Evolutionary Multitasking for Single-objective Continuous Optimization: Benchmark Problems, Performance Metric, and Baseline Results

Bingshui Da[1], Yew-Soon Ong[1], Liang Feng[2], A.K. Qin[3], Abhishek Gupta[1],
Zexuan Zhu[4], Chuan-Kang Ting[5], Ke Tang[6], and Xin Yao[7]

[1]School of Computer Science and Engineering, Nanyang Technological University
[2]College of Computer Science, Chongqing University
[3]Computer Science and Information Technology, RMIT University
[4]College of Computer Science and Software Engineering, Shenzhen University
[5]Department of Computer Science and Information Engineering, National Chung Cheng University
[6]School of Computer Science and Technology, University of Science and Technology of China
[7]School of Computer Science, University of Birmingham

I. INTRODUCTION

The notion of *evolutionary multitasking* has recently emerged as a promising approach for automatically exploiting the latent synergies between distinct (but possibly similar) optimization problems, simply by solving them together in a unified representation space [1]. Inspired by human's cognitive ability to multitask, evolutionary multitasking aims to improve convergence characteristics across multiple optimization problems at once by seamlessly transferring knowledge between them. A particular mode of knowledge transfer that has recently been studied in [1] and [2] is that of *implicit genetic transfer* through chromosomal crossover [3]. While the concept of multitask optimization may seem to bear similarity with that of multitask learning [4], it is noted that the former focuses on exploiting shared knowledge for improved *problem-solving*, in contrast to merely learning.

One of the key practical motivations behind our proposition is to effectively deal with the volume, velocity, and complexity of present-day challenges faced in various industrial environments. In particular, multitasking provides a means to *enhance productivity*. To elaborate, in many real-world settings, it is common for problems with essentially similar properties to recur. In such cases, it makes little sense to re-explore the same design space repeatedly from scratch via *tabula rasa* optimization techniques. Herein, evolutionary multitasking enables autonomous transfer of knowledge from one problem to the next, thereby naturally augmenting the effectiveness of the search.

In addition to enhancing productivity, the process of multitasking can often provide superior quality solutions in comparison to traditional single-task optimization. This is most often true in cases when the fitness landscapes of the constitutive tasks are appropriately complementary. A first step toward quantifying the synergy (or correlation) between optimization problems (with known fitness landscapes) has recently been presented in [5]. It is contended that investigating the correspondence between the measured synergy between problems, and the performance of a multitasking engine, can provide necessary insights into the mechanisms of various genetic operators and how they may be enhanced to better suit the practice of evolutionary multitasking in the future. To this end, the present technical report provides a diverse set of benchmark problems, with associated performance metric and baseline results, to support the future development of the field.

In [5], a metric is proposed for quantifying inter-task correlations between tasks, which involves the calculation of partial derivatives and integrals over the entire search space (and can therefore be time consuming to compute). Accordingly, in this technical report, we propose a simpler derivative/integration-free alternative for rapidly computing inter-task synergies, one that is based on Spearman's rank correlation coefficient (as is described in Section IV). This metric is then used to quantify the synergy between variants of popularly used single-task optimization functions. Based on the aforestated results, a diverse set of nine problem combinations are constructed, that serve as the first set of composite benchmark problems for evolutionary multitasking. A detailed description of the benchmark problem specifications are provided in Section V of this document, with MATLAB implementations of the same made available as supplementary material. Further, sample baseline results have also been provided herein for all the examples.

II. FORMALIZATION OF EVOLUTIONARY MULTITASKING

The concept of evolutionary multitasking has only recently been formalized in [1] under the label of *multifactorial optimization* (MFO), where each constitutive task is considered to impart an additional influence on the evolutionary process of a single population of individuals. In this section, we briefly introduce the associated *multifactorial evolutionary algorithm*

(MFEA) employed to generate baseline results in this report - source code of the MFEA is made available as supplementary material.

A. Preliminary

Consider a hypothetical situation wherein K self-contained optimization tasks are to be performed concurrently. Without loss of generality, all tasks are assumed to be minimization problems. The jth task, denoted T_j, is considered to have a search space X_j on which the objective function is defined as $F_j : \boldsymbol{X}_j \to \mathbb{R}$. In addition, each task may be constrained by several equality and/or inequality conditions that must be satisfied for a solution to be considered feasible. In such a setting, we define MFO as an evolutionary multitasking paradigm that aims to simultaneously navigate the design space of all tasks, constantly building on the implicit parallelism of population-based search so as to rapidly deduce $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{K-1}, \boldsymbol{x}_K\} = argmin\{F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \ldots, F_{K-1}(\boldsymbol{x}), F_K(\boldsymbol{x})\}$, where $\boldsymbol{x}_j$ is a feasible solution in $\boldsymbol{X}_j$. As suggested by the nomenclature, herein each F_j is treated as an additional factor influencing the evolution of a single population of individuals. For this reason, the composite problem may also be referred to as a K-factorial problem.

While designing evolutionary solvers for MFO, it is necessary to formulate a general technique for comparing population members in a multitasking environment. To this end, we first define a set of properties for every individual p_i, where $i \in \{1, 2, , |P|\}$, in a population P. Note that the individuals are encoded in a unified search space $\boldsymbol{Y}$ encompassing $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_K$, and can be decoded into a task-specific solution representation with respect to each of the K optimization tasks. The decoded form of p_i can thus be written as $\{\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{iK}\}$, where $\boldsymbol{x}_{i1} \in \boldsymbol{X}_1$, $\boldsymbol{x}_{i2} \in \boldsymbol{X}_2$, ..., and $\boldsymbol{x}_{iK} \in \boldsymbol{X}_K$.

- *Definition 1*(*Factorial Cost*): For a given task T_j, the *factorial cost* Ψ_{ij} of individual p_i is given by $\Psi_{ij} = \lambda \cdot \delta_{ij} + F_{ij}$; where λ is a large penalizing multiplier, F_{ij} and δ_{ij} are the objective value and the total constraint violation, respectively, of p_i with respect to T_j. Accordingly, if p_i is feasible with respect to T_j (zero constraint violation), we have $\Psi_{ij} = F_{ij}$.
- *Definition 2*(*Factorial Rank*): The *factorial rank* r_{ij} of p_i on task T_j is simply the index of p_i in the list of population members sorted in ascending order with respect to factorial cost Ψ_{ij}.

Note that, while assigning factorial ranks, whenever $\Psi_{1j} = \Psi_{2j}$ for a pair of individuals p_1 and p_2, the parity is resolved by random tie-breaking.

- *Definition 3*(*Skill Factor*): The *skill factor* τ_i of p_i is the one task, amongst all other tasks in a K-factorial environment, with which the individual is associated. If p_i is evaluated for all K tasks then $\tau_i = argmin_j\{r_{ij}\}$, where $j \in \{1, 2, \ldots, K\}$.
- *Definition 4*(*Scalar Fitness*): The *scalar fitness* of p_i in a multitasking environment is given by $\varphi_i = 1/r_{iT}$, where $T = \tau_i$. Notice that $max\{\varphi_i\} = 1$.

Once the fitness of every individual has been scalarized according to Definition 4, performance comparison can then be carried out in a straightforward manner. For example, individual p_1 will be considered to dominate individual p_2 in multifactorial sense simply if $\varphi_1 > \varphi_2$.

