

Covariance Local Search for Memetic Frameworks: A Fitness Landscape Analysis Approach

Ferrante Neri

*COL Laboratory, School of Computer Science
University of Nottingham
Nottingham, United Kingdom
ferrante.neri@nottingham.ac.uk*

Yuyang Zhou

*School of Computer Science
University of Nottingham
Nottingham, United Kingdom
psyyz14@nottingham.ac.uk*

Abstract—The design of each agent composing a Memetic Algorithm (MA) is a delicate task which often requires prior knowledge of the problem to be effective. This paper proposes a method to analyse one feature of the fitness landscape, that is the epistasis, with the aim of designing efficient local search algorithms for Memetic Frameworks. The proposed Analysis of Epistasis performs a sampling of points within the basin of attraction and builds a data set containing those candidate solutions whose objective function value falls below a threshold.

The covariance matrix associated with this data set is then calculated. The eigenvectors of this covariance matrix are then computed and used as the reference system for the local search: a change of variables is performed and then the local search is performed on the new variables. The Analysis of Epistasis has been implemented on the three local search algorithms composing a popular MA called Multiple Trajectory Search (MTS). Numerical results show that the three modified local search algorithms outperform their original counterparts.

Index Terms—Memetic Algorithms, Fitness Landscape Analysis, Epistasis, Covariance Matrix, Local Search

I. INTRODUCTION

Local search is a fundamental element of Memetic Computing. Within Memetic Algorithms (MAs), properly designed local search algorithms may have a major impact on the performance of the entire Memetic Framework [1].

Within the context of the algorithmic design, studies on Fitness Landscape Analysis (FLA) [2], [3] showed that testing and understanding the optimisation problem enable the design of powerful problem specific heuristics [4]. For example, the Fitness Cloud allows to estimate the potential of the search [5] while the negative slope [6] and hardness measure [7] estimates the difficulty of the problem. Some other examples of fundamental features of an optimisation problem are modality, epistasis, ruggedness and deceptiveness [8]. Although most of the work on FLA currently present in the literature tends to focus on the combinatorial domain, recent studies showed that FLA has great potential also on continuous problems, see [4], [8].

In Memetic Computing, FLA provides precious pieces of information to design global and local search algorithms employed within its framework [9]–[11]. For example, the balance between global and local search should be related to the modality of the problem [12]. The most studied feature in the continuous domain is perhaps the ruggedness, that is

the number and location of local optima, see [13], [14]. A popular way to analyse the ruggedness of a problem consists of building up a random walk which, while moving according to a pattern, samples and saves the points of the landscape, see e.g. [8].

The present paper focuses on local search design based on the analysis of epistasis. Furthermore, this study refers to optimisation in the continuous domain. The concept of epistasis is borrowed from biology where it refers to the dependency between genes and phenotype. In Optimisation and more specifically in Genetic Algorithms (GAs) epistasis has been reinterpreted by means of an analogy: epistasis refers to the degree of dependency between genes in a chromosome and its objective function/fitness value [15].

For optimisation problems in general (when we do not necessarily have the GA metaphor of chromosome), epistasis is the degree of interdependency between variables with respect to the objective function value. In other words, a multi-variate optimisation problem is characterised by low epistasis if it can be decomposed into multiple independent problems with lower dimensionality while it is characterised by high epistasis if this decomposition is not possible. This property is also referred to as operational/non-linear separability and is broadly used to address large scale optimisation problems, see [16]–[18].

This paper proposes a method for handling highly epistatic local problems by means of FLA. The proposed FLA at first generates a data set of candidate solutions whose fitness value is below (minimisation) a certain threshold. This data set describes the geometry of the basin of attraction [19]. The covariance matrix of the points composing this set is calculated and its eigenvectors extracted. A linear transformation is performed: the directions of the eigenvectors are used as the new reference system. The local search is then designed to move alongside the new system of coordinates. In this paper, we demonstrate and test the proposed approach on the three local search algorithms composing an MA called Multiple Trajectory Search (MTS) proposed in [20].

The remainder of this paper is organised in the following way. Section II describes the proposed analysis of epistasis and consequent change of reference system. Section III shows how the analysis of epistasis is used to modify the local search algorithms of MTS. Section IV displays the numerical results

of this study and demonstrates that local search designed after the analysis outperforms the original local search for the three algorithms composing MTS. Section V provides the conclusion of this study.

II. ANALYSIS OF EPISTASIS AND NEW REFERENCE SYSTEM

In order to clarify the notation used in this article, we will refer to the minimisation problem of an objective function $f(\mathbf{x})$ in the continuous domain. The candidate solution \mathbf{x} is a vector of n design variables in a hyper-cubical decision space $D \subset \mathbb{R}^n$:

$$\mathbf{x} = (x_1, x_2, \dots, x_n).$$

Each vector $\mathbf{x} \in D$ is expressed in the orthonormal reference system of its variables. In other words, any vector \mathbf{x} can be interpreted as the linear combination of the orthonormal basis $B_e = \{\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^n\}$ where \mathbf{e}^k is a vector of length n composed of all zeros and a 1 in the k^{th} design variable, see [21].

Let us now consider that, at some point of the functioning of an MA, a basin of attraction has been detected and a local search must be activated, see [19], [22], [23].

The proposed analysis of epistasis consists of the following steps. A number of candidate solutions/points are sampled in the decision space D and their objective function values are calculated. The function values are compared with a threshold thr and those values that are below thr are saved in a data structure, while the others are discarded. The purpose of this step is to have a sample of points whose distribution describes the geometry of the problem.

The m vectors/candidate solutions whose objective function value is below the threshold thr

$$\begin{aligned} \mathbf{x}^1 &= (x_1^1, x_2^1, \dots, x_n^1) \\ \mathbf{x}^2 &= (x_1^2, x_2^2, \dots, x_n^2) \\ &\dots \\ \mathbf{x}^m &= (x_1^m, x_2^m, \dots, x_n^m) \end{aligned}$$

are allocated in the data structure \mathbf{V} .

