The lower subplot, on the other hand, depicts the root mean square fitness prediction errors of the respective surrogates. The results from Figure 4 suggest that no single approximation methodology has served as most suitable throughout the different stages of the search. For instance, RBF is noted to be used more prominently than the other approximation methodologies at the initial stages of the search but then exhibits a decreasing trend in frequency of usage at the later stage of the search on both the Rosenbrock and rotated shifted Rastrigin, see Figures 4(c), (d). Likewise, the fitness prediction qualities of GP are noted to be significantly low at the initial stage of the search on Griewank, as shown in Figure 4(b), but the error eases as the search evolves. Note that such variations in fitness prediction qualities of the PR model can also be observed on Ackley, as shown in Figure 4(a). These results in a way strengthen our aspiration toward the notion of evolvability and hence the significance of EvoLS in facilitating productive cooperation among diverse approximation methodologies, working together to accomplish the shared goal of effective and efficient global optimization in the evolutionary search.

### 3.3 Assessment Against Other Surrogate-Assisted and Evolutionary Search Approaches

In this section, we further assess the proposed EvoLS against other recent state of the art evolutionary approaches reported in the literature. In particular, GS-SOMA (Lim et al., 2010) and IPOP-CMA-ES (Auger and Hansen, 2005), which are well established EAs designed for numerical optimization under the scenarios of limited computational budget, are included here as the state of the art algorithms for comparison. Using a statistical Wilcoxon test at a 95% confidence level, the search performance of each algorithm is pitted against the EvoLS on solving the set of benchmark functions described in Table 1, where the results are tabulated in Table 6. Note that all algorithms considered



Evolution by Adapting Surrogates

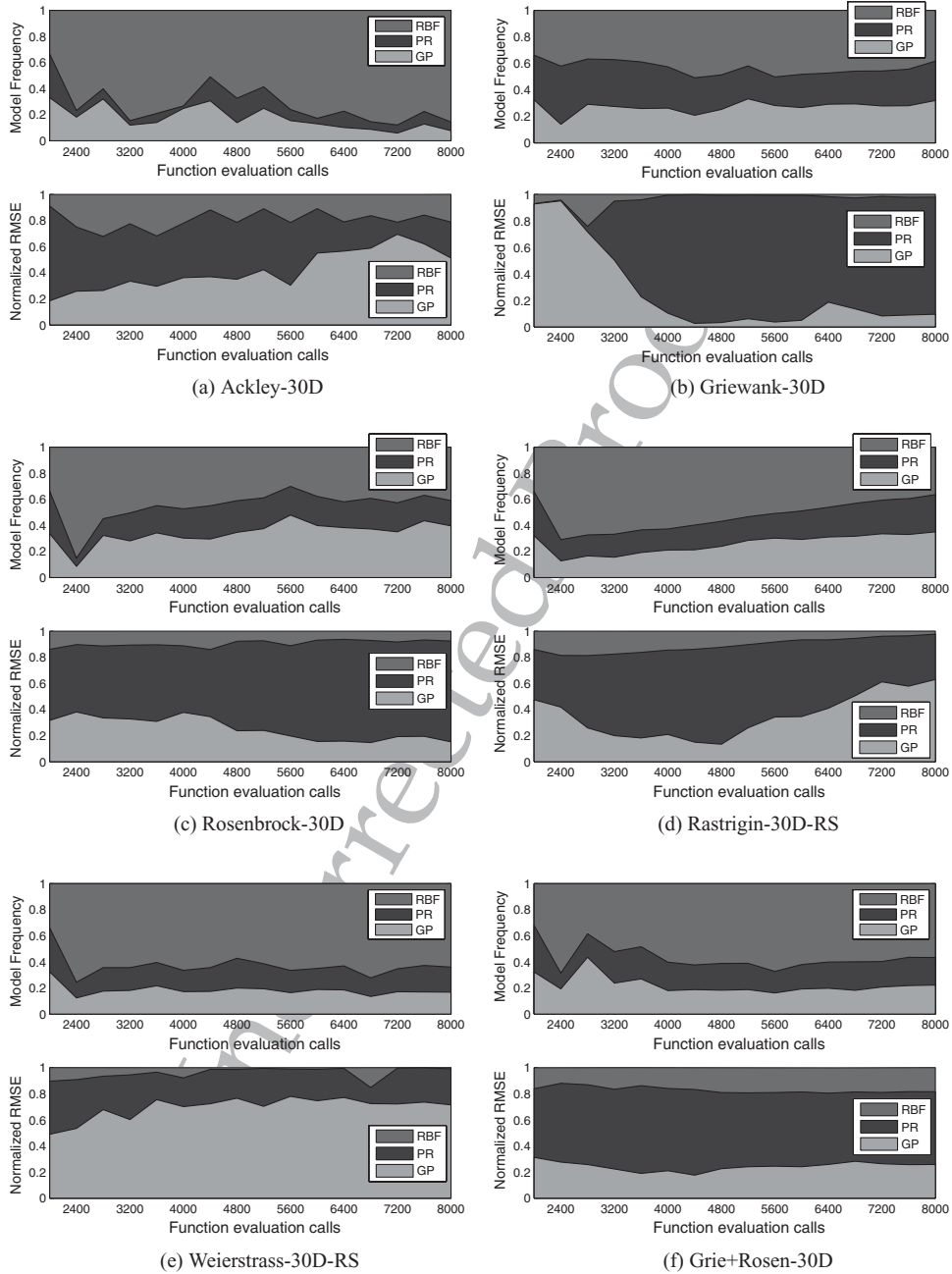


Figure 4: Surrogate frequencies of usage and fitness prediction errors in EvoLS search.

are configured to their default parametric settings as in Lim et al. (2010) and in Auger and Hansen (2005) to facilitate a fair comparison in our study.<sup>8</sup> The detailed statistics

<sup>8</sup>Note that the population size was configured to 100 for GS-SOMA and (7, 14) for IPOP-CMA-ES in the study.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

Table 6: Results of Wilcoxon test at 95% confidence level for EvoLS, GS-SOMA, and IPOP-CMA-ES in solving the 30D benchmark problems. Note that s+, s-, or ≈ indicates that EvoLS is significantly statistically better, worse, or indifferent, respectively.