It is important to note that the procedure described heretofore for comparing individuals is not absolute. As the factorial rank of an individual, and implicitly its scalar fitness, depends on the performance of every other individual in the population, the comparison is in fact population dependent. Nevertheless, the procedure guarantees that if an individual p^* attains the global optimum of any task then $\varphi^* = 1$, which implies that $\varphi^* \geq \varphi_i$ for all $i \in \{1, 2, \ldots, |P|\}$. Therefore, it can be said that the proposed technique is indeed consistent with the ensuing definition of multifactorial optimality.

- *Definition 5*(*Multifactorial Optimality*): An individual p^* is considered to be optimum in multifactorial sense if there exists at least one task in the K-factorial environment which it globally optimizes.

B. Multifactorial Evolution: A Framework for Effective Multitasking

In this subsection we describe the Multifactorial Evolutionary Algorithm (MFEA), an effective multitasking framework that draws upon the bio-cultural models of multifactorial inheritance [6]. As the workings of the approach are based on the transmission of biological as well as cultural building blocks from parents to their offspring, the MFEA is regarded as belonging to the realm of *memetic computation* [7], [8] – a field that has recently emerged as a successful computational paradigm synthesizing Darwinian principles of natural selection with Dawkins notion of a meme as the basic unit of cultural evolution [9]. An overview of the procedure is provided next.

As shown in Algorithm 1, the MFEA starts by randomly creating a population of n individuals in the unified search space $\boldsymbol{Y}$. Moreover, each individual in the initial population is pre-assigned a specific skill factor (see Definition 3) in a manner that guarantees every task to have fair number of representatives. We would like to emphasize that the skill factor of an individual (i.e., the task with which the individual is associated) is viewed as a computational representation of its pre-assigned cultural trait. The significance of this step is to ensure that an individual is only evaluated with respect to a single task (i.e., only its skill factor) amongst all other tasks in the multitasking environment. Doing so is considered practical since evaluating every individual exhaustively for every task will generally be computationally demanding, especially when K (the number of tasks in the multitasking environment) becomes large. The remainder of the MFEA proceeds similarly to any standard evolutionary procedure. In fact, it must be mentioned here that the underlying genetic mechanisms may be borrowed from any of the plethora of population-based algorithms available in the literature, keeping in mind the properties and requirements of the multitasking problem at hand. The only significant deviation from a traditional approach occurs in terms of offspring evaluation which accounts for cultural traits via individual skill factors.

Algorithm 1 Pseudocode of the MFEA

1: Randomly generate n individuals in $\mathbf{Y}$ to form initial population P_0
2: **for every** p_j in P_0 **do**
3:    Assign skill factor $\tau_j = \mod(j, K) + 1$, for the case of K tasks
4:    Evaluate p_j for task τ_j only
5: **end for**
6: Compute scalar fitness φ_j for every p_j
7: Set $t = 0$
8: **while** stopping conditions are not satisfied **do**
9:    $C_t = \text{Crossover} + \text{Mutate}(P_t)$
10:    **for every** c_j in C_t **do**
11:       Determine skill factor $\tau_j \rightarrow$ Refer Algorithm 2
12:       Evaluate c_j for task τ_j only
13:    **end for**
14:    $R_t = C_t \cup P_t$
15:    Update scalar fitness of all individuals in R_t
16:    Select N fittest members from R_t to form P_{t+1}
17:    Set $t = t + 1$
18: **end while**

Algorithm 2 Vertical cultural transmission via selective imitation

---

Consider offspring $c \in C_t$ where $c = \text{Crossover} + \text{Mutate}(p_1, p_2)$

1: Generate a random number $rand$ between 0 and 1
2: **if** $rand \leq 0.5$ **then**
   c imitates skill factor of p_1
3: **else**
   c imitates skill factor of p_2
4: **end if**

Following the memetic phenomenon of *vertical cultural transmission* [10], [11], [7], offspring in the MFEA experience strong cultural influences from their parents, in addition to inheriting their genes. In gene-culture co-evolutionary theory, vertical cultural transmission is viewed as a mode of inheritance that operates in tandem with genetics, and leads to the phenotype of an offspring being directly influenced by the phenotype of its parents. The algorithmic realization of the aforementioned notion is achieved in the MFEA via a *selective imitation strategy*. In particular, selective imitation is used to mimic the commonly observed phenomenon that offspring tend to imitate the cultural traits (i.e., skill factors) of their parents. Accordingly, in the MFEA, an offspring is only decoded (from the unified genotype space $\mathbf{Y}$ to a task-specific phenotype space) and evaluated with respect to a single task with which at least one of its parents is associated. As has been mentioned earlier, selective evaluation plays a role in managing the computational expense of the MFEA. A summary of the steps involved is provided in Algorithm 2.

C. Search Space Unification

The core motivation behind the evolutionary multitasking paradigm is the autonomous exploitation of known or latent commonalities and/or complementarities between distinct (but possibly similar) optimization tasks for achieving faster and better convergence characteristics. One of the possible means of harnessing the available synergy, at least from an evolutionary perspective, is through implicit genetic transfer during crossover operations. However, for the relevant knowledge to be transferred across appropriately, i.e., to ensure effective multitasking, it is pivotal to first describe a genotypic unification scheme that suits the requirements of the multitasking problem at hand. In particular, the unification serves as a higher-level abstraction that constitutes a *meme space*, wherein building blocks of encoded knowledge are processed and shared across different optimization tasks. This perspective is much in alignment with the workings of the human brain, where knowledge pertaining to different tasks are abstracted, stored, and re-used for relevant problem solving exercises whenever needed.

Unification implies that genetic building blocks [12] corresponding to different tasks are contained within a single pool of genetic material, thereby facilitating the MFEA to process them in parallel. To this end, assuming the search space dimensionality of the jth optimization task (in isolation) to be D_j, a unified search space $\mathbf{Y}$ comprising K (traditionally distinct) tasks may be defined such that $D_{multitask} = max_j\{D_j\}$, where $j \in \{1, 2, \ldots, K\}$. In other words, while handling K optimization tasks simultaneously, the chromosome $\mathbf{y} \in \mathbf{Y}$ of an individual in the MFEA is represented by a vector of $D_{multitask}$ variables. While addressing the jth task, we simply extract D_j variables from the chromosome and decode them into a meaningful solution representation for the underlying optimization task. In most cases, an appropriate selection of D_j task-specific variables from the list of $D_{multitask}$ variables is crucial for the success of multitasking. For instance, if two distinct variables belonging to two different tasks have similar phenotypic (or contextual) meaning, then they should intuitively be associated to the same variable in the unified search space $\mathbf{Y}$. On the other hand, in many naive cases where no a priori understanding about the phenotype space is available, simply extracting the *first* D_j variables from the chromosome can oftentimes be a viable option [1].

In what follows, we demonstrate how chromosomes in a unified genotype space can be decoded into meaningful task-specific solution representations when a *random-key unification scheme* [13] is adopted. According to the random-key scheme, each variable of a chromosome is simply encoded by a *continuous value* in the range [0, 1]. The salient feature of this representation is that it elegantly accommodates a wide variety of problems in continuous as well as discrete optimization, thereby laying the foundation for a cross-domain multitasking platform. Note that this report only focuses on synthetic continuous optimization problems, thereby decoding can be achieved in a straight-forward manner by linearly mapping each random-key from the genotype space to the box-constrained phenotype space of the relevant optimization task [1]. For some guidelines on decoding for discrete optimization problems, the reader is referred to [6].