These points can be interpreted as a multivariate statistical distribution characterised by a mean vector:

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n) = \frac{1}{m} \left(\sum_{i=1}^m x_1^i, \sum_{i=1}^m x_2^i, \dots, \sum_{i=1}^m x_n^i \right)$$

and a covariance matrix:

$$\mathbf{C} = \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,n} \\ c_{2,1} & c_{2,2} & \dots & c_{2,n} \\ \dots & \dots & \dots & \dots \\ c_{n,1} & c_{n,2} & \dots & c_{n,n} \end{pmatrix}$$

where

$$c_{j,j} = \frac{1}{m} \sum_{i=1}^m ((x_j^i - \mu_j)(x_j^i - \mu_j))$$

and

$$c_{j,l} = \frac{1}{m} \sum_{i=1}^m ((x_j^i - \mu_j)(x_l^i - \mu_l)).$$

Subsequently, the n eigenvectors of the matrix \mathbf{C} are calculated by means of Cholesky Factorisation, see [21], [24]. These eigenvectors are the columns \mathbf{p}^i of a matrix \mathbf{P} :

$$\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n).$$

The directions of the eigenvectors are then used by the algorithms to perform the search. Algorithm 1 displays the pseudocode of the Analysis of Epistasis.

Algorithm 1 Analysis of Epistasis

INPUT objective function $f(\mathbf{x})$, decision space D , and parameters thr and $sampleSize$
 $h = 1$
for $s = 1 : sampleSize$ **do**
 sample a point \mathbf{x} in the decision space D
 if $f(\mathbf{x}) < thr$ **then**
 Insert \mathbf{x} into the data structure \mathbf{V} : $\mathbf{V}(h) = \mathbf{x}$
 $h = h + 1$
 end if
end for
Process the data structure \mathbf{V} to calculate the mean vector $\boldsymbol{\mu}$ and covariance matrix \mathbf{C}
Apply Cholesky Factorisation to extract the eigenvectors $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$
OUTPUT Matrix \mathbf{P}

From a theoretical standpoint, the analysis of epistasis is a procedure that returns a new reference system that is identified by the eigenvectors \mathbf{p}^k of the matrix \mathbf{C} . From linear algebra, we know that a change of coordinates is the multiplication of a matrix (whose columns are vectors of the new reference system) by the vector in the old reference system [21]. Hence, to express \mathbf{x} in the reference system identified by \mathbf{P} , we calculate the vector \mathbf{x}' in the new system as $\mathbf{x}' = \mathbf{P}\mathbf{x}$. Equivalently, every point \mathbf{x}' in the new reference system can be interpreted as a linear combination of the vectors of the basis $B_P = \{\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n\}$.

In order to justify the proposed new system of coordinates, we have to consider the fundamental meaning of the covariance matrix of a multivariate distribution, see [25]. The diagonal elements of the matrix directly describe the geometry of the problem since they represent the extent of the distribution along a variable. A diagonal element much larger in value than that of other diagonal elements means that the distribution is stretched along a design variable.

The extra-diagonal elements represent the correlation between pairs of variables. A large value means high correlation while zero means no correlation. In our case, since the points are sampled below a threshold, the correlation is meant with respect to the objective function: zero means that the function can be decomposed over the variables while a large value means that this decomposition is not possible. In order to intuitively visualise this fact, the covariance matrix associated with a sphere or a non-rotated ellipsoid would be diagonal while that associated with a rotated ellipsoid would be full.

The proposed analysis of epistasis detects that basis of vectors (system of coordinates) \mathbf{P} that diagonalises the covariance matrix \mathbf{C} [21], since

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{C}\mathbf{P}$$

where \mathbf{D} is a diagonal matrix whose elements are the eigenvalues of the matrix \mathbf{C} . Then, within this basin of attraction, we propose to perform the local search alongside the directions of the new variables. This operation can be interpreted as a problem transformation by the change of reference system. Since in the new system the problem is characterised by low epistasis, it can be addressed by moving alongside the new variables. This concept is broadly used in other contexts, especially in Data Science, and is closely related to Principal Component Analysis (PCA) [26].

In order to graphically represent the analysis of epistasis of the landscape let us consider the following objective functions in two dimensions within $[-100, 100]^2$, see e.g. [27].

$$\begin{aligned} \text{Sphere} & f(\mathbf{x}) = x_1^2 + x_2^2 \\ \text{Ellipsoid} & f(\mathbf{x}) = 50x_1^2 + 200x_2^2 \\ \text{Bent Cigar} & f(\mathbf{x}) = x_1^2 + 10^6x_2^2 \\ \text{Rosenbrock} & f(\mathbf{x}) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 \end{aligned}$$

Let us consider that these problems have been shifted by means of a bias and rotated by means of an orthogonal matrix. Fig. 1 shows the data set of points below the threshold (in blue) and the corresponding new reference system (dashed black lines) identified by the columns of the matrix \mathbf{P} .

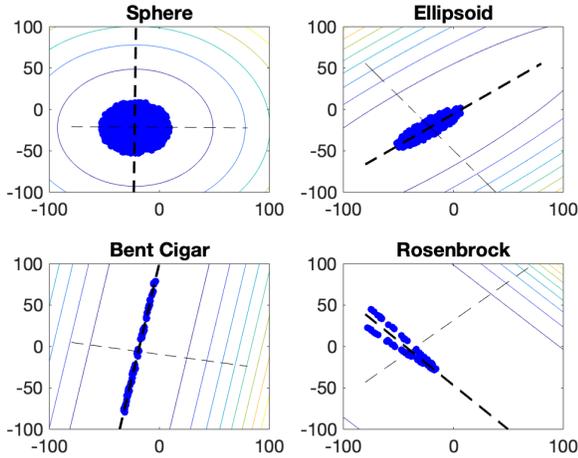


Fig. 1. Sampling of points (blue) within $[-100, 100]^2$ and new reference system (dashed black lines) for shifted and rotated Sphere, Ellipsoid, Bent Cigar, and Rosenbrock functions below the threshold values 10^3 , 3×10^4 , 10^6 , and 5×10^3 , respectively.

Furthermore, we may easily observe that the covariance matrix \mathbf{C} is symmetric, i.e. with reference to its definition above $c_{j,l} = c_{l,j}, \forall j, l$. Hence, the following observation is valid, see [21].

Observation 1: Since the covariance matrix \mathbf{C} is symmetric, it follows that

- 1) \mathbf{C} is always diagonalisable and hence there always exists a non-singular transformation matrix \mathbf{P} that diagonalises \mathbf{C} .
- 2) each pair of the eigenvectors of \mathbf{C} (column of matrix \mathbf{P}) is orthogonal ($\mathbf{P}^{-1} = \mathbf{P}^T$), the proposed reference system is orthogonal.