Algorithm	GS-SOMA	IPOP-CMA-ES
$F_{Ackley}$	s+	s+
$F_{Griewank}$	s+	s+
$F_{Rosenbrock}$	s+	s+
$F_{Rastrigin-SR}$	s+	≈
$F_{Weierstrass-SR}$	s+	s+
$F_{Grie+Rosen}$	s+	≈

Table 7: Statistics of the converged solution quality at the end of 8,000 exact function evaluations for GS-SOMA, IPOP-CMA-ES, and EvoLS on benchmark problems. Note that the value will be rounded to zero for the performance quality lower than the precision accuracy of  $1 E - 8$ .

Optimization algorithm	Mean	SD	Median	Best	Worst
Ackley (F1)					
GS-SOMA	3.58E+00	5.09E-01	3.67E+00	2.87E+00	4.28E+00
IPOP-CMA-ES	5.89E+00	9.24E+00	9.67E-04	4.14E-08	2.03E+01
EvoLS	7.43E-05	9.16E-05	4.22E-05	7.48E-06	3.71E-04
Griewank (F2)					
GS-SOMA	2.20E-03	4.60E-03	0.00E+00	0.00E+00	1.54E-02
IPOP-CMA-ES	1.97E-03	5.45E-03	0.00E+00	0.00E+00	2.22E-02
EvoLS	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Rosenbrock (F3)					
GS-SOMA	4.63e+01	2.92e+01	3.02e+01	2.83e+01	1.26e+02
IPOP-CMA-ES	2.24E+01	1.82E+00	2.21E+01	1.98E+01	2.63E+01
EvoLS	2.01E+01	1.71E+00	2.00E+01	1.71E+01	2.31E+01
Shifted rotated Rastrigin (F4)					
GS-SOMA	2.04E+02	1.60E+01	2.07E+02	1.66E+02	2.30E+02
IPOP-CMA-ES	6.11E+01	4.16E+01	5.11E+01	3.28E+01	2.31E+02
EvoLS	4.72E+01	1.79E+01	4.75E+01	2.36E+01	7.48E+01
Shifted rotated Weierstrass (F5)					
GS-SOMA	2.60E+01	3.05E+00	2.90E+01	2.40E+01	3.40E+01
IPOP-CMA-ES	2.22E+01	2.62E+00	2.24E+01	1.70E+01	2.95E+01
EvoLS	1.96E+01	2.14E+00	1.97E+01	1.52E+01	2.27E+01
Expanded Griewank plus Rosenbrock (F6)					
GS-SOMA	1.80E+01	1.05E+00	1.77E+01	1.70E+01	1.90E+01
IPOP-CMA-ES	3.42E+00	6.27E-01	3.44E+00	1.93E+00	4.70E+00
EvoLS	3.34E+00	4.32E-01	3.41E+00	2.46E+00	4.09E+00

of the algorithms from 20 independent runs on the numerical errors with respect to the global optimum are provided separately in Table 7. From the statistical results in Table 6, EvoLS is observed to fare competitively or significantly outperforms the state of the art methods GS-SOMA and IPOP-CMA-ES, at a 95% confidence level on the 30-dimensional benchmark problems, thus highlighting the robustness and superiority of EvoLS.

#### 4 Real-World Application: Aerodynamic Optimization of the Rear of a Car Model

Our ultimate motivation of the present work lies in the difficulties and challenges posed by computationally expensive real-world applications. In this section, we consider the proposed EvoLS for the design of a quasi-realistic aerodynamic car rear using a simplified model version of the Honda Civic. The objective is to minimize the aerodynamic performance calculation of the design, that is, the total drag of the car rear. The design model of the Honda Civic used in the present study is labeled here as the Baseline-C-Model. Aerodynamic car rear design is an extremely complex task that is normally undertaken over an extended time period and at different levels of complexity. In this application, an aerodynamic performance calculation of the design requires a CFD simulation that generally takes around 60 min of wall-clock time on a Quad-Core machine. For the calculation, the OpenFOAM CFD flow solver (Jasak et al., 2007; Jasak and Rusche, 2009) allows a tight integration into the optimization process, providing an automatic meshing procedure as well as parallelization.

The choice of an adequate geometrical representation of the car simulation model that can be reasonably coupled to the optimization algorithm is also crucial. In the presented experiments, the state of the art free form deformation (FFD; Sederberg and Parry, 1986; Coquillart, 1990) has been chosen as the geometrical representation since it provides a fair trade-off between design flexibility and scalable number of optimization parameters. FFD is a shape morphing technique that allows smooth design changes using an arbitrary number of parameters which are intuitively adjustable to the problem at hand. The benefits of FFD can be found in Menzel and Sendhoff (2008).