III. Individual Tasks in Composite Benchmark Problems

In this technical report, we use 7 commonly used optimization functions as ingredients to form synthetic multitask benchmark problems. The definitions of these individual test functions are shown in the following, where D denotes the dimensionality of the search space and $\boldsymbol{x} = [x_1, x_2, \ldots, x_D]$ is the design variable vector.

1) Sphere:

$$F_1(\boldsymbol{x}) = \sum_{i=1}^{D} x_i^2, \quad \boldsymbol{x} \in [-100, 100]^D \tag{1}$$

2) Rosenbrock:

$$F_2(\boldsymbol{x}) = \sum_{i=1}^{D-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2), \tag{2}$$
$$\boldsymbol{x} \in [-50, 50]^D$$

3) Ackley:

$$F_3(\boldsymbol{x}) = -20 \exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} x_i^2}\right)$$
$$- \exp\left(\frac{1}{D}\sum_{i=1}^{D} \cos(2\pi x_i)\right) + 20 + e, \ \boldsymbol{x} \in [-50, 50]^D \tag{3}$$

4) Rastrgin:

$$F_4(\boldsymbol{x}) = \sum_{i=1}^{D} \left(x_i^2 - 10\cos(2\pi x_i) + 10\right), \tag{4}$$
$$\boldsymbol{x} \in [-50, 50]^D$$

5) Griewank:

$$F_5(\boldsymbol{x}) = 1 + \frac{1}{4000}\sum_{i=1}^{D} x_i^2 - \prod_{i=1}^{D} \cos\left(\frac{x_i}{\sqrt{i}}\right), \tag{5}$$
$$\boldsymbol{x} \in [-100, 100]^D$$

6) Weierstrass:

$$F_6(\boldsymbol{x}) = \sum_{i=1}^{D} \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos\left(2\pi b^k(x_i + 0.5)\right)\right]\right)$$
$$- D\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5)\right] \tag{6}$$
$$a = 0.5, b = 3, k_{\max} = 20, \quad \boldsymbol{x} \in [-0.5, 0.5]^D$$

7) Schwefel:

$$F_7(\boldsymbol{x}) = 418.9829 \times D - \sum_{i=1}^{D} x_i \sin\left(|x_i|^{\frac{1}{2}}\right), \tag{7}$$
$$\boldsymbol{x} \in [-500, 500]^D$$

IV. Construction of Benchmark Problems

From the previous studies on evolutionary multitasking [1], [5], [6], the degree of intersection of the global optima, and the correspondence in the fitness landscape, are two important ingredients that lead to the complementarity between different optimization tasks. To quantify the overall inter-task synergy, the Spearman's rank correlation is proposed herein (as described next).

A. Spearman's rank correlation coefficient

Consider a large number n of well distributed solutions in a unified representation space. Let the ith solution $\boldsymbol{y}_i$ decode as $\boldsymbol{x}_{i1}$ and $\boldsymbol{x}_{i2}$ in the phenotype space of T_1 and T_2, respectively, in a multitasking environment. Further, let $r(\boldsymbol{x}_{i1})$ and $r(\boldsymbol{x}_{i2})$ be the *factorial rank* of the ith solution with respect to the two tasks. Then, the Spearman's rank (or ordinal) correlation coefficient [14] can be computed as the Pearson correlation between $r(\boldsymbol{x}_{i1})$ and $r(\boldsymbol{x}_{i2})$, stated as follows:

$$R_s = \frac{cov(r(\boldsymbol{x}_1), r(\boldsymbol{x}_2))}{std(r(\boldsymbol{x}_1))std(r(\boldsymbol{x}_2))}$$

In our experiments, we randomly generate 1,000,000 points in the unified search space to calculate R_s, which is viewed as a representation of the synergy between tasks.

In order to demonstrate the significance of the ordinal correlation measure in evolutionary multitasking, we consider the case of two minimization tasks T_1 and T_2 with objective/cost functions f_1 and f_2 that share high ordinal correlation (i.e., high similarity). To elaborate, we make the extreme assumption that for any pair of solutions i, j in the unified search space, we have $f_1(\boldsymbol{x}_{i1}) < f_1(\boldsymbol{x}_{j1}) \Leftrightarrow f_2(\boldsymbol{x}_{i2}) < f_2(\boldsymbol{x}_{j2})$. Thus, it can easily be seen that on bundling these two tasks together in a single multitasking environment, any series of steps leading to a cost reduction for T_1 will automatically lead to a cost reduction for T_2 *for free*, and vice versa, without the need for any additional function evaluations. In other words, at least within the family of functions characterized by high ordinal correlation, multitask optimization is guaranteed to be effective and provides the scope for *free lunches* [15].

B. Similarity Tuning and Classification of Benchmark Problems

Based on the aforementioned similarity metric, it is found the similarity between test functions changes as we shift or rotate the test problems. For example, in Fig. 1, we have a 1-D Ackley function and Sphere function (as described in Section III). When we shift the Sphere function, R_s almost monotonically decreases as the distance between the global optima increases. In another aspect, function rotation could generate the alignment and misalignment of the local minimum, thereby providing another means of manipulating the similarity between test problems. To elaborate, in this report, the problem pair with $R_s < 0.2$ is considered in low similarity, the problem pair with $0.2 \leq R_s < 0.8$ is considered in medium similarity, and the problem pair with $R_s \geq 0.8$ is considered in high similarity.

We first classify the composite benchmark problems into three categories based on the degree of intersection of the global optima of the constitutive optimization tasks, i.e., complete intersection, partial intersection, and no intersection, as described in Table I. Note that the intersection of optima occurs in the unified search space, and thus may not necessarily be apparent in the original search space of the optimization tasks. Within each category of optima intersection, there are three sub-categories of inter-task synergy based on the aforestated Spearman's rank correlation similarity metric, i.e.,
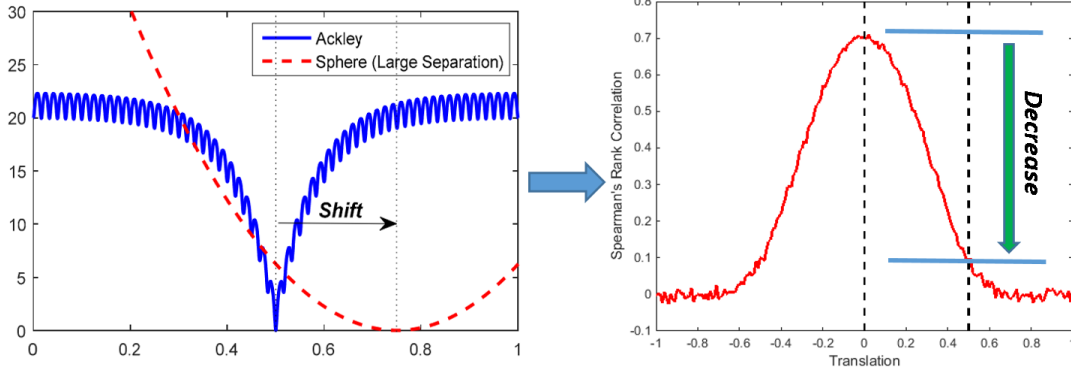
Fig. 1: The similarity decreases as we shift the Sphere function.

high, medium and low similarity. We design a single problem per category, as a result, nine problem sets are formed in total (as shall be detailed in Section V).

