Landscape Synergy in Evolutionary Multitasking

A. Gupta[1], Y. S. Ong[1,2], B. Da[2], L. Feng[3], and S. D. Handoko[4]

[1]Rolls-Royce@NTU Corporate Lab c/o, School of Computer Engineering
Nanyang Technological University, Singapore
Email: {abhishekg, asysong}@ntu.edu.sg
[2]Computational Intelligence Lab, Nanyang Technological University, Singapore
Email: DA0002UI@e.ntu.edu.sg
[3]College of Computer Science, Chongqing University, China
Email: liangf@cqu.edu.cn
[4]School of Information Systems, Singapore Management University, Singapore
Email: dhandoko@smu.edu.sg

***Abstract* — Over the years, the algorithms of evolutionary computation have emerged as popular tools for tackling complex real-world optimization problems. A common feature among these algorithms is that they focus on efficiently solving a single problem at a time. Despite the availability of a population of individuals navigating the search space, and the implicit parallelism of their collective behavior, seldom has an effort been made to *multitask*. Considering the power of implicit parallelism, we are drawn to the idea that population-based search strategies provide an idyllic setting for leveraging the underlying synergies between objective function landscapes of seemingly distinct optimization tasks, particularly when they are solved together with a single population of evolving individuals. As has been recently demonstrated, allowing the principles of evolution to autonomously exploit the available synergies can often lead to accelerated convergence for otherwise complex optimization tasks. With the aim of providing deeper insight into the processes of evolutionary multitasking, we present in this paper a conceptualization of what, in our opinion, is one possible interpretation of the complementarity between optimization tasks. In particular, we propose a *synergy metric* that captures the correlation between objective function landscapes of distinct tasks placed in synthetic multitasking environments. In the long run, it is contended that the metric will serve as an important guide toward better understanding of evolutionary multitasking, thereby facilitating the design of improved multitasking engines.**

***Index Terms* — Evolutionary Optimization, Evolutionary Multitasking, Landscape Synergy, Memetic Computation.**

I. Introduction

EVOLUTIONARY algorithms (EAs) are bio-inspired stochastic optimization techniques [1], [2] that are characterized by a population of virtual agents or individuals. Their remarkable feature is that from a collection of simple rules mimicking Darwinian evolution emerges an implicitly parallel search engine capable of dealing with a variety of complex optimization problems. Although significant research has been carried out over several years towards the advancement of EAs, we find that the majority of these works are limited to the case of handling a single optimization problem (usually belonging to a specific domain) at a time. Despite the known power of implicit parallelism [3], seldom has an effort been made toward exploring the implications of *evolutionary multitasking*, i.e., to solve multiple optimization problems concurrently using a single population of evolving individuals [4]. It is contended that the potential for multitasking is in fact a feature exclusive to population-based search algorithms, one that undeniably sets them apart from their classical mathematical counterparts. Moreover, the benefits of appropriately harnessing this potential can be numerous. From a theoretical point of view, it may be possible to leverage upon the underlying synergies between objective function landscapes of distinct optimization tasks in an implicit manner, thereby enabling accelerated convergence towards the global optimum of multiple tasks at once. In fact, in the long run, an *ideal* evolutionary multitasking engine is envisioned to be a complex adaptive system with its performance being at least comparable to that of standard serial evolutionary optimizers of the present day.

A strong practical motivation for the evolutionary multitasking paradigm is derived from the rapidly growing field of cloud computing. In general, most cloud-based services face the natural phenomenon wherein multiple jobs are received from multiple customers at the same time. For the purpose of this study, we conceive a cloud-based on-demand service providing customers with access to state-of-the-art optimization software. Thus, it is likely for a situation to arise wherein the service provider is faced with multiple distinct optimization tasks to be addressed at once. The tasks may either have similar properties or may even belong to entirely different domains. Nevertheless, we claim that it may so happen that there exist certain complementarities, albeit unknown to the service provider, which exist between some of these tasks. The source of the complementarity could, for instance, lie in the mutually supporting objective function landscapes of the respective tasks. Therefore, in such scenarios, it becomes possible to comprehend the utility of an optimization solver that is capable of autonomously leveraging on the available synergies, i.e., without the need for any external intervention. The potential improvement in the provided solutions, either in terms of speed and/or quality, with the end user as well as the service provider being the eventual beneficiaries, highlights the real-world significance of the suggested multitasking paradigm.

Until recently, the notion of evolutionary multitasking had largely eluded researchers in the field of evolutionary

978-1-5090-0623-6/16/$31.00 ©2016 IEEE

computation (EC). In [5], the problem of simultaneously dealing with multiple optimization tasks was introduced under the label of *multifactorial optimization* (MFO), wherein each task contributes a unique *factor* influencing the evolution of a single population of individuals. In order to simulate the evolution of the population in a composite landscape, the *multifactorial evolutionary algorithm* (MFEA) proposed in [5] as a computational analogue of the bio-cultural models of *multifactorial inheritance* [6], [7], is considered herein. These models essentially describe how complex developmental traits in offspring emerge from the interactions of various genetic and cultural factors. The central ingredient of the MFEA is that it makes use of a *unified search space* encompassing all tasks in a multitasking environment. As a result, the building blocks corresponding to different tasks are contained within a unified pool of genetic material. This feature is in fact the key element enabling the discovery and implicit transfer of useful genes from one task to another in an effective manner.

While the concept of multitasking has been popular in the field of machine learning for several years [8], to the best of our knowledge, little has been done to explore an equivalent concept in the context of numerical optimization. Further, in contrast to machine learning, it is recognized that in many real-world applications of evolutionary multitasking, there is unlikely to exist any *a priori* task-specific knowledge to leverage upon. Nevertheless, it was demonstrated in [5] that the notion of *implicit genetic transfer* during evolution provides the scope for autonomously exploiting the underlying synergies between optimization tasks. For facilitating further advancements in the field of evolutionary multitasking*, it is considered vital to develop a sensible theoretical explanation of when and why implicit genetic transfer may lead to performance enhancements*. In particular, it is important to devise a measure of the inter-task complementarity which gets harnessed during the process of multitasking. To this end, we propose a *synergy metric* that captures and quantifies what, in our opinion, is a meaningful interpretation of the complementarity between any two optimization tasks in a multitasking environment. Since the magnitude of synergy is analyzed from the perspective of complementarity in the objective function landscapes, we consider synthetic problems where knowledge about the landscapes is available beforehand for the purpose of analyzing the proposed metric.