The meaning of this observation is that we are theoretically guaranteed that Algorithm 1 always returns a coherent output i.e. a new orthogonal reference system that diagonalises the covariance matrix.

A. Considerations about the Directional Derivatives

In order to better highlight the functioning of the proposed Analysis of Epistasis, we tested the values of the directional derivatives alongside the directions of the variables as well as those alongside the directions of the eigenvectors of the matrix \mathbf{C} (columns on the matrix \mathbf{P}).

With reference to Bent Cigar in Fig. 1 by using the detected minimum as a start point we performed ten steps of size 0.01 along the directions of the variables and those of the eigenvectors. At each step, we saved the objective function value and calculated the numerical directional derivative. Fig. 2 illustrates the plots of the directional derivatives alongside the four directions under consideration.

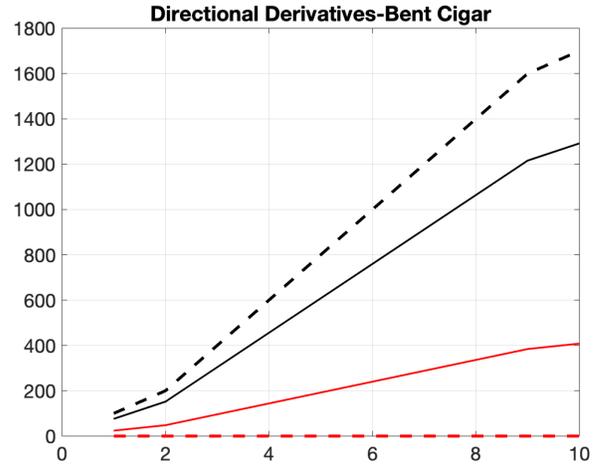


Fig. 2. Plot of the directional derivatives along the direction of the variables (continuous lines) and eigenvectors \mathbf{p}^1 and \mathbf{p}^2 (dashed lines) for the rotated bent cigar in two dimensions.

The test shows that one of the directional derivatives along one eigenvector has a high value while the other has nearly zero. The landscape analysis identifies a search direction along which small steps quickly lead to large improvements (steep slope) and another direction along which the problem is nearly “flat”. The directional derivatives along the variables present intermediate features.

We repeated the same test on a number of functions and multiple dimensions ranging from 2 to 100 dimensions. Apart from the sphere where, due to the central symmetry, all the lines of the derivative plots collapse in one line, the analysis

of epistasis systematically detected a direction whose gradient was higher than that of any variable directions. In other words, the proposed analysis of epistasis appears to be able to detect some preferential directions along which a local search displays a high convergence rate.

In order to show an example of the gradient in higher dimensions, Fig. 3 illustrates the results of the test for the rotated ill-conditioned ellipsoid in ten variables:

$$f(\mathbf{x}) = \sum_{i=1}^{10} (10^6)^{\frac{i-1}{9}} z_i^2.$$

where the design variables z_i are obtained from x_i after shifting and rotation, see [27]. The plots in Fig. 3 clearly show that the proposed analysis yields to the detection of one direction whose derivative is higher than that of the directional derivative along any variable directions. We may intuitively observe the correlation between this result and PCA [26].

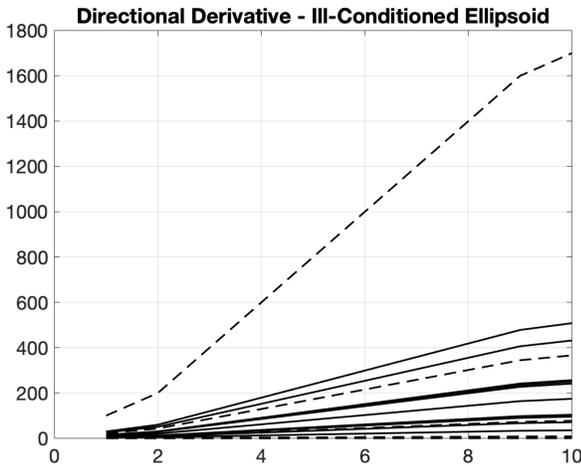


Fig. 3. Plot of the directional derivatives along the direction of the variables (continuous lines) and eigenvectors $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^{10})$ (dashed lines) for the rotated ill-conditioned ellipsoid in ten dimensions.

B. Limitation of the proposed Analysis of Epistasis

An evident limitation of the proposed Analysis of Epistasis outlined in Algorithm 1 is that its success depends on the threshold parameter thr . This parameter should allow the generation of a data set that describes the geometry of the problem. With reference to minimisation, a value that is too low would result in a small or empty data set, while a value that is too high would generate a dataset that does not contain the geometrical features of the basin of attraction. Furthermore, this parameter is problem dependent as it depends on the specific objective function values. Besides being inelegant, the sampling procedure can also be computationally expensive in the high dimensional domain.

On the other hand, the latter is an issue common to other FLA methods and is a consequence of the concept of FLA. Moreover, in the context of a Memetic Framework, we may suggest that part of the points visited in other phases of the

search can be saved in an archive and used for a local analysis of the fitness landscape. Finally, we observed that the selection of a reasonably good thr could be automatically performed with modest computational effort: we sample a small number of points ($10 \times n$) in the domain and selected the lowest objective function value among those of the sampled points as a threshold thr .

III. LOCAL SEARCH DESIGN OF MULTIPLE TRAJECTORY SEARCH

MTS is an MA composed of three local search algorithms, each of them improving upon the performance of a single solution \mathbf{x} . These three local search algorithms are supervised by a framework that processes a population of solutions and activates the local search on single solutions. The selection of the local search is adaptive and based on a mechanism that rewards the most successful local search algorithms, see [28].

The three local search algorithms, namely Local Search 1 (LS1), Local Search 2 (LS2), Local Search 3 (LS3) are briefly outlined in the following.