TABLE I: Implication of degree of intersection in the unified search space for a pair of optimization tasks

Category	Description
Complete intersection	The global optima of the two tasks are identical in the unified search space with respect to all variables
Partial intersection	The global optima of the two tasks are identical in the unified search space with respect to a subset of variables only, and different with respect to the remaining variables
No intersection	The global optima of the two tasks are different in the unified search space with respect to all variables

Note that many practical settings give rise to a third condition for categorizing potential multitask optimization settings, namely, based on the phenotypic overlap of the decision variables [6]. To elaborate, a pair of variables from distinct tasks may bear the same semantic (or contextual) meaning, which leads to the scope of knowledge transfer between them. However, due to the lack of substantial contextual meaning in the case of synthetic benchmark functions, such a condition for describing the overlap/similarity between tasks is not applied in this technical report.

V. PROBLEM SPECIFICATIONS AND BASELINE RESULTS

In this section, we present details of the nine MFO benchmark problems. Baseline results with MFEA and SOEA are also provided (implementations of both solvers are made available as supplementary material). In the current experimental setup, we employ a population of 100 individuals in the MFEA and SOEA. The total number of function evaluations for a composite problem of 2 distinct optimization tasks is restricted to 100,000 (no separate local search steps are performed). Note that "a function evaluation" here means a calculation of the objective function of a particular task T_i, and the function evaluations on different tasks are not distinguished. In MFEA, the random mating probability (*rmp*) is set to 0.3 [1]. The simulated binary crossover (SBX) [16] and polynomial mutation operators are employed for offspring creation for both
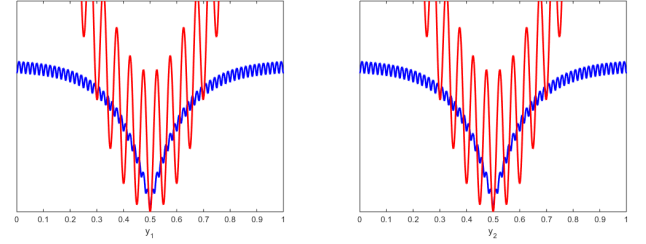

Fig. 2: In the complete intersection case, all corresponding optimum variables of the two tasks have identical values in the unified search space (see $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$)

MFEA and SOEA. Further note that no uniform crossover-like random variable swap between offspring is applied in our experiments (thereby reducing schema disruption), so as to clearly bring out the unadulterated effects of implicit genetic transfer during multitasking. All results presented hereafter are averages of 20 runs of the two solvers. All parameter settings are kept identical for both solvers to ensure fair comparison.

In what follows, $\boldsymbol{M}$ is used to represent rotation matrix (details provided below and in supplementary material). Further, $\boldsymbol{o}$ represents the location of the global optimum of an optimization task in the original (not unified) search space (details below). Table II summarizes the properties of the test problems.

A. Benchmark Problems with Complete Intersection of Global Optima

For elaboration, the notion of complete intersection of global optima is illustrated in Fig. 2 when $D = 2$. Each figure represents a projection onto the first and the second axis in the unified space, and the global optima of the two tasks are completely intersecting in the unified search space.

1) Complete Intersection and High Similarity (CI+HS): Task 1 is 50D rotated Griewank, and Task 2 is 50D rotated Rastrigin, which are defined as follows. The global optima of Task 1 and Task 2 are $\boldsymbol{o}_{ch1} = \boldsymbol{o}_{ch2} = (0, 0, \dots, 0) \in \mathbb{R}^{50}$,

TABLE II: Summary of properties of Problem Pairs for Evolutionary Multitasking

Category	Task	Landscape	Degree of intersection	Inter-task similarity R_s
CI+HS	Griewank (T_1)	multimodal, nonseparable	Complete intersection	1.0000
	Rastrigin (T_2)	multimodal, nonseparable		
CI+MS	Ackley (T_1)	multimodal, nonseparable	Complete intersection	0.2261
	Rastrigin (T_2)	multimodal, nonseparable		
CI+LS	Ackley (T_1)	multimodal, nonseparable	Complete intersection	0.0002
	Schwefel (T_2)	multimodal, separable		
PI+HS	Rastrigin (T_1)	multimodal, nonseparable	Partial intersection	0.8670
	Sphere (T_2)	unimodal, separable		
PI+MS	Ackley (T_1)	multimodal, nonseparable	Partial intersection	0.2154
	Rosenbrock (T_2)	multimodal, nonseparable		
PI+LS	Ackley (T_1)	multimodal, nonseparable	Partial intersection	0.0725
	Weierstrass (T_2)	multimodal, nonseparable		
NI+HS	Rosenbrock (T_1)	multimodal, nonseparable	No intersection	0.9434
	Rastrigin (T_2)	multimodal, nonseparable		
NI+MS	Griewank (T_1)	multimodal, nonseparable	No intersection	0.3669
	Weierstrass (T_2)	multimodal, nonseparable		
NI+LS	Rastrigin (T_1)	multimodal, nonseparable	No intersection	0.0016
	Schwefel (T_2)	multimodal, separable		

and the similarity between them is 1.0000.

$$Task\ 1:\ \min\ f_1(x) = 1 + \frac{1}{4000}\sum_{i=1}^{D} z_i^2 - \prod_{i=1}^{D}\cos\left(\frac{z_i}{\sqrt{i}}\right),$$
$$z = M_{ch1}x,\ D = 50,\ x \in [-100, 100]^D$$

$$Task\ 2:\ \min\ f_2(x) = \sum_{i=1}^{D}\left(z_i^2 - 10\cos(2\pi z_i) + 10\right),$$
$$z = M_{ch2}x,\ D = 50,\ x \in [-50, 50]^D \tag{8}$$

2) Complete Intersection and Medium Similarity (CI+MS): Task 1 is 50D rotated Ackley, and Task 2 is 50D rotated Rastrigin, which are defined as follows. The global optima of Task 1 and Task 2 are $o_{cm1} = o_{cm2} = (0, 0, \ldots, 0) \in \mathbb{R}^{50}$. And the similarity between the two tasks is 0.2261.

$$Task\ 1:\ \min\ f_1(x) = -20\exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D}z_i^2}\right)$$
$$-\exp\left(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)\right) + 20 + e,$$
$$z = M_{cm1}x,\ D = 50,\ x \in [-50, 50]^D$$

$$Task\ 2:\ \min\ f_2(x) = \sum_{i=1}^{D}\left(z_i^2 - 10\cos(2\pi z_i) + 10\right),$$
$$z = M_{cm2}x,\ D = 50,\ x \in [-50, 50]^D \tag{9}$$

3) Complete Intersection and Low Similarity (CI+LS): Task 1 is 50D rotated Ackley, and Task 2 is 50D Schwefel, which are defined as follows. The global optimal of Task 1 (search space [-50, 50]) is located at $o_{cl1} = (42.0969, \ldots, 42.0969) \in \mathbb{R}^{50}$ and the global optimal of Task 2 (search space [-500, 500]) is $o_{cl2} = (420.9687, \ldots, 420.9687) \in \mathbb{R}^{50}$, so that the global optima of the two tasks are in fact completely intersected in the
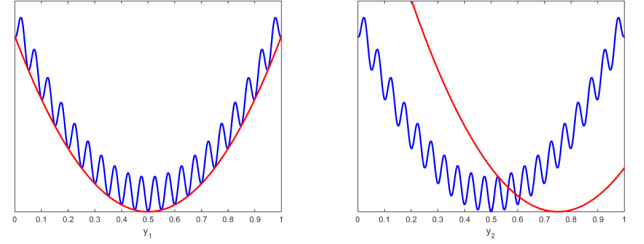

Fig. 3: In the partial intersection case, some of the corresponding optimum variables (such as y_1) for the two tasks have identical values in the unified space. Other optimum variables (such as y_2) may have distinct values in the unified space.

unified search space (based on the random-key representation scheme). The similarity between the two tasks is 0.0002.

$$Task\ 1:\ \min\ f_1(x) = -20\exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D}z_i^2}\right)$$
$$-\exp\left(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)\right) + 20 + e,$$
$$z = M_{cl1}(x - o_{cl1}),\ D = 50,\ x \in [-50, 50]^D$$

$$Task\ 2:\ \min\ f_2(x) = 418.9829 \times D - \sum_{i=1}^{D}x_i\sin\left(|x_i|^{\frac{1}{2}}\right),$$
$$D = 50,\ x \in [-500, 500]^D \tag{10}$$

B. Benchmark Problems with Partial Intersection of Global Optima

The notion of partial intersection of global optima is illustrated in Fig. 3 when $D = 2$.