For a comprehensive exposition on the formulation of the synergy metric, this paper has been organized as follows. Section II contains an overview of the basic concepts of MFO and briefly highlights the distinction between the proposed multitasking paradigm and the field of multi-objective optimization. The MFEA is discussed in Section III, emphasizing the means of implicit genetic transfer in multitasking. Thereafter, the construction of the synergy metric is carried out in Section IV, and its implications are studied in Section V via computational experiments. Finally, a summary of the work is contained in Section VI.

II. An Overview of Multifactorial Optimization

Consider a situation wherein K distinct optimization tasks to be performed simultaneously. Without loss of generality, all tasks are assumed to be *minimization* problems. The j^{th} task, denoted T_j, has a search space $\boldsymbol{X}_j$ and an objective function $F_j: \boldsymbol{X}_j \rightarrow \mathbb{R}$. In such a setting, we define MFO as an *evolutionary multitasking paradigm* that builds on the implicit parallelism of population-based search with the aim of finding $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_{K-1}, \boldsymbol{x}_K\} = argmin\ \{F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), ..., F_{K-1}(\boldsymbol{x}), F_K(\boldsymbol{x})\}$, where $\boldsymbol{x}_j \in \boldsymbol{X}_j$. Here, each F_j is treated as an added *factor* influencing the evolution of a single population of individuals. Thus, the composite problem is also referred to as a K-factorial problem.

Since the design fundamentals of an EA are based on the Darwinian principle of natural selection, it is necessary to first quantify the "fitness" of an individual in a multitasking environment. Accordingly, we define a set of properties for every individual p_i, where $i \in \{1, 2, \ldots, |P|\}$, in a population P. Keep in mind that since the MFEA is endowed with a unified search space $\boldsymbol{Y}$ encompassing $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, …, and $\boldsymbol{X}_K$, every individual can be decoded into a task-specific solution representation with respect to each of the K tasks.

Definition 1 (*Factorial Rank*)*:* The *factorial rank* r_j^i 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 F_j.

Definition 2 (*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 tasks then $\tau_i = argmin_j\{r_j^i\}$ where $j \in \{1, 2, .., K\}$.

Definition 3 (*Scalar Fitness*)*:* The scalar fitness of p_i in a multitasking environment is given by $\varphi_i = 1/r_{\tau_i}^i$.

Once the fitness of every individual has been scalarized according to Definition 3, they can be compared in a straightforward manner. For e.g., individual p_1 is considered to dominate p_2 in multifactorial sense simply if $\varphi_1 > \varphi_2$. However, note that the scalar fitness assignment and comparison procedures are not absolute. Since the factorial rank of an individual depends on the performance of every other individual in the population, the comparison is in fact population dependent. However, the procedure guarantees that if an individual $p*$ maps to the global optimum of any task, then, $\varphi* \geq \varphi_i$ for all $i \in \{1, 2, \ldots, |P|\}$. Thus, the proposed technique is indeed compatible with the ensuing definition of multifactorial optimality.

Definition 4 (*Multifactorial Optimality*)*:* An individual $p*$, with a list of objective values $\{F_1^*, F_2^*, \ldots, F_K^*\}$, is considered *optimum* in multifactorial sense iff $\exists j \in \{1, 2, \ldots, K\}$ such that $F_j^* \leq F_j(\boldsymbol{x}_j)$, for all feasible $\boldsymbol{x}_j \in \boldsymbol{X}_j$.

A. Multifactorial vs. Multi-Objective Optimization

As multi-objective optimization and MFO are both involved with optimizing a set of objective functions, conceptual similarities may be seen to exist between them. However, we emphasize that a fundamental distinction exists between the two paradigms. While evolutionary multitasking aims to leverage the implicit parallelism of population-based search to exploit latent complementarities between distinct tasks, multi-objective optimization deals with efficiently resolving conflicts among competing objectives of the same task. An illustration summarizing the statement is depicted in Fig. 1. It is shown therein that simultaneous existence of multiple heterogeneous search spaces occurs in the case of multitasking. On the other hand, for multi-objective

optimization, there typically exists a single search space for a given task, with all objective functions depending on variables contained within that space. As a point of further interest (although not specifically studied in the present paper), note that a multitasking environment could in fact include a multi-objective optimization task as one among many other concurrent tasks. This highlights the greater generality of the proposed multitasking paradigm.

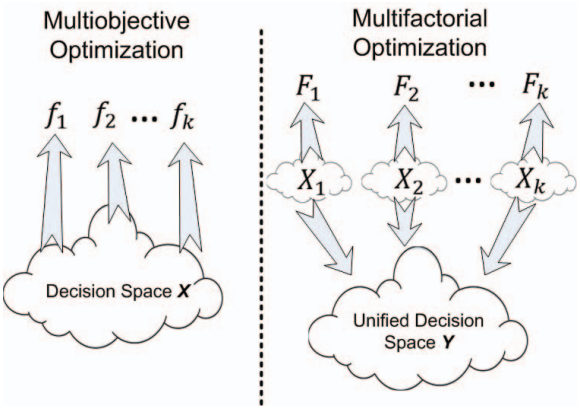

Fig. 1. Highlighting the distinction between multi-objective and multifactorial optimization. While multi-objective optimization typically comprises a single search/decision space for all objective functions, MFO combines multiple heterogeneous spaces at once.