For the technical realization, the application of FFD requires a control volume, that is, a lattice of control points, in which the model geometry is embedded. In the next step, the geometry is transferred to the spline parameter space of the control volume, a numerical process which is called freezing. After the object is frozen, it is possible to select and move single or several control points to generate deformed shape variations. To amplify the surface deformations, it is important to position the control points close to the car body. Figure 5 illustrates the Baseline-C-Model as well as the implemented FFD control volume. Based on this setup, 10 optimization parameters  $p_i$  were chosen. Each parameter is composed of a group of 22 control points within one layer of the control volume as marked by the dashed box in the lower left image of Figure 5. Since the model is symmetric to the center plane in the  $y$  direction, the parameters affect the left and right side of the car in the same way.

During the evaluation step of the optimization process, the aerodynamic performance of each design solution, that is, the total drag of the car, is calculated. Therefore, each of the 10 parameters stored in the design solution or individual are extracted and added as an offset on the corresponding layer of control points, either in the  $x$  direction ( $p_1, p_3, p_5, p_7, p_9$ ) or in the  $z$  direction ( $p_2, p_4, p_6, p_8, p_{10}$ ). Second, based on the modified control points, the car shape is updated using FFD and stored as a triangulated mesh, that is, the STL file format. Third, the external airflow around the car is computed by the OpenFOAM CFD solver. Therefore, a CFD calculation grid around the updated car shape has to be generated and is automatically carried out using the snappyHexMesh tool of OpenFOAM. Based on this grid, the external airflow is calculated, resulting in a drag value that is monitored every 10th time step of the simulation. After the calculation has sufficiently converged, the total drag value is extracted and assigned to the individual.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

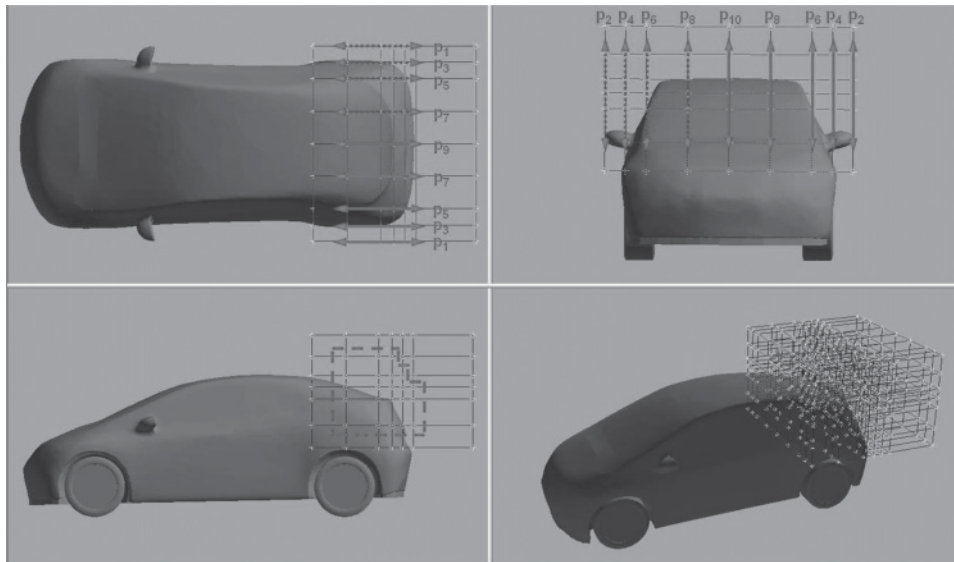


Figure 5: Experiment setup.

#### 4.1 Fitness Landscape Analysis

To conduct a preliminary analysis on the conditions of the problem fitness landscape, the effect of design variables  $p_i$  on drag at the car rear is investigated in this section. In particular, the design solutions generated randomly using the Latin hypercube sampling technique and the corresponding aerodynamic performance (i.e., total drag) from the SimpleBase-C-Model landscape are first grouped into different bins based on each control variable  $p_i$ . The average fitness and standard deviation of each bin is plotted against the control variable for the sensitivity analysis. The plots of 700 samples drawn from the search range of  $[-0.4, 0.4]^{10}$  for each dimension  $p_i$ , which is divided into 10 bins, are presented in Figure 6.

The sensitivity analysis of the design variables presented in Figure 6 depicts the relationships between each input variable and the objective function. In this case, the nonmonotonic sensitivity of the problem landscape with respect to each design variable was observed, especially in Figures 6(a), 6(b), 6(c), 6(f), 6(h), and 6(j). The dependencies of the objective function to the interactions of multiple design variables as implied by the nonmonotonic trends in Figure 6 thus highlighted the landscape ruggedness of the quasi-realistic aerodynamic car rear design problem.

#### 4.2 Optimization Results

Due to the highly computationally expensive CFD simulations, a computational budget of 200 exact evaluations (200 hr) is used for one optimization run in our study. A small population size of 10 was considered in EvoLS. Note that no database building phase (i.e.,  $G_{db} = 0$ ) is required by EvoLS in this case since sufficient sample data of the search space was available at hand. Figure 7 shows convergence trend of the best run for the aerodynamic car rear design problem obtained by EvoLS (as described in Section 3) after 200 evaluations on the aerodynamic car rear design problem. For comparison, the optimization results obtained previously based on the covariance matrix