1) Partial Intersection and High Similarity (PI+HS): Task 1 is 50D rotated Rastrigin, and Task 2 is 50D shifted Sphere, which are defined as follows. The global optima of Task 1 and Task 2 are located at $o_{ph1} = (0, 0, \ldots, 0) \in \mathbb{R}^{50}$ and $o_{ph2} = (0, \ldots, 0, 20, \ldots, 20) \in \mathbb{R}^{50}$ (first twenty five variables are 0s,

and the rest are 20s), respectively. The similarity between the two tasks is 0.8670.

$$Task\ 1:\quad \min\ f_1(x) = \sum_{i=1}^{D}\left(z_i^2 - 10\cos(2\pi z_i) + 10\right),$$
$$z = M_{ph1}x,\ D = 50,\ x \in [-50, 50]^D$$
$$Task\ 2:\quad \min\ f_2(x) = \sum_{i=1}^{D} z_i^2,$$
$$z = x - o_{ph2},\ D = 50,\ x \in [-100, 100]^D \tag{11}$$

2) Partial Intersection and Medium Similarity (PI+MS):
Task 1 is 50D rotated and shifted Ackley, and Task 2 is 50D Rosenbrock, which are defined as follows. The global optima of Task 1 and Task 2 are located at $o_{pm1} = (0, 0, \ldots, 0, 1, 1, \ldots, 1) \in \mathbb{R}^{50}$ (first twenty five variables are 0s, and the rest are 1s) and $o_{pm2} = (1, 1, \ldots, 1) \in \mathbb{R}^{50}$, respectively. The similarity between the two tasks is 0.2154.

$$Task\ 1:\quad \min\ f_1(x) = -20\exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} z_i^2}\right)$$
$$-\exp\left(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)\right) + 20 + e,$$
$$z = M_{pm1}x,\ D = 50,\ x \in [-50, 50]^D$$
$$Task\ 2:\quad \min\ f_2(x) = \sum_{i=1}^{D-1}(100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2),$$
$$D = 50, x \in [-50, 50]^D \tag{12}$$

3) Partial Intersection and Low Similarity (PI+LS): Task 1 is 50D rotated Ackley, and Task 2 is 25D rotated Weierstrass, which are defined as follows. The global optima of Task 1 and Task 2 are located at $o_{pl1} = (0, 0, \ldots, 0) \in \mathbb{R}^{50}$ and $o_{pl2} = (0, 0, \ldots, 0) \in \mathbb{R}^{25}$, respectively. The similarity between the two tasks is 0.0725.

$$Task\ 1:\quad \min\ f_1(x) = -20\exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} z_i^2}\right)$$
$$-\exp\left(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)\right) + 20 + e,$$
$$z = M_{pl1}x,\ D = 50,\ x \in [-50, 50]^D$$
$$Task\ 2:\quad \min\ f_2(x) = \sum_{i=1}^{D}\left(\sum_{k=0}^{k_{\max}}\left[a^k\cos\left(2\pi b^k(z_i + 0.5)\right)\right]\right)$$
$$-D\sum_{k=0}^{k_{\max}}\left[a^k\cos(2\pi b^k \cdot 0.5)\right]$$
$$a = 0.5, b = 3, k_{\max} = 20,$$
$$z = M_{pl2}x,\ D = 25,\ x \in [-0.5, 0.5]^D \tag{13}$$
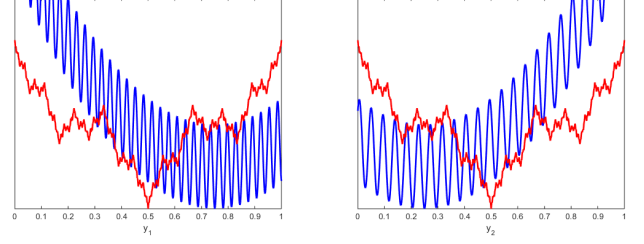
Fig. 4: In the no intersection case, the global optima of the two tasks are distinct with respect to all variables in the unified search space (see y_1 and y_2)

C. Benchmark Problems with No Intersection of Global Optima

The notion of no intersection of global optima is illustrated in Fig. 2 when $D = 2$. From the figures, we see that there is no intersection of the global optima of the two tasks in the unified search space.

1) No Intersection and High Similarity (NI+HS): Task 1 is 50D Rosenbrock, and Task 2 is 50D rotated Rastrigin, which are defined as follows. The global optimal of Task 1 and Task 2 are located at $o_{nh1} = (1, 1, \ldots, 1) \in \mathbb{R}^{50}$ and $o_{nh2} = (0, \ldots, 0) \in \mathbb{R}^{50}$, respectively. The similarity between the two tasks is 0.9434.

$$Task\ 1:\quad \min\ f_1(x) = \sum_{i=1}^{D-1}(100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2),$$
$$D = 50, x \in [-50, 50]^D$$
$$Task\ 2:\quad \min\ f_2(x) = \sum_{i=1}^{D}\left(z_i^2 - 10\cos(2\pi z_i) + 10\right),$$
$$z = M_{nh2}x,\ D = 50,\ x \in [-50, 50]^D \tag{14}$$

2) No Intersection and Medium Similarity (NI+MS): Task 1 is 50D rotated and shifted Griewank, and Task 2 is 50D rotated Weierstrass, which are defined as follows. The global optimal of Task 1 and Task 2 are located at $o_{nm1} = (10, 10, \ldots, 10) \in \mathbb{R}^{50}$ and $o_{nm2} = (0, \ldots, 0) \in \mathbb{R}^{50}$, respectively. The similarity between the two tasks is 0.3669.

$$Task\ 1:\quad \min\ f_1(x) = 1 + \frac{1}{4000}\sum_{i=1}^{D} z_i^2 - \prod_{i=1}^{D}\cos\left(\frac{z_i}{\sqrt{i}}\right),$$
$$z = M_{nm1}(x - o_{nm1}),\ D = 50,\ x \in [-100, 100]^D$$
$$Task\ 2:\quad \min\ f_2(x) = \sum_{i=1}^{D}\left(\sum_{k=0}^{k_{\max}}\left[a^k\cos\left(2\pi b^k(z_i + 0.5)\right)\right]\right)$$
$$-D\sum_{k=0}^{k_{\max}}\left[a^k\cos(2\pi b^k \cdot 0.5)\right]$$
$$a = 0.5, b = 3, k_{\max} = 20,$$
$$z = M_{nm2}x,\ D = 25,\ x \in [-0.5, 0.5]^D \tag{15}$$

TABLE III: Average performances (mean and bracketed standard deviation) of MFEA and SOEA on different tasks. Better scores are shown in bold.