- LS1 is a greedy, deterministic local search [29] belonging to the family of pattern search [30] that explores all the variables of the problem in both directions
- LS2 is a randomised version of LS1 where only some of the design variables (about 25% of the variables) in a randomly selected direction are explored at each iteration
- LS3 is a steepest descent local search [31] that concurrently performs two actions: the deterministic perturbation and a randomised exploration along all the variables

This section reports the implementation of LS1, LS2, and LS3 in the new system of coordinates determined by the Analysis of Epistasis in Algorithm 1. Let us consider that a local basin of attraction has been identified and a matrix $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$ has been determined. The three modified local search algorithms, namely Covariance Local Search 1 (CLS1), Covariance Local Search 2 (CLS2), and Covariance Local Search 3 (CLS3) are reported in Algorithms 2, 3, and 4, respectively.

Although the pseudocodes are straightforward and only slightly modify those in [20], it is worthwhile mentioning that the change of variables alters the update rule. While in its original version the local search varies one single design variable at the time, in the proposed scheme the update of one variable in the new reference system corresponds to a sum of vectors. For example, with reference to Algorithm 2, the update of a single design variable

$$x_i^t = x_i - \rho$$

is equivalent to

$$\mathbf{x}^t = \mathbf{x} - \rho \cdot \mathbf{e}^i$$

that is a sum of numbers.

In the new reference system, the corresponding proposed update rule is

$$\mathbf{x}^t = \mathbf{x} - \rho \cdot \mathbf{p}^i$$

that is, unavoidably, a sum of vectors.

Algorithm 2 Covariance Local Search 1

```
INPUT  $\mathbf{x}$  and  $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$ 
while budget condition do
   $\rho = 0.4 \times \text{width of } D$ 
  while  $\rho > 10^{-15}$  do
     $\mathbf{x}^t = \mathbf{x}$ 
    for  $i = 1 : n$  do
       $\mathbf{x}^t = \mathbf{x} - \rho \cdot \mathbf{p}^i$ 
      if  $f(\mathbf{x}^t) < f(\mathbf{x})$  then
         $\mathbf{x} = \mathbf{x}^t$ 
      else if  $f(\mathbf{x}^t) > f(\mathbf{x})$  then
         $\mathbf{x}^t = \mathbf{x} + \frac{\rho}{2} \cdot \mathbf{p}^i$ 
        if  $f(\mathbf{x}^t) < f(\mathbf{x})$  then
           $\mathbf{x} = \mathbf{x}^t$ 
        end if
      end if
    end for
    if  $\mathbf{x}$  has not been updated then
       $\rho = \frac{\rho}{2}$ 
    end if
  end while
end while
RETURN  $\mathbf{x}$ 
```

It is worth mentioning the initialisation of ρ Algorithms 2 and 3 which means that 80% of the space is reachable with the first move. The stopping condition $\rho > 10^{-15}$ refers to the minimal radius allowed. The other parameters in Algorithms 2, 3, and 4 are the same parameters used in [20].

IV. NUMERICAL RESULTS

In order to experimentally demonstrate the effectiveness of the proposed Analysis of Epistasis, we tested the performance of CLS1, CLS2, and CLS3 and compared it against their original versions, LS1, LS2, LS3, respectively. For each experiment in this paper, LS1, LS2, and LS3 have been executed with a budget of $5000 \times n$ function calls where n is the problem dimensionality. In order to guarantee a fair comparison, the budget of CLS1, CLS2, and CLS3 has been split into two parts: $2500 \times n$ function calls have been used to build the covariance matrix \mathbf{C} whilst $2500 \times n$ function calls have been spent to execute the algorithm. The bound handling has been performed by saturating the design variable to the bound. We preferred the saturation to the bound over the toroidal insertion or reflection [32] since the latter two mechanisms would be equivalent to the sampling of a point. We evaluated that this sampling would modify the original logic of the local search algorithms.

Each algorithm for each problem had been run 30 times. Since the purpose of this experimental study is to check multiple pairwise comparisons LS# vs CLS# for each problem, we strengthen the statistical significance of the tests by the application of the Wilcoxon rank-sum test, see [33], [34]. In the Tables in this section, a “+” indicates that CLS# significantly outperforms LS#, a “-” indicates that LS#

Algorithm 3 Covariance Local Search 2

```
INPUT  $\mathbf{x}$  and  $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$ 
while budget condition do
   $\rho = 0.4 \times \text{width of } D$ 
  while  $\rho > 10^{-15}$  do
     $\mathbf{x}^t = \mathbf{x}$ 
    for  $i = 1 : n$  do
      generate a random integer  $r_i \in \{0, 1, 2, 3\}$ 
      generate a random integer  $d_i \in \{-1, 1\}$ 
      if  $r_i == 0$  then
         $\mathbf{x}^t = \mathbf{x}^t - \rho \cdot d_i \cdot \mathbf{p}^i$ 
      end if
    end for
    if  $f(\mathbf{x}^t) < f(\mathbf{x})$  then
       $\mathbf{x} = \mathbf{x}^t$ 
    else if  $f(\mathbf{x}^t) > f(\mathbf{x})$  then
       $\mathbf{x}^t = \mathbf{x}$ 
      for  $i = 1 : n$  do
        if  $r_i == 0$  then
           $\mathbf{x}^t = \mathbf{x}^t + \frac{\rho}{2} \cdot d_i \cdot \mathbf{p}^i$ 
        end if
      end for
      if  $f(\mathbf{x}^t) < f(\mathbf{x})$  then
         $\mathbf{x} = \mathbf{x}^t$ 
      end if
    end if
    if  $\mathbf{x}$  has not been updated then
       $\rho = \frac{\rho}{2}$ 
    end if
  end while
end while
RETURN  $\mathbf{x}$ 
```

significantly outperforms CLS#, and a “=” indicates that there is no significant difference in performance.

To test the enhancement in performance due to the new reference system (directions of eigenvectors) we built the following testbed. We simulated the local search conditions we considered the unimodal functions of CEC2008 [35] and CEC2013 [27]. More specifically the following testbed has been used.