III. The Multifactorial Evolutionary Algorithm

In this section, we turn our focus towards discussing an effective EA for the purpose of multitasking. In [5], the multifactorial evolutionary algorithm (MFEA) was shown to be inspired by bio-cultural models of multifactorial inheritance [6], [7]. The algorithm can in fact be classified under the umbrella of *memetic computation* [9], [10] as it considers the transmission of biological as well as cultural building blocks (genes and memes [11]) from parents to their offspring. In particular, cultural effects are incorporated via two aspects of multifactorial inheritance acting in concert, namely, (a) *nonrandom* or *assortative mating* [6], [7]: which states that individuals prefer to mate (crossover) with those exhibiting similar characteristics or belonging to the same cultural background, and (b) *vertical cultural transmission* [12]: which states that the phenotype of an offspring gets directly affected by the phenotype of its parents.

The basic structure of the MFEA is presented in Algorithm 1. While the majority of the algorithm resembles a standard EA, there are some distinctive features that follow the principles of multifactorial inheritance for effective evolutionary multitasking. For details, the reader is referred to [5]. The same has not been reproduced in this paper for the sake of brevity. An additional feature of the algorithm that was considered as one of the ingredients in the MFEA proposed in [5] is that every offspring undergoes a local learning step with respect to its skill factor. In this study, a *gradient based* local learning procedure is considered. This feature shall be utilized while constructing the synergy metric.

A. The Unified Search Space

For implicit genetic transfer to take place efficiently in a multitasking environment, it is essential to describe a common or unified genotype space that encompasses the heterogeneous design spaces of all the constitutive tasks. By doing so, the genetic building blocks corresponding to different tasks are unified into a single pool of genetic material, thereby enabling the MFEA to process them in parallel. Accordingly, assuming the design space dimensionality of task T_j to be D_j, we define a unified search space $\boldsymbol{Y}$ such that $D_{multitask} = max_j\{D_j\}$ for $j \in \{1, 2, \ldots, K\}$. The chromosome of an individual in the MFEA is therefore defined by a vector of $D_{multitask}$ random-keys [13] with each random-key bounded within the fixed range [0, 1]. While addressing task T_j, we simply refer to D_j random-keys contained in the chromosome.

With the above in mind, we now describe the decoding procedure of a chromosome $\boldsymbol{y} \in \boldsymbol{Y}$ into a meaningful task-specific solution representation. In the case of continuous optimization problems, this can be achieved in a straightforward manner by linearly mapping each random-key from the genotype space to the design space of the optimization task. For instance, consider a task T_j in which the i^{th} variable (x_i) is bounded in the range [L_i, U_i]. If the i^{th} random-key of the chromosome $\boldsymbol{y}$ takes a value $y_i \in [0, 1]$, then the decoding procedure is simply given by,

$$x_i = L_i + (U_i - L_i) \cdot y_i. \quad (1)$$

Algorithm 1: Pseudocode of the MFEA

1. Randomly generate n individuals in $\boldsymbol{Y}$ to form initial population P_0
2. **for every** p_i in P_0 **do**
   Evaluate p_i for all tasks in the multitasking environment
3. **end for**
4. Compute scalar fitness φ_i for every p_i
5. Set $t = 0$
6. **while** (stopping conditions are not satisfied) **do**
   C_t = Offspring (P_t) → Based on assortative mating
   for every c_i in C_t **do**
   Assign skill factor τ_i as per vertical cultural transmission
   Evaluate c_i for task τ_i only
   end
   $R_t = C_t \cup P_t$
   Update scalar fitness of all individuals in R_t.
   Select N fittest members from R_t to form P_{t+1}.
   Set $t = t + 1$
7. **end while**

IV. Constructing the Synergy Metric

In this paper, we aim to develop a deeper theoretical understanding of when and why the process of implicit genetic transfer (during multitasking) may lead to performance enhancements. Thus, of key interest is a formal description of the manner in which one task may complement another in a multitasking environment. To this end, based on a correlation analysis of the objective function landscapes, we propose a

synergy metric (ξ) for capturing and quantifying a promising mode of complementarity between distinct optimization tasks.

Let us consider a 2-factorial problem for minimizing the objective functions $F_1 : \boldsymbol{X}_1 \rightarrow \mathbb{R}$ and $F_2 : \boldsymbol{X}_2 \rightarrow \mathbb{R}$. For ease of conceptualization, it is assumed in the subsequent steps that both tasks belong to the domain of continuous optimization and possess the same search space dimensionality. For cases where the dimensionality differs, the chromosome decoding procedure permits genetic exchange to only occur between the first $D_{overlap}$ dimensions, where $D_{overlap} = min\{D_1, D_2\}$.

During construction of the synergy metric, the gradients of the constitutive functions are expected to provide clues toward estimating the movement of individuals in a population. Recall that gradient information is incorporated in the MFEA during local search refinements of offspring via some gradient-based local optimizer. Accordingly, we denote the gradient of a function F in its original search space as $\nabla_{\boldsymbol{X}} F$, and its transformation to the unified search space as $\nabla_{\boldsymbol{Y}} F$. The relation between $\nabla_{\boldsymbol{X}} F$ and $\nabla_{\boldsymbol{Y}} F$ can then be stated as,

$$\nabla_{\boldsymbol{Y}} f = \boldsymbol{J}^T \nabla_{\boldsymbol{X}} f, \tag{2}$$

where $\boldsymbol{J}$ is the Jacobian matrix. For the adopted chromosome decoding procedure described in Section III-A, $\boldsymbol{J}$ takes the following form,

$$\boldsymbol{J} = \begin{bmatrix} U_1 - L_1 & 0 & \dots \\ 0 & U_2 - L_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \tag{3}$$

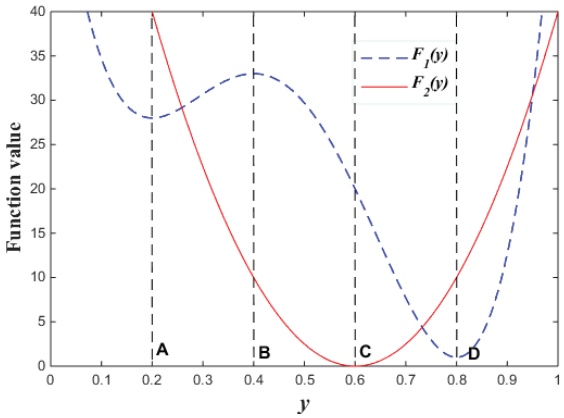

Fig. 2. Hypothetical functions F_1 and F_2 in a unified one-dimensional search space. The functions are *locally conflicting* in regions **AB** and **CD**. Yet, F_2 is globally beneficial for F_1 in **AB**.