Category	MFEA			SOEA		
	T_1	T_2	score	T_1	T_2	score
CI+HS	**0.3732** (0.0617)	**194.6774** (34.4953)	**-37.6773**	0.9084 (0.0585)	410.3692 (49.0439)	37.6773
CI+MS	**4.3918** (0.4481)	**227.6537** (52.2778)	**-25.2130**	5.3211 (1.2338)	440.5710 (65.0750)	25.2130
CI+LS	**20.1937** (0.0798)	**3700.2443** (429.1093)	**-25.7157**	21.1666 (0.2010)	4118.7017 (657.2786)	25.7157
PI+HS	613.7820 (131.0438)	**10.1331** (2.4734)	**-6.8453**	**445.1040** (57.2891)	83.9985 (17.1924)	6.8453
PI+MS	**3.4988** (0.6289)	**702.5026** (267.8558)	**-33.1556**	5.0665 (0.4417)	23956.6394 (10487.2597)	33.1556
PI+LS	20.0101 (0.1302)	19.3731 (1.7291)	36.1798	**5.0485** (0.6299)	**13.1894** (2.3771)	**-36.1798**
NI+HS	**1008.1740** (346.1264)	**287.7497** (92.4182)	**-33.7021**	24250.9184 (5842.0394)	447.9407 (61.1624)	33.7021
NI+MS	**0.4183** (0.0654)	**27.1470** (2.6883)	**-35.2738**	0.9080 (0.0702)	36.9601 (3.4558)	35.2738
NI+LS	650.8576 (98.6871)	**3616.0492** (325.0275)	4.2962	**437.9926** (62.6339)	4139.8903 (524.4335)	**-4.2962**

3) No Intersection and Low Similarity (NI+LS): Task 1 is 50D rotated Rastrigin, and Task 2 is 50D Schwefel, which are defined as follows. The global optimal of Task 1 and Task 2 are located at $\boldsymbol{o}_{nl1} = (0, 0, \ldots, 0) \in \mathbb{R}^{50}$ and $\boldsymbol{o}_{nl2} = (420.9687, 420.9687, \ldots, 420.9687) \in \mathbb{R}^{50}$, respectively. The similarity between the two tasks is 0.0016.

$$Task\ 1: \quad \min\ f_1(x) = \sum_{i=1}^{D} \left(z_i^2 - 10\cos(2\pi z_i) + 10\right),$$
$$\boldsymbol{z} = \boldsymbol{M}_{ph1}\boldsymbol{x},\ D = 50,\ \boldsymbol{x} \in [-50, 50]^D$$
$$Task\ 2: \quad \min\ f_2(x) = 418.9829 \times D - \sum_{i=1}^{D} x_i \sin\left(|x_i|^{\frac{1}{2}}\right),$$
$$D = 50,\ \boldsymbol{x} \in [-500, 500]^D \tag{16}$$

VI. PERFORMANCE EVALUATION AND BASELINE RESULTS

In this section, we provide a performance metric to facilitate the comparisons of algorithms by summarizing their overall performance over multiple tasks.

A. Performance Metric

In order to compare the performance of different multi- and single-tasking algorithms, a simple performance metric is proposed herein. Say we have N stochastic algorithms, $A_1, A_2, \ldots, A_N$ for a specific test case with K minimization tasks $T_1, T_2, \ldots, T_K$, and each algorithm is run for L repetitions. Suppose $I(i, j)_l$ denotes the best obtained result on the lth repetition by Algorithm A_i on the task T_j. Note that for fairness of comparison, the amount of computational effort (measured in terms of total function evaluations) spent on T_j should be the same for all algorithms (i.e. both single and multitasking). Next, let μ_j and σ_j be the mean and the standard deviation with respect to task T_j over all the repetitions of all algorithms. Thereafter, consider the normalized performance $I'(:, j)_l = (I(:, j)_l - \mu_j)/\sigma_j$. This normalization procedure is repeated for all tasks on all the repetitions.

Based on the above, for each algorithm A_i, its final performance score is given as:

$$\text{score}_i = \sum_{j=1}^{K} \sum_{l=1}^{L} I'(i, j)_l. \tag{17}$$

As is clear, a smaller score suggests that the corresponding algorithm has a superior overall performance over all tasks in the multitasking environment.

B. Baseline Results

In Table III, we present the *scores* of the MFEA and SOEA when the termination condition of 100,000 total function evaluations (for both tasks combined) is satisfied. Best scores are shown in bold. Table III demonstrates that multitasking (using MFEA) is in fact superior to single-task optimization in seven out of the nine benchmark examples.

Figs. 9-17 show the sample performances of MFEA and SOEA, where the first two panels are the convergence trends with respect to the two constitutive optimization tasks, and the third panel depicts the "score" trends of MFEA and SOEA. From these figures as well, it can be inferred that the MFEA demonstrates superior overall search performance in seven out of nine benchmark cases. Note that the two cases where MFEA is inferior, are characterized by very low similarity (Spearman's rank correlation) between the constitutive tasks, thereby highlighting the importance of inter-task synergy to the success of multitasking. The lack of any latent synergy can potentially render multitasking ineffective due to the threat of predominantly negative transfer of genetic material in the unified search space.

VII. CONCLUSIONS

This report presents a total of 9 benchmark problems, and a performance metric, which can be used for future development and testing of multitask optimization algorithms. Baseline results are also provided. As supplementary material, we make available MATLAB implementations of all the toy examples.
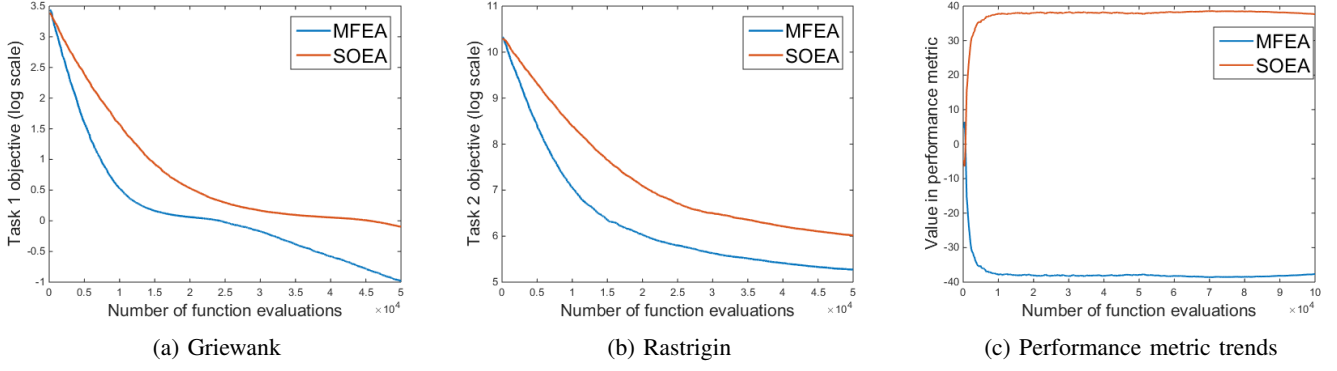
(a) Griewank   (b) Rastrigin   (c) Performance metric trends