Evolution by Adapting Surrogates

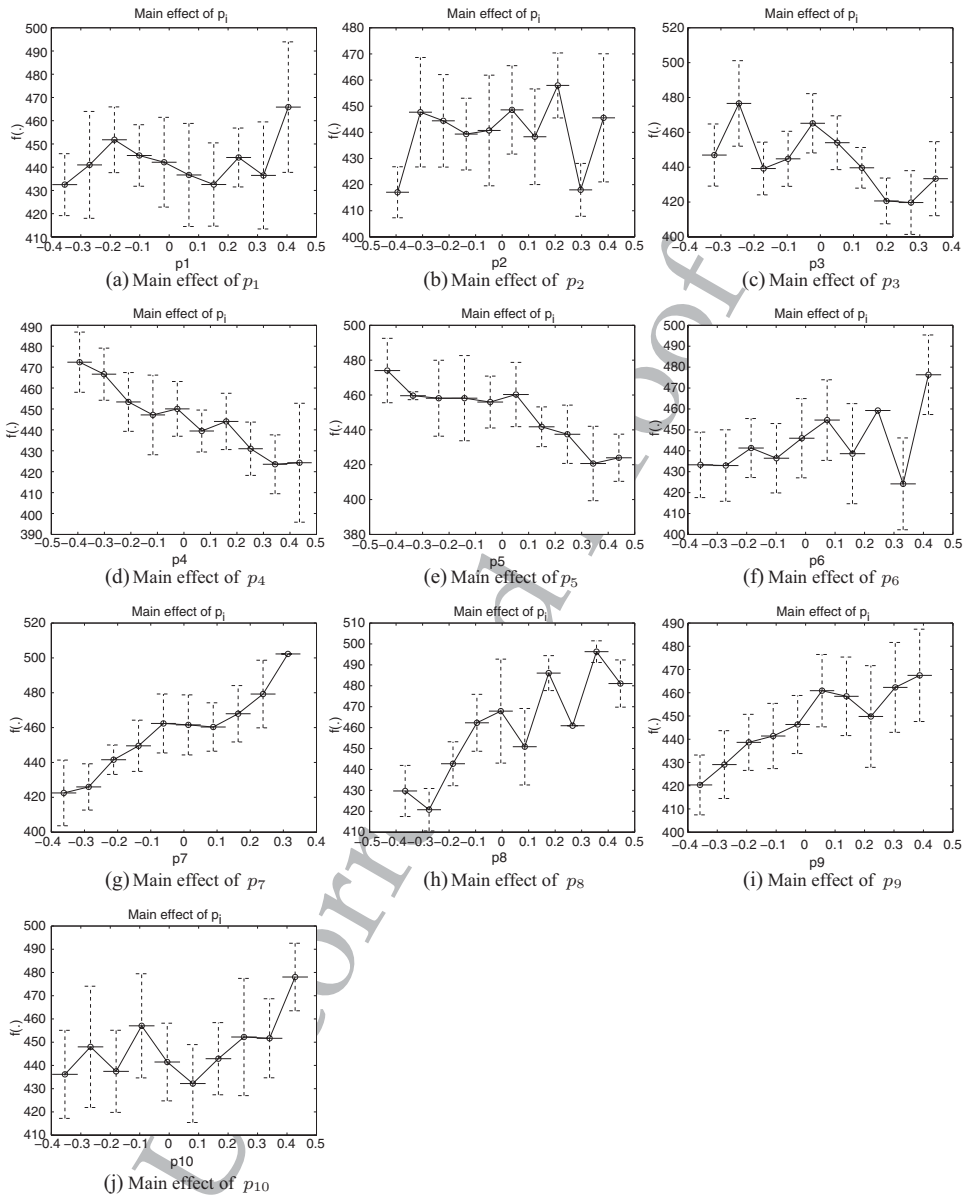


Figure 6: Sensitivity analysis of the aerodynamic car rear design problem.

adaptation evolution strategy (Hansen and Kern, 2004), including CMA-ES(5, 10) and CMA-ES(1, 10), are also reported in Figure 7. As shown in the figure, due to the complexity of the problem landscape, CMA-ES(1, 10) as an individual-based search strategy performed worst as compared to the other population-based approaches. Among the algorithms considered, EvoLS exhibited the best performance by locating the car rear design with the lowest drag value of 403.573. The search trends also show that the proposed EvoLS arrived at the best design solution discovered by CMA-ES(5, 10)

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

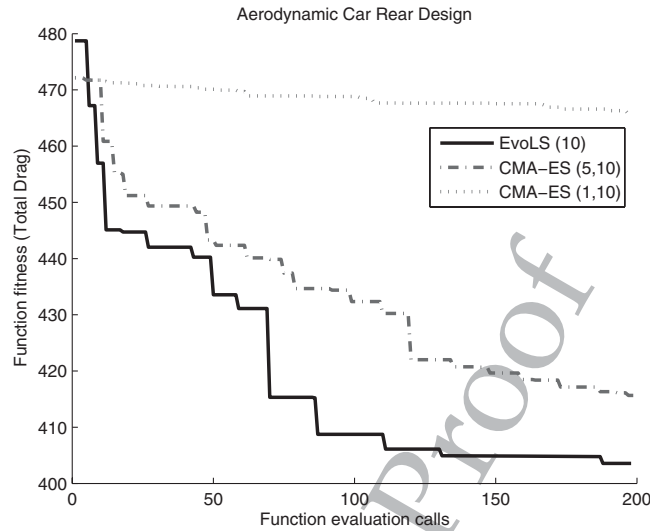


Figure 7: Convergence trace on aerodynamic car rear design.

using only 1/2 of the computational budget incurred by the latter. The computational saving of more than 50% and improved solution quality attained thus show that in a practical setting the usage of the proposed EvoLS can be highly beneficial as compared to standard techniques for solving challenging computationally expensive design optimization.