Based on the aforementioned preliminaries, we first devise a means to determine whether one task is complementing another *at a given point* in the unified search space of a multitasking problem. In order to develop some intuition about the concept of complementarity, we begin by discussing the opposing notion of conflicting functions that frequently surfaces in the context of multi-objective optimization. Conventionally, a relationship in which the performance of one function deteriorates as the performance of another improves is described as being *conflicting* [14]. This leads to a formal definition of *locally conflicting* functions, as is summarized below.

Definition 5: Functions F_1 and F_2 are said to be *locally conflicting* at some point $\boldsymbol{y} \in \boldsymbol{Y}$ if at that point the inner product of their gradients, given by $\nabla_{\boldsymbol{Y}} F_1 \cdot \nabla_{\boldsymbol{Y}} F_2$, is negative.

Keeping Definition 5 in mind, we refer to Fig. 2, which depicts the two hypothetical functions in a unified one-dimensional search space. The global optimum of function F_1 is located at **D**, while that of function F_2 is located at **C**. Within the region **CD**, the two functions are locally conflicting according to Definition 5 as an increase in F_2 is accompanied by a decrease in F_1 (and vice versa). Similarly, in the region **AB**, the functions can once again be designated as being locally conflicting as an increase in F_1 is associated with a decrease in F_2. However, notice that although minimizing F_2 in **AB** pushes candidate solutions *away* from the *local optimum* of F_1 located at **A**, it effectively pushes the solutions *towards* the *global optimum* of F_1 at **D**. In other words, minimizing F_2 in **AB** can in fact be considered globally beneficial for F_1 despite the two functions being locally conflicting in that region. This observation is summarized by Definition 6, wherein we introduce the idea of positively complementing functions.

Definition 6: The function F_2 is said to be *positively complementing* F_1 at some point $\boldsymbol{y} \in \boldsymbol{Y}$ if at that point $-\nabla_{\boldsymbol{Y}} F_2 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y}) > 0$, where $\boldsymbol{y}_1^* \in \boldsymbol{Y}$ is the global optimum of F_1. Moreover, a local *complementarity coefficient* of F_2 towards F_1 at $\boldsymbol{y}$ is defined as,

$$G_{21}(\boldsymbol{y}) = \frac{-\nabla_{\boldsymbol{Y}} F_2 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y})}{|\nabla_{\boldsymbol{Y}} F_2| \cdot |\boldsymbol{y}_1^* - \boldsymbol{y}|}. \tag{4}$$

The complementarity coefficient G_{21} in Eq. (4) is simply the cosine of the angle between the gradient of F_2 and the vector pointing towards the global optimum of F_1. Thus, G_{21} is bounded in the range [-1, 1]. $G_{21} > 0$ at some point $\boldsymbol{y}$ implies that minimizing F_2 in the neighborhood of $\boldsymbol{y}$ pushes candidate solutions in the general direction of the global optimum ($\boldsymbol{y}_1^*$) of F_1. On the other hand, $G_{21} < 0$ implies that minimizing F_2 in the neighborhood of $\boldsymbol{y}$ causes solutions to get pushed away from the global optimum of F_1. A multitasking EA that suitably exploits the former scenario, especially in the case of complex multimodal functions, can potentially achieve the desired performance improvements.

In addition to the notions of complementarity between distinct functions, a concept that turns out to be of much importance in the construction of the synergy metric is that of *self-complementarity*. The notion of self-complementarity, as formally stated below, is defined as a measure of how the local landscape of a given function helps in progressing the search process towards its own global optimum.

Definition 7: The local landscape of a function F_1, at some point $\boldsymbol{y} \in \boldsymbol{Y}$, is said to be *positively self-complementing*, if at that point $-\nabla_{\boldsymbol{Y}} F_1 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y}) > 0$. Conversely, the local landscape is *negatively self-complementing* if $-\nabla_{\boldsymbol{Y}} F_1 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y}) < 0$. Furthermore, the local *complementarity coefficient* of F_1 towards itself is defined as,

$$G_{11}(\boldsymbol{y}) = \frac{-\nabla_{\boldsymbol{Y}} f_1 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y})}{|\nabla_{\boldsymbol{Y}} f_1| \cdot |\boldsymbol{y}_1^* - \boldsymbol{y}|}. \quad (5)$$

G_{11} is interpreted in exactly the same manner as the complementarity coefficient between distinct functions. Thus, $G_{11} > 0$ at some point $\boldsymbol{y}$ implies that performing local search for function F_1 at $\boldsymbol{y}$ drives candidate solutions towards the global optimum of the same function. On the other hand, $G_{11} < 0$ indicates that the local improvement step in fact causes candidate solutions to be pushed away from the global optimum of the function.