Fig. 5: CIHS: Greiwank and Rastrigin

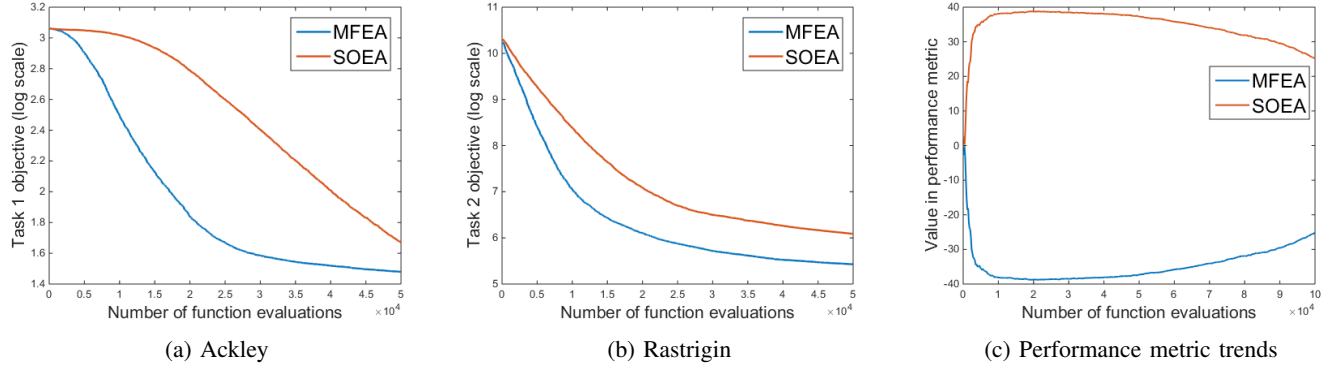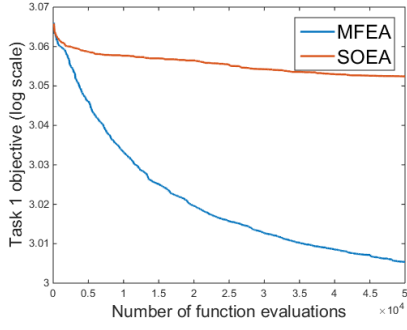(a) Ackley   (b) Rastrigin   (c) Performance metric trends
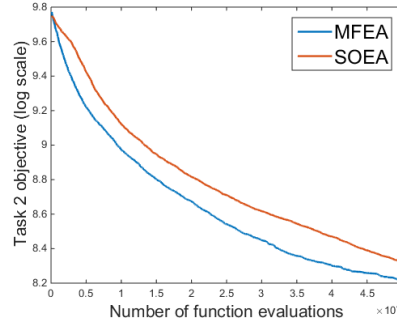
Fig. 6: CIMS: Ackley and Rastrigin

To conclude, it must be mentioned that the main aim of this work, at least in the long run, is to support the development of an *ideal* evolutionary multitasking engine which is envisioned to be a complex *adaptive* system with performance being comparable (if not consistently superior) to state-of-the-art serial evolutionary optimizers of the present day.

REFERENCES

[1] A. Gupta, Y.-S. Ong, and L. Feng, "Multifactorial evolution: toward evolutionary multitasking," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 3, pp. 343–357, 2016.

[2] A. Gupta, Y.-S. Ong, L. Feng, and K. C. Tan, "Multiobjective multifactorial optimization in evolutionary multitasking," 2016.

[3] A. Gupta and Y.-S. Ong, "Genetic transfer or population diversification? deciphering the secret ingredients of evolutionary multitask optimization," *arXiv preprint arXiv:1607.05390*, 2016.

[4] R. Caruana, "Multitask learning," in *Learning to learn*. Springer, 1998, pp. 95–133.

[5] A. Gupta, Y. Ong, B. Da, L. Feng, and S. Handoko, "Measuring complementarity between function landscapes in evolutionary multitasking."

[6] Y.-S. Ong and A. Gupta, "Evolutionary multitasking: A computer science view of cognitive multitasking," *Cognitive Computation*, vol. 8, no. 2, pp. 125–142, 2016.

[7] X. Chen, Y.-S. Ong, M.-H. Lim, and K. C. Tan, "A multi-facet survey on memetic computation," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 5, p. 591, 2011.

[8] Y.-S. Ong, M. H. Lim, and X. Chen, "Research frontier-memetic computationpast, present & future," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, p. 24, 2010.

[9] R. Dawkins *et al.*, *The selfish gene*. Oxford university press, 2016.

[10] L. L. Cavalli-Sforza and M. W. Feldman, "Cultural versus biological inheritance: phenotypic transmission from parents to children.(a theory of the effect of parental phenotypes on children's phenotypes)." *American journal of human genetics*, vol. 25, no. 6, p. 618, 1973.

[11] M. W. Feldman and K. N. Laland, "Gene-culture coevolutionary theory," *Trends in Ecology & Evolution*, vol. 11, no. 11, pp. 453–457, 1996.

[12] M. Iqbal, W. N. Browne, and M. Zhang, "Reusing building blocks of extracted knowledge to solve complex, large-scale boolean problems," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 465–480, 2014.

[13] J. C. Bean, "Genetic algorithms and random keys for sequencing and optimization," *ORSA journal on computing*, vol. 6, no. 2, pp. 154–160, 1994.

[14] J. H. Zar, "Significance testing of the spearman rank correlation coefficient," *Journal of the American Statistical Association*, vol. 67, no. 339, pp. 578–580, 1972.

[15] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE transactions on evolutionary computation*, vol. 1, no. 1, pp. 67–82, 1997.

[16] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 3, pp. 1–15, 1994.
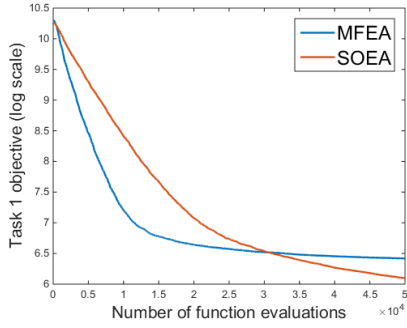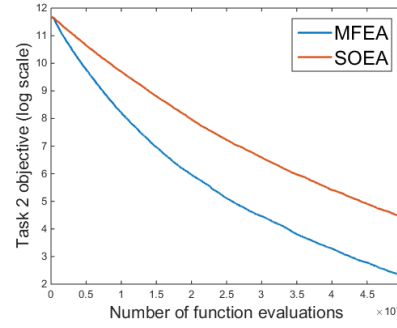
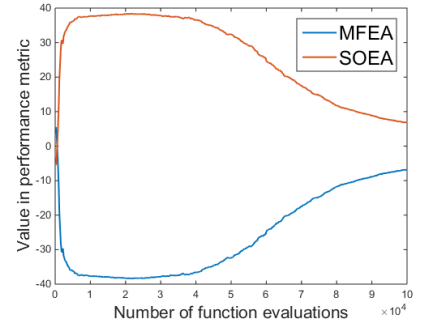(a) Ackley

(b) Schwefel
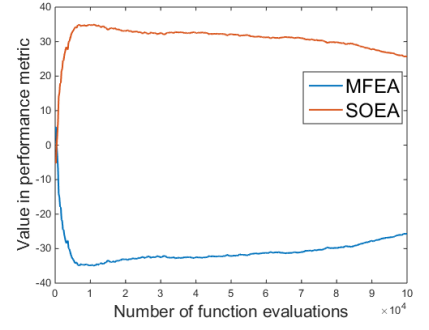
(c) Performance metric trends
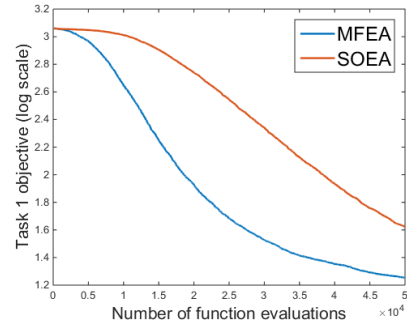
Fig. 7: CILS: Ackley and Schwefel