## 5 Conclusions

In this paper, we have presented a novel EvoLS framework that operates on multiple approximation methodologies of diverse characteristics. The suitability of an approximation method in the construction of surrogates for guiding the search is assessed by the evolvability metric instead of solely focusing on the fitness prediction error. By constructing the respective surrogate using the most productive approximation method inferred for each individual solution in the population, EvoLS serves as a self-configurable surrogate-assisted memetic algorithm for optimizing computationally expensive problems at improved search performance.

Numerical study of EvoLS with an assessment made against the use of single approximation methodology or the imaginary perfect surrogates as well as other state of the art algorithms on commonly used benchmark problems demonstrated the robustness and efficiency of EvoLS. Further analysis on the suitability of surrogates operating within EvoLS also confirmed our motivations to introduce the evolvability measure that take into account the current state of the search, properties of the search operators, and characteristics of the fitness landscape, while the search progresses online. EvoLS thus serves as a novel effort toward the design of frameworks that promote competition and cooperation among diverse approximation methodologies, working together to accomplish the shared goal of global optimization in the evolutionary search. Last but not least, the results on the real-world computationally expensive problem of aerodynamic car rear design further highlighted the competitiveness of EvoLS in attaining improved search performance.

## Acknowledgments

M. N. Le is grateful to the financial support of Honda Research Institute Europe. The authors would like to thank the editor and reviewers for their thoughtful suggestions and constructive comments.

## References

- Acar, E., and Rais-Rohani, M. (2009). Ensemble of metamodels with optimized weight factors. *Structural and Multidisciplinary Optimization*, 37(3):279–294.
- Alexandrov, N. M., Dennis, J. E., Lewis, R. M., and Torczon, V. (1998). A trust-region framework for managing the use of approximation models in optimization. *Structural and Multidisciplinary Optimization*, 15(1):16–23.
- Auger, A., and Hansen, N. (2005). A restart CMA evolution strategy with increasing population size. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation, CEC'05*, Vol. 2, pp. 1769–1776.
- Bishop, C. M. (1996). *Neural networks for pattern recognition*. Oxford, UK: Oxford University Press.
- Booker, A. J., Dennis, J. E., Frank, P. D., Serafini, D. B., Torczon, V., and Trosset, M. W. (1999). A rigorous framework for optimization of expensive functions by surrogates. *Structural and Multidisciplinary Optimization*, 17(1):1–13.
- Branke, J., and Schmidt, C. (2005). Faster convergence by means of fitness estimation. *Soft Computing—A Fusion of Foundations, Methodologies and Applications*, 9(1):13–20.
- Buche, D., Schraudolph, N. N., and Koumoutsakos, P. (2005). Accelerating evolutionary algorithms with Gaussian process fitness function models. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 35(2):183–194.
- Coquillart, S. (1990). Extended free-form deformation: A sculpturing tool for 3D geometric modeling. *Computer Graphics*, 24(4):187–196.
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.
- DaCosta, L., Fialho, A., Schoenauer, M., and Sebag, M. (2008). Adaptive operator selection with dynamic multi-armed bandits. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation*, pp. 913–920.
- Donoho, D. L. (2000). High-dimensional data analysis: The curses and blessings of dimensionality. In *Proceedings of the American Mathematical Society Conference on Math Challenges of the 21st Century*.
- Emmerich, M., Giotis, A., Özdemir, M., Bäck, T., and Giannakoglou, K. (2002). Metamodel-assisted evolution strategies. *Parallel Problem Solving from Nature, PPSN VII*, Vol. 2439, pp. 361–370.
- Fielding, A. H., and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1):38–49.
- Giannakoglou, K. C. (2002). Design of optimal aerodynamic shapes using stochastic optimization methods and computational intelligence. *Progress in Aerospace Sciences*, 38(1):43–76.
- Goel, T., Haftka, R. T., Shyy, W., and Queipo, N. V. (2007). Ensemble of surrogates. *Structural and Multidisciplinary Optimization*, 33(3):199–216.
- Goldberg, D. E. (1990). Probability matching, the magnitude of reinforcement, and classifier system bidding. *Machine Learning*, 5(4):407–425.
- Hansen, N., and Kern, S. (2004). Evaluating the CMA evolution strategy on multimodal test functions. In *Parallel Problem Solving from Nature, PPSN VIII*, pp. 282–291.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