In order to evaluate the overall complementarity of F_2 towards F_1, we only consider the contribution of F_2 in those regions of the unified search space where the local landscape of F_1 is negatively self-complementing (as per Definition 7). The union of these regions is denoted as $\overline{\boldsymbol{Y}}_{F_1}$. The total volume ($\bar{V}_{F_1}$) of this possibly disjoint subspace ($\overline{\boldsymbol{Y}}_{F_1} \subset \boldsymbol{Y}$) is given by,

$$\bar{V}_{F_1} = \int \max\{0, sign(-G_{11})\} \cdot dV, \quad (6)$$

where dV is a differential volume element of the unified search space $\boldsymbol{Y}$. With this background, the *cumulative complementarity* of F_2 towards F_1 is represented by the synergy metric (ξ_{21}) as,

$$\xi_{21} = \frac{\int \omega(\boldsymbol{y}) \cdot d\bar{V}_{F_1}}{\bar{V}_{F_1}}. \quad (7)$$

where,

$$\omega(\boldsymbol{y}) = \begin{cases} 1 & if\ G_{11}(\boldsymbol{y}) \leq 0\ and\ G_{21}(\boldsymbol{y}) > 0 \\ -1 & if\ G_{11}(\boldsymbol{y}) \leq 0\ and\ G_{21}(\boldsymbol{y}) < G_{11}(\boldsymbol{y}) \\ 0 & otherwise \end{cases}. \quad (8)$$

Note that although we exclusively address the case of F_2 complementing F_1 in this section, the same equations hold in the converse situation wherein the complementarity of F_1 toward F_2 is to be evaluated. The only adjustment necessary is the swapping of indices 1 and 2 in Eqs. (4) - (8).

A. Some theoretical implications of the synergy metric

As is clear from Eq. (7), ξ_{21} must take a value between -1 and 1. In the extreme case, when $\xi_{21} = 1$, it is indicated that F_2 positively complements F_1 throughout $\overline{\boldsymbol{Y}}_{F_1}$, and can therefore be expected to substantially accelerate the convergence process for F_1 in a multitasking environment.

As an illustration of this claim, consider a 2-factorial problem in which F_1 is a complex multimodal function with a unique global optimum, and F_2 is any convex function. Further, assume that the search space dimensionality of both functions match, and that their global optima intersect (or, are located in close proximity of one another) in the unified search space. Since F_2 is convex, we know $-\nabla_{\boldsymbol{Y}} F_2 \cdot (\boldsymbol{y}_2^* - \boldsymbol{y}) > 0$ everywhere in $\boldsymbol{Y}$. Combining this fact with the assumption that $\boldsymbol{y}_1^* \approx \boldsymbol{y}_2^*$, it follows that $-\nabla_{\boldsymbol{Y}} F_2 \cdot (\boldsymbol{y}_1^* - \boldsymbol{y})$ must also be positive everywhere in $\boldsymbol{Y}$. Therefore, according to Eqs. (7) and (8), in such scenarios $\xi_{21} \rightarrow 1$ regardless of the complexity of F_1. Now, we imagine solving the aforementioned problem with the MFEA. Notice that since F_2 is convex, it will get optimized comparatively quickly as a result of the local optimization step. Thereafter, the refined genes will get efficiently transferred across for F_1 by the process of implicit genetic transfer. In other words, both functions are likely to be optimized within a short span of time.

While examples with $\xi_{21} > 0$ are expected to showcase a majority of positive (i.e., beneficial) genetic transfer, cases with $\xi_{21} < 0$ are expected to experience much unwanted negative transfer [15]-[17] during multitasking. However, the nice property of evolution is that when the latter occurs, the negatively transferred genes get automatically eliminated from the population (over the course of a few generations) by the process of natural selection or survival of the fittest.

V. Computational Experiments

Having devised the synergy metric, we carry out an empirical analysis of the relationship between its numerical value and the observed performance of the MFEA. Before proceeding further, notice that the synergy measure is based purely on the correlation between objective function landscapes, and is independent of the mechanisms of the multitasking engine used. Thus, it is recognized that the correspondence we are seeking can only be of a "qualitative" nature as the metric is only implicitly related to the expected performance of the MFEA. In fact, how well the MFEA actually exploits underlying synergies is entirely subject to the efficacy of its mechanisms. As of now, a higher value of synergy simply highlights the possibility of improved convergence due to increased positive genetic transfer, while a drop in value suggests inferior convergence from predominantly negative transfer. Nevertheless, the metric serves as the means to understand when and why evolutionary multitasking may lead to potential performance improvements.

The computational experiments in this section have been designed to fully reveal the implications of the synergy metric. To begin with, we consider the following set of continuous optimization benchmark functions that have been commonly in use in the literature;

a) Sphere function,

$$\sum_{i=1}^{D} z_i^2\ ;\ \boldsymbol{z} = (\boldsymbol{x} - \boldsymbol{O}_S). \quad (9)$$

b) Ackley function [18],

$$20 + e - 20\exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} z_i^2}\right) - \exp\left(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)\right);\ \boldsymbol{z} = (\boldsymbol{x} - \boldsymbol{O}_A). \quad (10)$$

c) Rastrigin function [19],

$$\sum_{i=1}^{D}(z_i^2 - 10\cos(2\pi z_i) + 10);\ \boldsymbol{z} = (\boldsymbol{x} - \boldsymbol{O}_R). \quad (11)$$

d) Weierstrass function [20],

$$\sum_{i=1}^{D}\left(\sum_{k=0}^{20}\left[0.5^k \cos\left(2\pi 3^k (z_i + 0.5)\right)\right]\right) - D\sum_{k=0}^{20}[0.5^k \cos(\pi 3^k)];\ \boldsymbol{z} = (\boldsymbol{x} - \boldsymbol{O}_W). \quad (12)$$

In Eqs. (9) – (12), $\boldsymbol{O}_S$, $\boldsymbol{O}_A$, $\boldsymbol{O}_R$, and $\boldsymbol{O}_W$ represent the (adjustable) locations of the global optimums of the sphere, Ackley, Rastrigin, and Weierstrass functions, respectively. By

adjusting these locations, one can create new multitasking instances with varying levels of inter-task synergy.