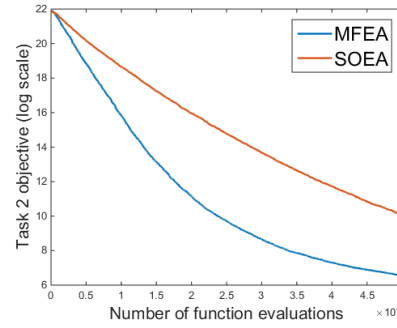

(a) Rastrigin

(b) Sphere

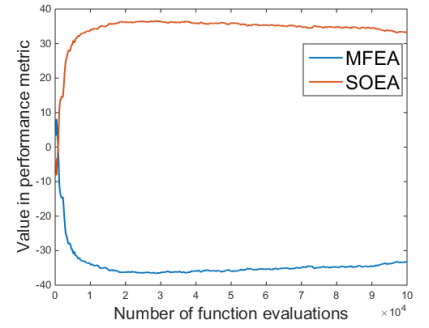(c) Performance metric trends

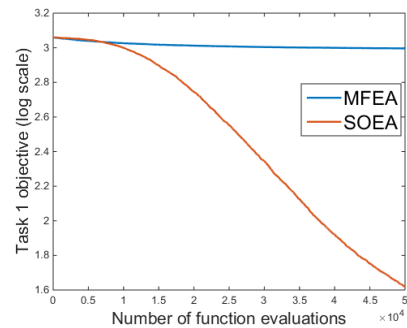Fig. 8: PIHS: Rastrigin and Sphere
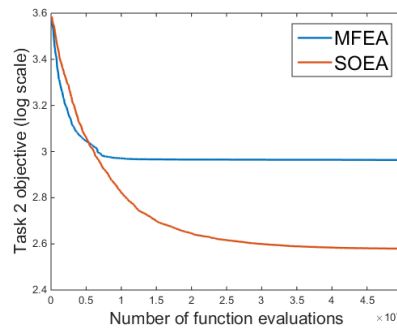

(a) Ackley

(b) Rosenbrock
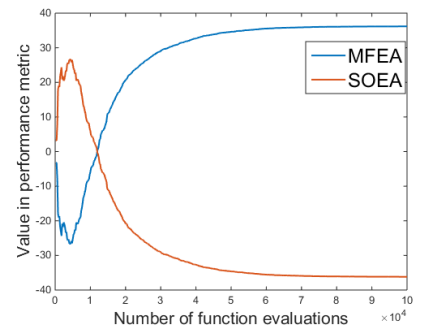
(c) Performance metric trends

Fig. 9: PIMS: Ackley and Rosenbrock
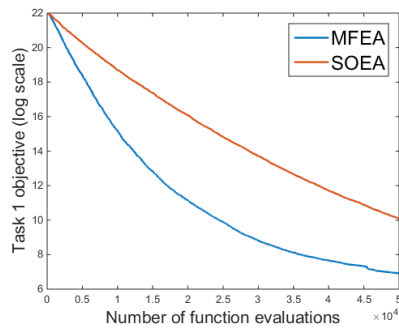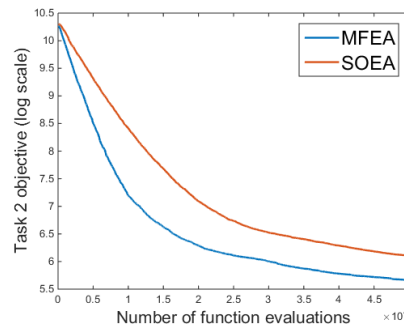

(a) Ackley

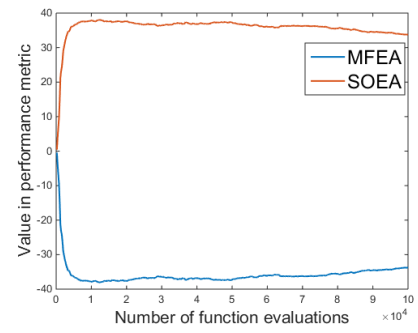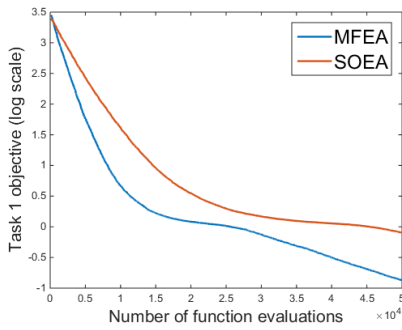(b) Weierstrass

(c) Performance metric trends

Fig. 10: PILS: Ackley and Weierstrass

(a) Rosenbrock

(b) Rastrigin

(c) Performance metric trends

Fig. 11: NIHS: Rosenbrock and Rastrigin
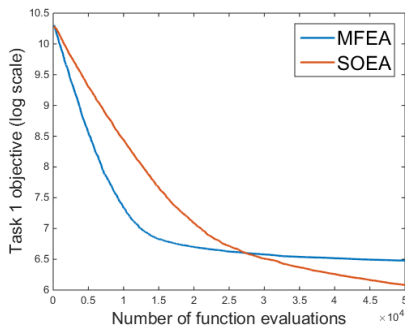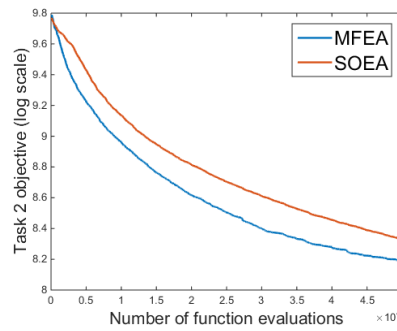

(a) Griewank
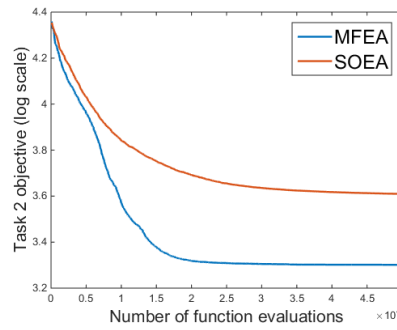
(b) Weierstrass
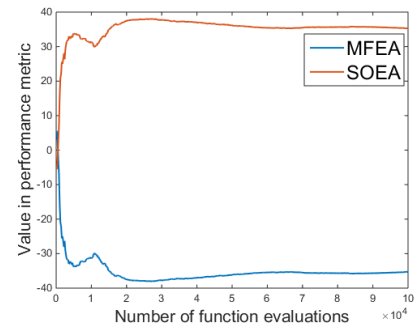
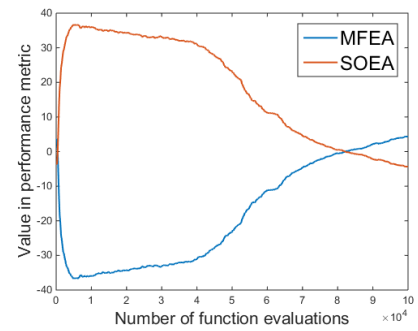(c) Performance metric trends

Fig. 12: NIMS: Griewank and Weierstrass

(a) Rastrigin

(b) Schwefel

(c) Performance metric trends

Fig. 13: NILS: Rastrigin and Schwefel