- Hart, W. E., Krasnogor, N., and Smith, J. E. (2004). Editorial introduction; Special issue on memetic algorithms. *Evolutionary Computation*, 12(3):v–vi.
- Herrera, F., Lozano, M., and Verdegay, J. L. (1998). Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial Intelligence Review*, 12(4):265–319.
- Hirsch, C. (2007). *Numerical computation of internal and external flows: Fundamentals of computational fluid dynamics*, 2nd ed. Woburn, MA: Butterworth-Heinemann.
- Jasak, H., Jemcov, A., and Tukovic, Z. (2007). OpenFOAM: A C++ library for complex physics simulations. *International Workshop on Coupled Methods in Numerical Dynamics*.
- Jasak, H., and Rusche, H. (2009). Dynamic mesh handling in OpenFOAM. In *Proceedings of the American Institute of Aeronautics and Astronautics*.
- Jin, J. (2002). *The finite element method in electromagnetics*, 2nd ed. New York: Wiley.
- Jin, R., Chen, W., and Simpson, T. W. (2001). Comparative studies of metamodeling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization*, 23(1):1–13.
- Jin, Y. (2005). A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing—A Fusion of Foundations, Methodologies and Applications*, 9(1):3–12.
- Jin, Y., Olhofer, M., and Sendhoff, B. (2000). On evolutionary optimization with approximate fitness functions. In *Proceedings of the Genetic and Evolutionary Computation*, pp. 786–793.
- Jin, Y., Olhofer, M., and Sendhoff, B. (2002). A framework for evolutionary optimization with approximate fitness functions. *IEEE Transactions on Evolutionary Computation*, 6(5):481–494.
- Jin, Y., and Sendhoff, B. (2004). Reducing fitness evaluations using clustering techniques and neural networks ensembles. In *Proceedings of the Genetic and Evolutionary Computation Conference. Lecture notes in computer science*, Vol. 3102 (pp. 688–699). Berlin: Springer-Verlag.
- Keane, A. J., and Petruzzelli, N. (2000). Aircraft wing design using GA-based multi-level strategies. In *Proceedings of the AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*.
- Kim, H.-S., and Cho, S.-B. (2002). An efficient genetic algorithm with less fitness evaluation by clustering. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 887–894.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1137–1145.
- Krasnogor, N., and Smith, J. (2005). A tutorial for competent memetic algorithms: Model, taxonomy, and design issues. *IEEE Transactions on Evolutionary Computation*, 9(5):474–488.
- Le, M. N., Ong, Y. S., Jin, Y., and Sendhoff, B. (2009). Lamarckian memetic algorithms: Local optimum and connectivity structure analysis. *Memetic Computing*, 1(3):175–190.
- Lesh, F. H. (1959). Multi-dimensional least-squares polynomial curve fitting. *Communications of the ACM*, 2(9):29–30.
- Liang, K.-H., Yao, X., and Newton, C. (1999). Combining landscape approximation and local search in global optimization. In *Proceedings of the 1999 Congress on Evolutionary Computation*, pp. 1514–1520.
- Lim, D., Jin, Y., Ong, Y. S., and Sendhoff, B. (2010). Generalizing surrogate-assisted evolutionary computation. *IEEE Transactions on Evolutionary Computation*, 14(3):329–355.
- Lim, D., Ong, Y. S., Jin, Y., and Sendhoff, B. (2007). A study on metamodeling techniques, ensembles, and multi-surrogates in evolutionary computation. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pp. 1288–1295.



## Evolution by Adapting Surrogates

- Lim, D., Ong, Y. S., Jin, Y., and Sendhoff, B. (2008). Evolutionary optimization with dynamic fidelity computational models. In *Proceedings of the International Conference on Advanced Intelligent Computing Theories and Applications with Aspects of Artificial Intelligence. Lecture notes in artificial intelligence*, Vol. 5227 (pp. 235–242). Berlin: Springer-Verlag.
- Mackay, D. J. (1998). Introduction to Gaussian processes. *Neural Networks and Machine Learning*, 168:133–165.
- Menzel, S., and Sendhoff, B. (2008). Representing the change—Free form deformation for evolutionary design optimisation. In *Evolutionary computation in practice* (pp. 63–86). Berlin: Springer-Verlag.
- Ong, Y. S., Lim, M. H., and Chen, X. S. (2010). Memetic computation—Past, present & future. *IEEE Computational Intelligence Magazine*, 5(2):24–31.
- Ong, Y. S., Lim, M. H., Zhu, N., and Wong, K. W. (2006). Classification of adaptive memetic algorithms: A comparative study. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 36(1):141–152.
- Ong, Y. S., Nair, P. B., and Keane, A. J. (2003). Evolutionary optimization of computationally expensive problems via surrogate modelling. *American Institute of Aeronautics and Astronautics Journal*, 41(4):687–696.
- Ong, Y. S., Nair, P. B., Keane, A. J., and Wong, K. W. (2004). Surrogate-assisted evolutionary optimization frameworks for high-fidelity engineering design problems. *Knowledge Incorporation in Evolutionary Computation* (pp. 307–332). Berlin: Springer-Verlag.
- Ong, Y. S., Zhou, Z., and Lim, D. (2006). Curse and blessing of uncertainty in evolutionary algorithm using approximation. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 2928–2935.
- Powell, M. J. D. (1987). Radial basis functions for multivariable interpolation: A review. In *Algorithms for approximation* (pp. 143–167). New York: Clarendon Press.
- Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41(1):1–28.
- Regis, R. G., and Shoemaker, C. A. (2004). Local function approximation in evolutionary algorithms for the optimization of costly functions. *IEEE Transactions on Evolutionary Computation*, 8(5):490–505.
- Sanchez, E., Pintos, S., and Queipo, N. V. (2008). Toward an optimal ensemble of kernel-based approximations with engineering applications. *Structural and Multidisciplinary Optimization*, 36(3):247–261.
- Santana-Quintero, L. V., Montaña, A. A., and Coello, C. A. C. (2010). A review of techniques for handling expensive functions in evolutionary multi-objective optimization. In Y. Tenne and C.-K. Goh (Eds.), *Computational Intelligence in Expensive Optimization Problems*, (pp. 29–59). Berlin: Springer-Verlag.
- Sederberg, T. W., and Parry, S. R. (1986). Free-form deformation of solid geometric models. *Computer Graphics*, 20(4):151–160.
- Serafini, D. B. (1998). A framework for managing models in nonlinear optimization of computationally expensive functions. PhD thesis, Rice University, Houston, Texas.
- Shang, Y.-W., and Qiu, Y.-H. (2006). A note on the extended Rosenbrock function. *Evolutionary Computation*, 14(1):119–126.
- Shi, L., and Rasheed, K. (2010). A survey of fitness approximation methods applied in evolutionary algorithms. In Y. Tenne and C.-K. Goh (Eds.), *Computational Intelligence in Expensive Optimization Problem* (pp. 3–28). Berlin: Springer-Verlag.