- f_1 : Shifted Sphere Function, f_1 from [27] or nearly identical F_1 from [35]
- f_2 : Rotated High Conditioned Elliptic Function, f_2 from [27]
- f_3 : Rotated Bent Cigar Function, f_3 from [27]
- f_4 : Rotated Discus Function, f_4 from [27]
- f_5 : Different Powers Function, f_5 from [27]
- f_6 : Schwefel’s Problem 2.21, F_2 from [35]

Each problem has been considered in 10, 30, and 50 dimensions. In order to obtain highly epistatic basins of attraction, all the problems have been rotated. For the problems that were rotated in their original definitions (f_2 , f_3 and f_4), we used the rotation matrix provided with the test-bed. For the remaining

Algorithm 4 Covariance Local Search 3

```
INPUT  $\mathbf{x}$  and  $\mathbf{P} = (\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$ 
while budget condition do
   $\mathbf{x}^t = \mathbf{x}$ 
   $\mathbf{x}^s = \mathbf{x}$ 
  for  $i = 1 : n$  do
     $\mathbf{X}^1 = \mathbf{x}^s + 0.1 \cdot \mathbf{p}^i$ 
     $\mathbf{Y}^1 = \mathbf{x}^s - 0.1 \cdot \mathbf{p}^i$ 
     $\mathbf{X}^2 = \mathbf{x}^s + 0.2 \cdot \mathbf{p}^i$ 
    if  $f(\mathbf{X}^1) < f(\mathbf{x})$  then
       $\mathbf{x} = \mathbf{X}^1$ 
    end if
    if  $f(\mathbf{Y}^1) < f(\mathbf{x})$  then
       $\mathbf{x} = \mathbf{Y}^1$ 
    end if
    if  $f(\mathbf{X}^2) < f(\mathbf{x})$  then
       $\mathbf{x} = \mathbf{X}^2$ 
    end if
     $D_1 = f(\mathbf{x}^s) - f(\mathbf{X}^1)$ 
     $D_2 = f(\mathbf{x}^s) - f(\mathbf{Y}^1)$ 
     $D_3 = f(\mathbf{x}^s) - f(\mathbf{X}^2)$ 
    generate a random real number  $a \in [0.4, 0.5]$ 
    generate a random real number  $b \in [0.1, 0.3]$ 
    generate a random real number  $c \in [0, 1]$ 
     $\delta = a(D_1 - D_2) + b(D_3 - 2D_1) + c$ 
     $\mathbf{x}^t = \mathbf{x}^t + \delta \cdot \mathbf{p}^i$ 
  end for
  if  $f(\mathbf{x}^t) < f(\mathbf{x})$  then
     $\mathbf{x} = \mathbf{x}^t$ 
  end if
end while
RETURN  $\mathbf{x}$ 
```

problems, we used a randomly generated rotation/orthogonal matrix, see [36], [37]. The following manually tuned threshold values thr were used.

| n | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 |
|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-------------------|
| 10 | 10^4 | 5×10^8 | 10^9 | 5×10^8 | 3×10^4 | 10^2 |
| 30 | 5×10^4 | 2×10^9 | 2×10^9 | 10^8 | 10^5 | 1.1×10^2 |
| 50 | 10^5 | 5×10^9 | 5×10^7 | 3×10^5 | 10^5 | 1.2×10^2 |

Tables I, II, and III display the average $\mu \pm$ the standard deviation σ over the 30 independent runs available and the results of the Wilcoxon test. The best results are emphasised in bold.

Numerical results displayed in Tables I, II, III show that the proposed Analysis of Epistasis is overall beneficial to the performance of the three local search algorithms. As a first comment, we should not expect an improvement for f_1 since a sphere is a lowly epistatic problem due to its central symmetry. In other words, the reference system will not affect the local search performance. The benefits of the proposed fitness landscape analysis are evident for LS1 where all the directions of the new basis of vectors $(\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n)$ are exploited. Results on LS3 display that the proposed approach

is beneficial also to for this local search. The performance of CLS3 is marginally but consistently better than that of LS3. It should be observed that also in this case the entire basis of vectors is exploited. However, most of the exploration is set in the close neighbourhood of the current best point and since the radius is never reduced there is no clear mechanism to follow the variations of the gradient. Hence, the potentials of the information provided by the Analysis of Epistasis are not fully exploited. Furthermore, we observed that the randomisation in LS3 appears to be detrimental to the proposed fitness landscape approach. The advantages of the proposed fitness landscape approach on LS2 are, albeit present, not substantial. In this case, the local search selects randomly only 25% (on average) of the variables/search directions. This mechanism does not exploit the potentials of the proposed fitness landscape analysis since the new variables with high directional derivatives may be excluded from the search, see Figs 2 and 3.

In order to better highlight the impact of the proposed approach on the local search performance, Figs 4, 5, 6 show the performance trends of LS1 vs CLS1, LS2 vs CLS2, and LS3 vs CLS3, respectively.

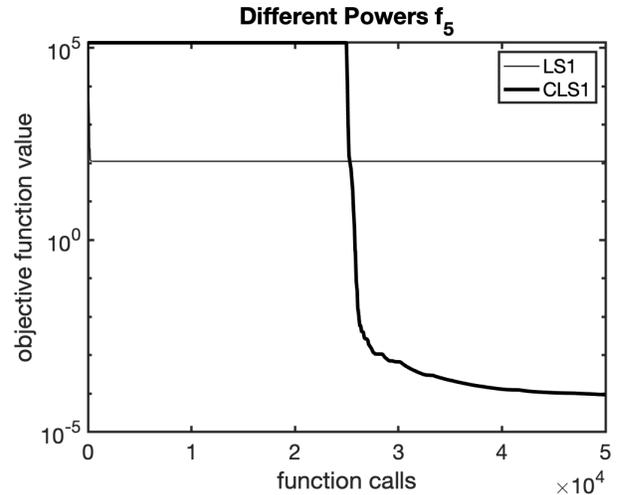


Fig. 4. Performance trend (logarithmic scale) of LS1 vs CLS1 for the Different Powers Function f_5 in 10D

The performance trends in Figs 4, 5, and 6 show that for these problems the use of the eigenvectors of the covariance matrix as new variables is very efficient. It should be highlighted that the CLS# algorithms require half of the computational budget to analyse the landscape and build the covariance matrix. This explains why there are no improvements in the fitness values for the first half of the trend. However, the displayed results show that in many cases this investment in the budget pays off with a rapid improvement in performance during the second half of the run. This consideration is aligned with the algorithmic philosophy of Fitness Landscape Analysis [4], [8], [14] and should be taken into account while local search elements are designed within Memetic Frameworks [38], [39].