In the computational experiments that follow, we shall consider 2-factorial problems. The extent and dimensionality of the search space for each function is reported in Table I. With regard to the MFEA, standard genetic operators, namely, Simulated Binary Crossover (SBX) [21] and Gaussian mutation [22], are consistently used throughout. With the SBX operator, no additional uniform crossover-like variable swap is performed so as to emphasize on the preservation of genetic linkage/building-blocks at the cost of reduced exploration. Although such swapping has proven useful for many instances with additively separable functions, it becomes much harder to decipher the true effects of multitasking. Further, a small population of $n = 30$ individuals is evolved for a period of 100 generations prior to termination of the algorithm. Recall that the MFEA incorporates individual learning into each task evaluation call. Every individual is thus improved by a local search move. Local improvements are identified via the BFGS quasi-Newton method (which is executed for a maximum of 5 iterations per individual) and the changes are incorporated into the chromosome (genetic material) of the individual in the spirit of Lamarckian learning [23].

All convergence trends obtained by the MFEA are presented alongside standard single-objective optimization (SOO). The corresponding EA for SOO employs the same population size and identical local search (BFGS), crossover (SBX), and mutation (Gaussian) operations as the MFEA, with the probability of an individual to mutate fixed at 10%.

TABLE I

DETAILS OF THE CHOSEN TEST FUNCTIONS

Functions	Dimensionality	Search Range	Description
Sphere	30	[-100, 100]	Unimodal
Ackley	30	[-32, 32]	Multimodal
Rastrigin	20	[-5, 5]	Multimodal
Weierstrass	20	[-0.5, 0.5]	Multimodal

A. Numerical approximation of the synergy metric

In order to rapidly compute the implication of the synergy metric, the integrals in Eqs. (6) & (7) are numerically approximated over a restricted two dimensional search space. The suppression of the remaining dimensions is considered viable as the test functions are separable. 10,000 uniformly distributed points are generated in the 2-D space and standard rectangular quadrature is employed for numerical integration. Also, with regard to the (numerical) computation of function gradients, note that although an infinite series form of the Weierstrass function is non-differentiable, the partial sum in Eq. (12) is differentiable everywhere.

In the subsection that follows, we show that the numerical approximation of the metric provides a good "qualitative" understanding of the observed performance of the MFEA.

B. Empirical revelations of the synergy metric

We construct three 2-factorial test examples by selecting from the list of benchmark functions provided in Table I. Each example consists of three multitasking instances which are formed by combining functions and/or adjusting the locations of their global optimum.

1) Example 1 (Ackley with sphere function)

In the first example, we combine the Ackley function and the sphere function into a single multitasking environment. Here, Ackley serves as a base function with its optimum fixed at $\boldsymbol{O}_A = [0, 0, \ldots, 0]$. On the other hand, three variants of the sphere function are considered. In the first case, the global optimum of the sphere function intersects with that of the Ackley function. We denote the resultant multitasking instance as $(A, S)_{ZS}$, where *ZS* stands for *zero separation*. In the second case, the optimum of the sphere function is shifted to $\boldsymbol{O}_S = [10, 10, \ldots, 10]$ and the resultant instance is denoted as $(A, S)_{MS}$ (here *MS* stands for *medium separation*). Finally, in the third case, the optimum of the sphere function is adjusted to $\boldsymbol{O}_S = [50, 50, \ldots, 50]$ and the resultant instance is denoted as $(A, S)_{LS}$ (here *LS* stands for *large separation*). A 1-D visualization of the shifted function is depicted in Fig. 3.

Table II summarizes the three multitasking instances and reports the numerical values of the synergy metric. For instance $(A, S)_{ZS}$ with intersecting optima, ξ_{SA}^{ZS} equates to 1 indicating that the overlapping sphere function positively complements the Ackley function everywhere in $\overline{\boldsymbol{Y}}_{F_A}$. In fact, this result follows directly from the discussion in Section IV-A as the sphere function is convex. On the other hand, for the instances with increasing optima separation, the results in Table II show a continuous drop in the value of the synergy metric. The experimental manifestation of the loss in synergy is displayed in Fig. 4.

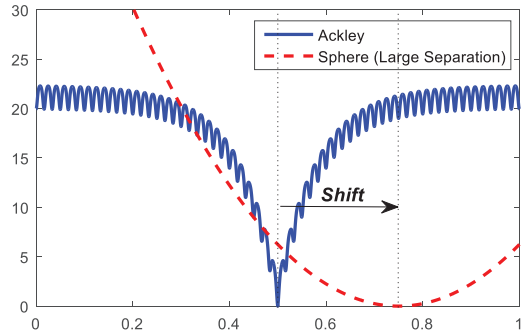

Fig. 3. A one-dimensional illustration of instance $(A, S)_{LS}$ with shifted functions mapped in the unified search space $\boldsymbol{Y}$.

TABLE II

EXAMPLE 1 (ACKLEY WITH SPHERE FUNCTION): SUMMARY OF INSTANCES AND SYNERGY METRIC VALUES. $(A, S)_{ZS}$: ZERO SEPARATION, $(A, S)_{MS}$: MEDIUM SEPARATION, $(A, S)_{LS}$: LARGE SEPARATION.

	$(A, S)_{ZS}$	$(A, S)_{MS}$	$(A, S)_{LS}$
$\boldsymbol{O}_A$	[0, 0, …, 0]	[0, 0, …, 0]	[0, 0, …, 0]
$\boldsymbol{O}_S$	[0, 0, …, 0]	[10, 10, …, 10]	[50, 50, …, 50]
ξ_{SA}	1	0.9978	0.8734

ξ_{SA} = Complementarity of the sphere function toward Ackley.

Fig. 4 presents averaged convergence trends of the Ackley function for instances (A, S)$_{ZS}$, (A, S)$_{MS}$, and (A, S)$_{LS}$. The correspondence between the numerical value of the synergy metric and the observed performance of the MFEA is revealed in this figure. As was expected, in instance (A, S)$_{ZS}$, the Ackley function gets optimized rapidly as it is positively complemented by the sphere function throughout the space $\overline{\boldsymbol{Y}}_{F_A}$. Moreover, the results in Table II indicate $\xi_{SA}^{MS} > \xi_{SA}^{LS}$, which substantiates the observation that the convergence rate of (A, S)$_{MS}$ is significantly faster than that of (A, S)$_{LS}$.