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

- Song, W. (2002). *Shape optimisation of turbine blade firtrrees*. PhD thesis, School of Engineering Sciences, University of Southampton, Southampton, UK.
- Tenne, Y. (2009). A model-assisted memetic algorithm for expensive optimization problems. In R. Chiong (Ed.), *Nature-inspired algorithms for optimisation* (pp. 133–169). Berlin: Springer-Verlag.
- Tenne, Y., and Armfield, S. W. (2008a). A versatile surrogate-assisted memetic algorithm for optimization of computationally expensive functions and its engineering applications. In A. Yang, Y. Shan, and L. T. Bui, *Success in evolutionary computation* (pp. 43–72). Berlin: Springer-Verlag.
- Tenne, Y., and Armfield, S. W. (2008b). Metamodel accuracy assessment in evolutionary optimization. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 1505–1512.
- Thierens, D. (2005). An adaptive pursuit strategy for allocating operator probabilities. In *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation*, pp. 1539–1546.
- Ulmer, H., Streichert, F., and Zell, A. (2004). Evolution strategies assisted by Gaussian processes with improved preselection criterion. In *Proceedings of the 2003 Congress on Evolutionary Computation, CEC '03*, pp. 692–699.
- Valiant, L. G. (2009). Evolvability. *Journal of the Association for Computing Machinery*, 56(1):1–21.
- Vapnik, V. N. (1998). *Statistical learning theory*. New York: Wiley.
- Voutchkov, I., and Keane, A. (2010). Multi-objective optimization using surrogates. In Y. Tenne and C.-K. Goh (Eds.), *Computational Intelligence in Optimization* (pp. 155–175). Berlin: Springer-Verlag.
- Wagner, G. P., and Altenberg, L. (1996). Perspective: Complex adaptations and the evolution of evolvability. *Evolution*, 50(3):967–976.
- Zerpa, L. E., Queipo, N. V., Pintos, S., and Salager, J.-L. (2005). An optimization methodology of alkaline-surfactant-polymer flooding processes using field scale numerical simulation and multiple surrogates. *Journal of Petroleum Science and Engineering*, 47(3–4):197–208.
- Zhou, Z., Ong, Y. S., Nair, P. B., Keane, A. J., and Lum, K. Y. (2006). Combining global and local surrogate models to accelerate evolutionary optimization. *IEEE Transactions on Systems, Man and Cybernetics, Part C*, 36(6):814–823.
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-BFGS-B: FORTRAN subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560.
- Zienkiewicz, O. C., and Taylor, R. L. (2006). *The finite element method for solid and structural mechanics*. 6th ed. Woburn, MA: Butterworth-Heinemann.
- Zurada, J. (1992). *Introduction to artificial neural systems*. St. Paul, MN: West Publishing.

## Appendix A: Surrogate Modeling

There exist a variety of approximation methodologies for constructing surrogate models that take their roots from the field of statistical and machine learning (Jin, 2005; Shi and Rasheed, 2010). One popular approach in the design optimization literature is PR or response surface methodology (Lesh, 1959). Neural networks which have shown to be effective tools for function approximation have also been employed as surrogates extensively in evolutionary optimization. These include support vector machines (Cortes and Vapnik, 1995; Vapnik, 1998), artificial neural networks (Zurada, 1992), and radial basis functions (Powell, 1987). A statistically sound alternative is GP

or Kriging (Mackay, 1998), referred to as design and analysis of computer experiment models. Next, we provide a brief overview on three different approximation methods used in the paper, PR, RBF, and GP.

### A.1 Polynomial Regression

The most widely used PR model is the quadratic model, which takes the form

$$\hat{f}_M(\mathbf{x}) = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{x}^{(i)} + \sum_{1 \leq i < j \leq n} \beta_{n-1+i+j} \mathbf{x}^{(i)} \mathbf{x}^{(j)} \quad (6)$$

where  $n$  is the number of input variables,  $\mathbf{x}^{(i)}$  is the  $i$ th component of  $\mathbf{x}$ , and  $\beta_i$  are the coefficients to be estimated. As the number of terms in the quadratic model is  $n_t = (n + 1)(n + 2)/2$  in total, the number of training sample points should be at least  $n_t$  for proper estimation of the unknown coefficients, by means of either least square or gradient-based methods (Jin, 2005).

### A.2 Radial Basic Function

The surrogate models in this category are interpolating radial basic function (RBF) networks of the form

$$\hat{f}_M(\mathbf{x}) = \sum_{i=1}^m \alpha_i K(\|\mathbf{x} - \mathbf{x}_i\|) \quad (7)$$

where  $K(\|\mathbf{x} - \mathbf{x}_i\|) : \mathbb{R}^d \rightarrow \mathbb{R}$  is a RBF and  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_m]^T \in \mathbb{R}^m$  denotes the vector of weights. The number of hidden nodes used in interpolating RBF is often assumed as equal to the number of training vector points.

Typical choices for the kernel function include linear splines, cubic splines, multiquadrics, thin-plate splines, and Gaussian functions (Bishop, 1996). Given a suitable kernel, the weight vector  $\boldsymbol{\alpha}$  can be computed by solving the linear algebraic system of equations

$$\mathbf{K}\boldsymbol{\alpha} = \mathbf{t}$$

where  $\mathbf{t} = [t_1, t_2, \dots, t_m]^T \in \mathbb{R}^m$  denotes the vector of outputs and  $\mathbf{K} \in \mathbb{R}^{m \times m}$  denotes the Gram matrix formed using the training inputs (i.e., the  $ij$ th element of  $\mathbf{K}$  is computed as  $K(\|\mathbf{x}_i - \mathbf{x}_j\|)$ ).