TABLE I
AVERAGE ERROR AVG \pm STANDARD DEVIATION σ OVER 30 RUNS FOR THE PROBLEMS $f_1 - f_7$ IN 10 DIMENSIONS.

| | LS1 | | CLS1 | | LS2 | | CLS2 | | LS3 | | CLS3 | |
|-------|--------------------------------|---|--------------------------------|--|--------------------------------|---|--------------------------------|--|--------------------------------|---|--------------------------------|--|
| | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | |
| f_1 | 0.00e+00 \pm 0.00e+00 | - | 1.12e-28 \pm 9.99e-29 | | 3.16e-30 \pm 1.18e-29 | - | 1.54e-28 \pm 1.45e-28 | | 9.03e+03 \pm 5.43e+03 | = | 8.00e+03 \pm 5.27e+03 | |
| f_2 | 2.51e+05 \pm 3.05e+05 | + | 1.49e+04 \pm 1.24e+04 | | 8.41e+04 \pm 4.74e+04 | + | 4.51e+04 \pm 3.38e+04 | | 1.02e+08 \pm 1.41e+08 | + | 3.55e+07 \pm 5.82e+07 | |
| f_3 | 1.28e+03 \pm 1.58e+03 | + | 4.93e+01 \pm 2.33e+02 | | 5.56e+03 \pm 5.56e+03 | + | 3.44e+02 \pm 1.13e+03 | | 2.54e+03 \pm 3.69e+03 | + | 2.53e+02 \pm 8.33e+02 | |
| f_4 | 1.19e+04 \pm 1.50e+04 | + | 2.15e-25 \pm 3.64e-25 | | 2.17e+04 \pm 2.67e+04 | + | 6.94e-25 \pm 1.91e-24 | | 1.99e+07 \pm 1.09e+08 | + | 7.02e-03 \pm 1.75e-02 | |
| f_5 | 4.75e+01 \pm 1.36e+02 | + | 5.08e-05 \pm 2.44e-05 | | 1.12e-04 \pm 2.36e-05 | = | 1.20e-04 \pm 3.62e-05 | | 6.36e+03 \pm 7.91e+03 | = | 7.99e+03 \pm 1.38e+04 | |
| f_6 | 2.34e+01 \pm 1.95e+01 | + | 5.31e+00 \pm 9.26e+00 | | 7.09e+00 \pm 1.17e+01 | + | 1.33e+00 \pm 5.87e+00 | | 6.20e+01 \pm 2.13e+01 | + | 5.58e+01 \pm 1.62e+01 | |

TABLE II
AVERAGE ERROR AVG \pm STANDARD DEVIATION σ OVER 30 RUNS FOR THE PROBLEMS $f_1 - f_7$ IN 30 DIMENSIONS.

| | LS1 | | CLS1 | | LS2 | | CLS2 | | LS3 | | CLS3 | |
|-------|--------------------------------|---|--------------------------------|--|--------------------------------|---|--------------------------------|--|-------------------------|---|--------------------------------|--|
| | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | |
| f_1 | 0.00e+00 \pm 0.00e+00 | - | 5.09e-28 \pm 2.81e-28 | | 1.18e-28 \pm 1.02e-28 | = | 2.76e-27 \pm 1.41e-27 | | 8.12e+04 \pm 2.14e+04 | = | 7.81e+04 \pm 2.12e+04 | |
| f_2 | 1.11e+06 \pm 2.29e+06 | + | 2.24e+05 \pm 1.21e+05 | | 1.80e+05 \pm 8.26e+04 | = | 1.78e+05 \pm 7.09e+04 | | 2.19e+09 \pm 1.26e+09 | + | 1.28e+09 \pm 1.02e+09 | |
| f_3 | 4.40e+03 \pm 6.24e+03 | + | 2.86e-04 \pm 1.56e-03 | | 6.67e+03 \pm 8.92e+03 | + | 1.12e-23 \pm 5.16e-23 | | 1.50e+03 \pm 2.30e+03 | + | 4.87e+01 \pm 1.83e+02 | |
| f_4 | 1.35e+04 \pm 1.57e+04 | + | 1.75e-26 \pm 4.37e-26 | | 3.51e-01 \pm 8.25e-01 | + | 1.23e-26 \pm 1.50e-26 | | 1.20e+09 \pm 1.94e+09 | + | 1.65e+06 \pm 9.01e+06 | |
| f_5 | 1.40e+02 \pm 2.22e+02 | + | 1.29e-04 \pm 3.90e-05 | | 5.60e-05 \pm 3.57e-06 | = | 5.80e-05 \pm 9.18e-06 | | 3.99e+05 \pm 4.13e+05 | + | 1.44e+05 \pm 1.43e+05 | |
| f_6 | 5.54e+01 \pm 1.13e+01 | + | 3.82e+01 \pm 1.33e+01 | | 2.96e+01 \pm 1.24e+01 | = | 2.74e+01 \pm 1.14e+01 | | 1.05e+02 \pm 1.97e+01 | = | 9.93e+01 \pm 1.51e+01 | |

TABLE III
AVERAGE ERROR AVG \pm STANDARD DEVIATION σ OVER 30 RUNS FOR THE PROBLEMS $f_1 - f_7$ IN 50 DIMENSIONS.