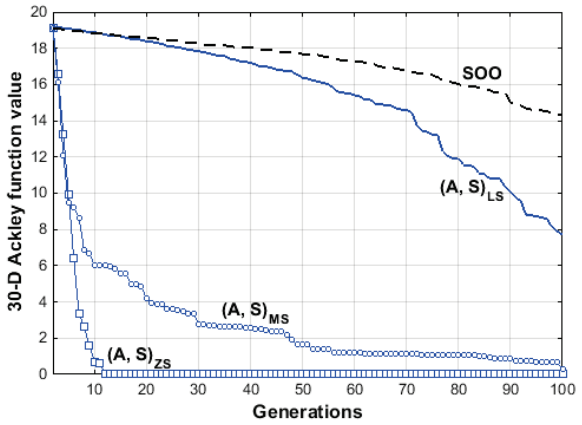

Fig. 4. Averaged convergence trends of the 30-D Ackley function in Example 1. SOO indicates the averaged performance of standard single-objective optimization.

2) Example 2 (Weierstrass with Rastrigin function)

The second 2-factorial example comprises the Weierstrass function (as the base function) and the Rastrigin function. In these instances, the global optimum of the Weierstrass function is kept fixed at $\boldsymbol{O}_W$ = [0, 0, …, 0] while the Rastrigin function occurs in the following three variants: (a) (W, R)$_{ZS}$ with $\boldsymbol{O}_R$ = [0, 0, …, 0], (b) (W, R)$_{MS}$ with $\boldsymbol{O}_R$ = [1, 1, …, 1], and finally (c) (W, R)$_{LS}$ with $\boldsymbol{O}_R$ = [2, 2, …, 2].

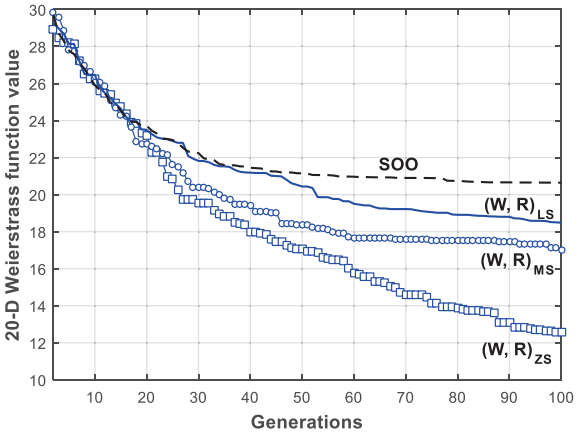

Fig. 5. Averaged convergence trends of 20-D Weierstrass function in Example 2. SOO indicates the averaged performance of standard single-objective optimization.

Table III summarizes the three instances in example 2 and reports the numerical values of the synergy metric. Convergence trends for the same are depicted in Fig. 5. As can be seen therein, the qualitative implications of the result $\xi_{RW}^{ZS} > \xi_{RW}^{MS} > \xi_{RW}^{LS}$ (in Table III) are borne out by the computational experiments. As is predicted by the synergy values, the instance (W, R)$_{ZS}$ is found to be most benefitted from the availability of positively transferrable genetic material. Furthermore, the subsequent loss in inter-task synergy successfully explains the decelerating convergence rate of the instances with increasing optima separation.

TABLE III

EXAMPLE 2 (WEIERSTRASS WITH RASTRIGIN FUNCTION): SUMMARY OF INSTANCES AND SYNERGY METRIC VALUES. (W, R)$_{ZS}$: ZERO SEPARATION, (W, R)$_{MS}$: MEDIUM SEPARATION, (W, R)$_{LS}$: LARGE SEPARATION.

	(W, R)$_{ZS}$	**(W, R)$_{MS}$**	**(W, R)$_{LS}$**
$\boldsymbol{O}_W$	[0, 0, …, 0]	[0, 0, …, 0]	[0, 0, …, 0]
$\boldsymbol{O}_R$	[0, 0, …, 0]	[1, 1, …, 1]	[2, 2, …, 2]
ξ_{RW}	0.3122	0.3057	0.2980

ξ_{RW} = Complementarity of the Rastrigin function toward Weierstrass.

3) Example 3

In the final example, the Rastrigin function is viewed as a base function with global optimum fixed at $\boldsymbol{O}_R$ = [0, 0, ….,0]. It is combined in turn with the sphere, Ackley, and Weierstrass functions. In the first instance, denoted as (R, S)$_{MS}$, the global optimum of the sphere function is shifted to [10, 10, …, 10]. The second and third instances, labeled as (R, A)$_{ZS}$ and (R, W)$_{ZS}$, comprise the Ackley and Weierstrass functions, respectively, with their global optimums intersecting with the Rastrigin function at [0, 0, …, 0]. The synergy metric values for all three instances are listed in Table IV in decreasing order of inter-task synergy. It is interesting to note that the synergy value turns out to be maximum for the sphere function (by a significant margin) despite the fact that its optimum is substantially separated from that of the Rastrigin fuction.

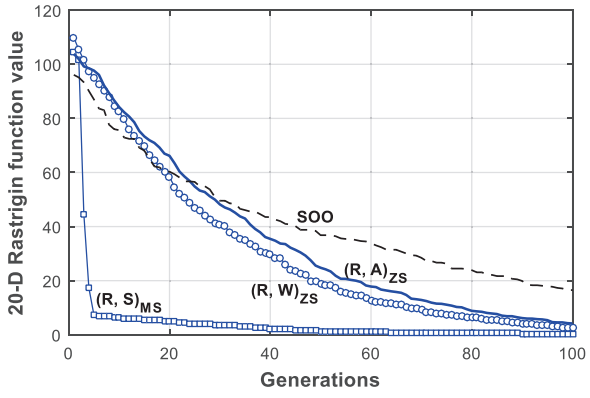

Fig. 6. Averaged convergence trends of 20-D Rastrigin function in Example 3. SOO indicates the averaged performance of standard single-objective optimization

With regard to the averaged convergence trends depicted in Fig. 6, the numerical values of the synergy metric once again prove to be a good indicator of the eventual performance of the MFEA. In particular, the correspondence between the result $\xi_{SR} > \xi_{WR} > \xi_{AR}$ (in Table IV) and the outcome of the MFEA are verified by the observed convergence trends.