### A.3 Kriging/Gaussian Process

The Kriging model or GP assumes the presence of a global model  $g(\mathbf{x})$  and an additive noise term  $Z(\mathbf{x})$  in the original function.

$$f(\mathbf{x}) = g(\mathbf{x}) + Z(\mathbf{x})$$

where  $g(\mathbf{x})$  is a known function of  $\mathbf{x}$  as a global model of the original function, and  $Z(\mathbf{x})$  is a Gaussian random function with zero mean and nonzero covariance that represents a localized noise or deviation from the global model. Usually,  $g(\mathbf{x})$  is a polynomial, and in many cases, it is reduced to a constant  $\beta$ . The approximation model of  $f(\mathbf{x})$ , given the  $m$  samples and the current input  $\mathbf{x}$ , is defined as:

$$\hat{f}_M(\mathbf{x}) = \hat{\beta} + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{t} - \hat{\beta}\mathbf{I}) \quad (8)$$

where  $\mathbf{t} = [t_1, t_2, \dots, t_m]^T$ ,  $\mathbf{I}$  is a unit vector of length  $m$ , and  $\mathbf{R}$  is the correlation matrix which can be obtained by computing the correlation function between any two samples, that is,  $\mathbf{R}_{i,j} = R(\mathbf{x}_i, \mathbf{x}_j)$ . While the correlation function can be specified by the user, the

M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff

Gaussian exponential correlation function, defined by correlation parameters  $\{\theta_k\}_{k=1}^n$ , has often been used:

$$R(\mathbf{x}_i, \mathbf{x}_j) = \exp \left[ - \sum_{k=1}^n \theta_k |\mathbf{x}_i^{(k)} - \mathbf{x}_j^{(k)}|^2 \right]$$

where  $\mathbf{x}_i^{(k)}$  and  $\mathbf{x}_j^{(k)}$  are the  $k$ th component of sample points  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , respectively.  $\mathbf{r}$  is the correlation vector of length  $m$  between the given input  $\mathbf{x}$  and the samples  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ , that is,  $\mathbf{r} = [R(\mathbf{x}, \mathbf{x}_1), R(\mathbf{x}, \mathbf{x}_2), \dots, R(\mathbf{x}, \mathbf{x}_m)]^T$ .

The estimation of the unknown parameters  $\beta$  and  $\{\theta_k\}_{k=1}^n$  can be carried out using the maximum likelihood method (Shi and Rasheed, 2010). Aside from the approximation values, the Kriging model or GP can also provide a confidence interval without much additional computational cost incurred. However, one main disadvantage of the GP is the significant increase of computational expense when the dimensionality becomes high, due to the matrix inversions in the estimation of parameters.

### Appendix B: Complexity Analysis and Parameters of EvoLS Framework

Here we comment on the computational complexity of present conventional surrogate selection schemes that have roots in the fields of statistical and machine learning (Fielding and Bell, 1997; Queipo et al., 2005; Tenne and Armfield, 2008a). In conventional surrogate selection schemes, multiple sets of sample data are generally segregated, typically into training sets and test sets. For each approximation methodology, the respective surrogate model is commonly constructed based on the training set and the true error is estimated using the test set in the prediction process. This procedure of computation cost  $C_M$  is typically repeated for  $k$  times on different training and test sets to arrive at a statistically sound estimation of the approximation error that is then used in the selection scheme. Although many error estimation approaches are in abundance, the major differences lie mainly on how the training and test sets are generated, which vary from random subsampling (holdout),  $k$ -fold cross-validation, or bootstrapping as described in Kohavi (1995). For  $ID$  number of approximation methodologies considered, the overall computational complexity of the conventional selection scheme in estimating the error of the surrogates can thus be derived as  $O(ID \times k \times C_M)$ .

Next, a complexity analysis of the EvoLS framework is detailed. Apart from the standard parameters of a typical SAEA (Lim et al., 2010), EvoLS has two additional parameters: database  $\Phi_{M_{id}}$  and density function  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$ . Typically, the form of  $P(\mathbf{y}|\mathbf{P}^t, \mathbf{x})$  can be explicitly defined according to the stochastic operator used (as illustrated in Section 3), while databases  $\Phi_{M_{id}}$  naturally follows a first in first out queue structure to favor more recently archived optimization data  $\{(\mathbf{y}, \varphi_{M_{id}}(\mathbf{y}))\}$ . The complexity for EvoLS can be derived as  $O(ID \times |\Phi_{M_{id}}| \times C_E)$  where  $|\Phi_{M_{id}}|$  denotes the database size,  $ID$  denotes the number of approximation methodologies used, and  $C_E$  is the computational effort incurred to determine  $P(\mathbf{y}_i|\mathbf{P}^t, \mathbf{x})$  for each  $\mathbf{y}_i$ . For each individual, since only the most productive approximation methodology inferred is used to construct a new surrogate at a computational requirement of  $C_M$ , the complexity of EvoLS can be derived as  $O(ID \times |\Phi_{M_{id}}| \times C_E + C_M)$ . Nevertheless, as  $(ID \times |\Phi_{M_{id}}| \times C_E) \ll C_M$  in practice, the computational complexity of EvoLS becomes  $O(C_M)$ . Thus, EvoLS offers an alternative to the conventional selection scheme with a significantly lower complexity of  $O(C_M)$  that is independent of the number of approximation methodologies considered in the framework.