| | LS1 | | CLS1 | | LS2 | | CLS2 | | LS3 | | CLS3 | |
|-------|--------------------------------|---|--------------------------------|--|--------------------------------|---|--------------------------------|--|-------------------------|---|--------------------------------|--|
| | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | | $\mu \pm \sigma$ | W | $\mu \pm \sigma$ | |
| f_1 | 0.00e+00 \pm 0.00e+00 | - | 1.19e-27 \pm 3.89e-28 | | 3.99e-28 \pm 2.08e-28 | = | 7.79e-27 \pm 3.06e-27 | | 1.47e+05 \pm 2.66e+04 | = | 1.40e+05 \pm 2.32e+04 | |
| f_2 | 9.75e+05 \pm 9.69e+05 | + | 2.12e+05 \pm 1.24e+05 | | 2.76e+05 \pm 8.72e+04 | = | 2.60e+05 \pm 1.68e+05 | | 6.64e+09 \pm 4.23e+09 | + | 3.40e+09 \pm 1.93e+09 | |
| f_3 | 3.47e+03 \pm 4.14e+03 | + | 5.21e-24 \pm 1.99e-23 | | 6.66e+03 \pm 6.96e+03 | + | 5.99e-24 \pm 2.17e-23 | | 1.93e+04 \pm 1.17e+04 | + | 7.36e+03 \pm 7.49e+03 | |
| f_4 | 7.04e+03 \pm 8.53e+03 | + | 2.33e-27 \pm 3.08e-27 | | 5.20e-21 \pm 4.69e-21 | + | 1.36e-26 \pm 3.52e-26 | | 3.27e+07 \pm 1.50e+08 | + | 1.79e+03 \pm 9.57e+03 | |
| f_5 | 1.64e+02 \pm 1.67e+02 | + | 1.87e-04 \pm 4.11e-05 | | 7.44e-05 \pm 3.33e-06 | = | 6.52e-05 \pm 6.46e-06 | | 2.47e+05 \pm 1.79e+05 | + | 1.67e+05 \pm 1.08e+05 | |
| f_6 | 5.94e+01 \pm 7.33e+00 | + | 4.13e+01 \pm 4.55e+00 | | 5.65e+01 \pm 1.09e+01 | = | 5.24e+01 \pm 4.32e+00 | | 1.10e+02 \pm 1.58e+01 | = | 1.08e+02 \pm 1.58e+01 | |

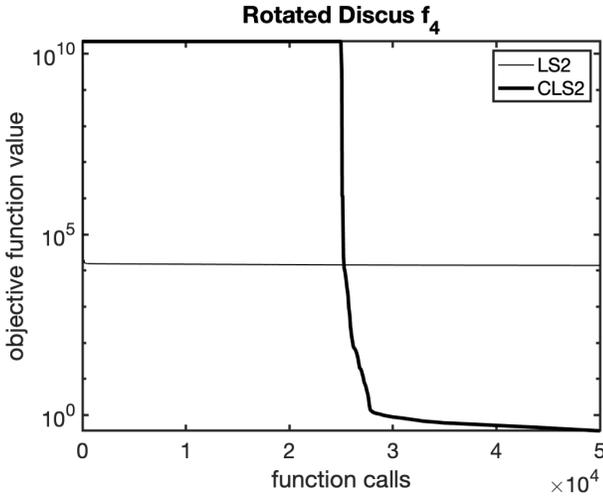


Fig. 5. Performance trend (logarithmic scale) of LS2 vs CLS2 for the Rotated Discus Function f_4 in 10D

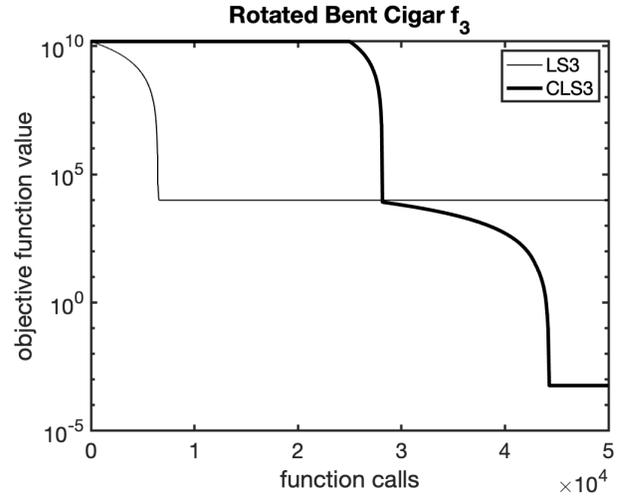


Fig. 6. Performance trend (logarithmic scale) of LS3 vs CLS3 for the Rotated Bent Cigar f_3 in 10D

V. CONCLUSION

This paper proposes a method to analyse the fitness landscape of continuous optimisation problems. The analysis aims to assess the epistasis of a basin of attraction to exploit this information to design an effective local search within

Memetic Frameworks. The proposed method makes use of sampled points describing the geometry of the problem and the covariance matrix associated with the distribution of these points. The local search explores the basin of attraction along the directions of the eigenvectors of this covariance matrix. In other words, the local search works on an alternative reference

system of variables. This alternative reference system ensures that the covariance matrix is diagonalised and then the problem appears to be lowly epistatic.

The proposed method has been implemented on the three local search algorithms composing a popular MA called Multiple Trajectory Search. Multiple experiments have been run on a number of rotated unimodal problems. These unimodal problems simulate basins of attraction detected during the search of an MA. Numerical results show that the proposed landscape fitness analysis detects a set of preferential search directions. One of these directions is associated with a directional derivative higher than that of the directional derivatives associated with the directions of all the variables. Furthermore, each local search designed on the proposed fitness landscape analysis outperforms its original counterpart.

The proposed fitness landscape analysis can be potentially integrated within Memetic Frameworks to automatically design problem-specific local search algorithms.