TABLE IV

EXAMPLE 3: SYNERGY METRIC VALUES.

ξ_{SR}	ξ_{WR}	ξ_{AR}
0.9995	0.4868	0.2927

(ξ_{SR}, ξ_{WR}, ξ_{AR}) = Complementarity of the (sphere, Weierstass, Ackley) functions toward Rastrigin.

VI. CONCLUSIONS

In this paper, we have described the concept of evolutionary multitasking as a novel means to fully unlocking the power of implicit parallelism of population-based search. In order to better understand the cause of inter-task complementarity in multitasking environments, we have conceptualized a synergy metric (ξ) that analyzes the correlation between objective function landscapes of distinct tasks. As is substantiated by various computational experiments, the metric explains (at least in a qualitative sense) when and why the notion of implicit genetic transfer in evolutionary multitasking may allow for accelerated convergence of complex optimization tasks.

In conclusion, the present work provides an interesting perspective towards the numerical representation of the underlying complementarity between tasks, that which potentially gets harnessed by the process of implicit genetic transfer during evolutionary multitasking. It is hoped that armed with this understanding, researchers will be in a better position to design multitasking engines of the future that are capable of exploiting synergies to the fullest.

ACKNOWLEDGEMENT

This work was conducted in the Rolls-Royce@NTU Corporate Lab with support from the National Research Foundation (NRF) Singapore under the Corp Lab@University Scheme.

REFERENCES

[1] T. Back, U. Hammel, and H. P. Schwefel, "Evolutionary computation: Comments on the history and current state," *IEEE Trans. Evo. Comp.*, vol. 1, no. 1, pp. 3-17, 1997.

[2] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*, Addison Wesley, 1989.

[3] A. H. Wright, M. D. Vose, and J. E. Rowe, "Implicit Parallelism," *Lecture Notes in Comp. Sci*., vol. 2724, pp. 1505-1517, 2003.

[4] K. Krawiec, and B. Wieloch, "Automatic generation and exploitation of related problems in genetic programming." In *Evolutionary Computation (CEC), 2010 IEEE Congress on*, pp. 1-8. IEEE, 2010.

[5] A. Gupta, Y. S. Ong, and L. Feng. "Multifactorial Evolution: Toward Evolutionary Multitasking," *IEEE Trans. Evo. Comp*, Accepted, 2015.

[6] J. Rice, C. R. Cloninger, and T. Reich, "Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of the unitary models," *Am. J. Hum. Genet*., vol. 30, pp. 618-643, 1978.

[7] C. R. Cloninger, J. Rice, and T. Reich, "Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance," *Am. J. Hum. Genet*., vol. 31, 1979.

[8] R. Caruana, "Multitask Learning," *Machine Learning*, vol. 28, no. 1, pp. 41-75, 1997.

[9] X. Chen, Y. S. Ong, M. H. Lim, and K. C. Tan, "A multi-facet survey on memetic computation," *IEEE Trans. Evo. Comp*., vol. 15, no. 5, pp. 591-606, Oct. 2011.

[10] Y. S. Ong, M. H. Lim, X. S. Chen, "Research Frontier: Memetic Computation – Past, Present & Future," *IEEE Comp. Intel. Mag*., vol. 5, no. 2, pp. 24-36, 2010.

[11] R. Dawkins, *The Selfish Gene*. Oxford, U.K.: Oxford University Press, 1976.

[12] L. L. Cavalli-Sforza and M. W. Feldman, "Cultural vs Biological inheritance: Phenotypic transmission from parents to children (A theory of the effect of parental phenotypes on children's phenotypes)," *Am. J. Hum. Genet*., vol. 25, 618-637, 1973.

[13] J. C. Bean, "Genetic algorithms and random keys for sequencing and optimization," *ORSA J. Comp*., vol. 6, no. 2, 1994.

[14] R. C. Purshouse, P. J. Fleming, "Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimization," *Evolutionary Multi-Criterion Optimization*, 2003.

[15] L. Feng, Y. S. Ong, M. H. Lim, and I. W. Tsang, "Memetic Search with Inter-Domain Learning: A Realization between CVRP and CARP," *IEEE Trans. Evo. Comp.*, Accepted, 2014.

[16] L. Feng, Y. S. Ong, A. H. Tan, and I. W. Tsang, "Memes as building blocks: A case study on evolutionary optimization + transfer learning for routing problems." *Memetic Computing*, vol. 7, no. 3, pp. 159-180, 2015.

[17] S. J. Pan, Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowledge & Data Engg*., vol. 22, no. 10, pp. 1345-1359, 2009.

[18] D. H. Ackley, *A connectionist machine for genetic hillclimbing*. Boston: Kluwer Academic Publishers, 1987.

[19] L. A. Rastrigin, *Systems of extreme control*. 1974.

[20] G. H. Hardy, "Weierstrass's non-differential function," *Trans. of the American Math. Society*, vol.17, no. 3, pp. 301-325, 1916.

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

[22] R. Hinterding, "Gaussian mutation and self-adaption for numeric genetic algorithms," *IEEE CEC*, 1995, vol. 1.

[23] Y. S. Ong and A. J. Keane, "Meta-Lamarckian learning in memetic algorithms," *IEEE Trans. Evo. Comp*., vol. 8, no. 2, pp. 99-110, April 2004.