REFERENCES

- [1] D. Molina, M. Lozano, C. Garcia-Martinez, and F. Herrera, "Memetic Algorithms for Continuous Optimization Based on Local Search Chains," *Evolutionary Computation*, vol. 18, no. 1, pp. 27–63, 2010.
- [2] C. Reeves and J. E. Rowe, *Genetic Algorithms: Principles and Perspectives*. Springer, 2002.
- [3] P. Merz and B. Freisleben, "Fitness landscape analysis and memetic algorithms for the quadratic assignment problem," *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 4, pp. 337–352, 2000.
- [4] K. M. Malan and A. P. Engelbrecht, "A survey of techniques for characterising fitness landscapes and some possible ways forward," *Information Sciences*, vol. 241, pp. 148 – 163, 2013.
- [5] L. Vanneschi, M. Clergue, P. Collard, M. Tomassini, and S. Vérel, "Fitness clouds and problem hardness in genetic programming," in *Genetic and Evolutionary Computation - GECCO 2004, Genetic and Evolutionary Computation Conference, Seattle, WA, USA, June 26-30, 2004, Proceedings, Part II*, pp. 690–701, 2004.
- [6] L. Vanneschi, M. Tomassini, P. Collard, and S. Vérel, "Negative slope coefficient: A measure to characterize genetic programming fitness landscapes," in *Genetic Programming, 9th European Conference, EuroGP, Budapest, Hungary, April 10-12, 2006, Proc.*, pp. 178–189, 2006.
- [7] Y. Borenstein and R. Poli, "Information landscapes and problem hardness," in *Genetic and Evolutionary Computation Conference, GECCO 2005, Proceedings, Washington DC, USA, June 25-29, 2005* (H. Beyer and U. O'Reilly, eds.), pp. 1425–1431, ACM, 2005.
- [8] N. D. Jana, J. Sil, and S. Das, "Continuous fitness landscape analysis using a chaos-based random walk algorithm," *Soft Computing*, vol. 22, pp. 921–948, 2018.
- [9] P. Merz, "Advanced fitness landscape analysis and the performance of memetic algorithms," *Evolutionary Computation*, vol. 12, no. 3, pp. 303–325, 2004.
- [10] Z. Zhu, S. Jia, S. He, Y. Sun, Z. Ji, and L. Shen, "Three-dimensional gabor feature extraction for hyperspectral imagery classification using a memetic framework," *Information Sciences*, vol. 298, pp. 274 – 287, 2015.
- [11] L. Ma, J. Li, Q. Lin, M. Gong, C. A. Coello Coello, and Z. Ming, "Cost-aware robust control of signed networks by using a memetic algorithm," *IEEE Transactions on Cybernetics*, pp. 1–14, 2019.
- [12] C. Liao and C. Ting, "A novel integer-coded memetic algorithm for the setk-cover problem in wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 48, no. 8, pp. 2245–2258, 2018.
- [13] K. M. Malan and A. P. Engelbrecht, "Quantifying ruggedness of continuous landscapes using entropy," in *2009 IEEE Congress on Evolutionary Computation*, pp. 1440–1447, 2009.
- [14] K. M. Malan and A. P. Engelbrecht, "A progressive random walk algorithm for sampling continuous fitness landscapes," in *2014 IEEE Congress on Evolutionary Computation (CEC)*, pp. 2507–2514, 2014.
- [15] Y. Davidor, "Epistasis variance: A viewpoint on ga-hardness," vol. 1 of *Foundations of Genetic Algorithms*, pp. 23 – 35, Elsevier, 1991.
- [16] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 378–393, 2014.
- [17] F. Caraffini, F. Neri, and L. Picinali, "An analysis on separability for memetic computing automatic design," *Information Sciences*, vol. 265, pp. 1–22, 2014.
- [18] X. Ma, Q. Zhang, G. Tian, J. Yang, and Z. Zhu, "On tchebycheff decomposition approaches for multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 226–244, 2018.
- [19] W. E. Hart, N. Krasnogor, and J. E. Smith, "Memetic Evolutionary Algorithms," in *Recent Advances in Memetic Algorithms* (W. E. Hart, N. Krasnogor, and J. E. Smith, eds.), (Berlin, Germany), pp. 3–27, Springer, 2004.
- [20] L.-Y. Tseng and C. Chen, "Multiple trajectory search for Large Scale Global Optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 3052–3059, 2008.
- [21] F. Neri, *Linear Algebra for Computational Sciences and Engineering*. Springer, second ed., 2019.
- [22] L. Ma, M. Gong, J. Yan, W. Liu, and S. Wang, "Detecting composite communities in multiplex networks: A multilevel memetic algorithm," *Swarm and Evolutionary Computation*, vol. 39, pp. 177 – 191, 2018.
- [23] M. Gong, C. Chen, Y. Xie, and S. Wang, "Community preserving network embedding based on memetic algorithm," *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–11, 2018.
- [24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C (2Nd Ed.): The Art of Scientific Computing*. New York, NY, USA: Cambridge University Press, 1992.
- [25] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1. Wiley, January 1968.
- [26] I. T. Jolliffe, *Principal Component Analysis*. Springer Series in Statistics, Springer, 2nd ed., 2002.
- [27] J. J. Liang, B. Y. Qu, P. N. Suganthan, and A. G. Hernández-Díaz, "Problem Definitions and Evaluation Criteria for the CEC 2013 Special Session on Real-Parameter Optimization," Tech. Rep. 201212, Zhengzhou University and Nanyang Technological University, Zhengzhou China and Singapore, 2013.
- [28] F. Caraffini, F. Neri, and M. Epitropakis, "Hyperspan: A study on hyperheuristic coordination strategies in the continuous domain," *Information Sciences*, vol. 477, pp. 186 – 202, 2019.
- [29] M. A. Montes de Oca, C. Cotta, and F. Neri, "Local search," in *Handbook of Memetic Algorithms* (F. Neri, C. Cotta, and P. Moscato, eds.), vol. 379 of *Studies in Computational Intelligence*, Springer, 2012.
- [30] V. Torczon, "On the convergence of pattern search algorithms," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 1–25, 1997.
- [31] M. A. Montes de Oca, T. Stutzle, M. Birattari, and M. Dorigo, "Frankenstein's PSO: A Composite Particle Swarm Optimization Algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 1120–1132, 2009.
- [32] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computation*. Berlin: Springer, 2015. second edition.
- [33] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.
- [34] S. García, A. Fernandez, J. Luengo, and F. Herrera, "A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability," *Soft Computing*, vol. 13, no. 10, pp. 959–977, 2008.
- [35] K. Tang, X. Yao, P. N. Suganthan, C. MacNish, Y. P. Chen, C. M. Chen, and Z. Yang, "Benchmark Functions for the CEC 2008 Special Session and Competition on Large Scale Global Optimization," tech. rep., Nature Inspired Computation and Applications Laboratory, USTC, China, 2007.
- [36] F. Caraffini and F. Neri, "Rotation invariance and rotated problems: An experimental study on differential evolution," in *Applications of Evolutionary Computation* (K. Sim and P. Kaufmann, eds.), vol. 10784 of *Lecture Notes in Computer Science*, pp. 597–614, Springer, 2018.
- [37] F. Caraffini and F. Neri, "A study on rotation invariance in differential evolution," *Swarm and Evolutionary Computation*, vol. 50, p. 100436, 2019.
- [38] Z. Zhu, Y. Ong, and M. Dash, "Wrapper-filter feature selection algorithm using a memetic framework," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 1, pp. 70–76, 2007.
- [39] F. Caraffini, F. Neri, G. Iacca, and A. Mol, "Parallel memetic structures," *Information Sciences*, vol. 227, no. 0, pp. 60 – 82, 2013.