



**HAL**  
open science

## Pareto Local Optima of Multiobjective NK-Landscapes with Correlated Objectives

Sébastien Verel, Arnaud Liefooghe, Laetitia Jourdan, Clarisse Dhaenens

► **To cite this version:**

Sébastien Verel, Arnaud Liefooghe, Laetitia Jourdan, Clarisse Dhaenens. Pareto Local Optima of Multiobjective NK-Landscapes with Correlated Objectives. 11th European Conference on Evolutionary Computation in Combinatorial Optimisation, Apr 2011, Turino, Italy. pp.226–237. hal-00563460

**HAL Id: hal-00563460**

**<https://hal.science/hal-00563460v1>**

Submitted on 18 Jul 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Pareto Local Optima of Multiobjective NK-Landscapes with Correlated Objectives

Sébastien Verel<sup>1,3</sup>, Arnaud Liefooghe<sup>2,3</sup>,  
Laetitia Jourdan<sup>3</sup>, and Clarisse Dhaenens<sup>2,3</sup>

<sup>1</sup> University of Nice Sophia Antipolis – CNRS, France

<sup>2</sup> Université Lille 1, LIFL – CNRS, France

<sup>3</sup> INRIA Lille-Nord Europe, France

verel@i3s.unice.fr, arnaud.liefooghe@univ-lille1.fr,  
laetitia.jourdan@inria.fr, clarisse.dhaenens@lifl.fr

**Abstract.** In this paper, we conduct a fitness landscape analysis for multiobjective combinatorial optimization, based on the local optima of multiobjective *NK*-landscapes with objective correlation. In single-objective optimization, it has become clear that local optima have a strong impact on the performance of metaheuristics. Here, we propose an extension to the multiobjective case, based on the Pareto dominance. We study the co-influence of the problem dimension, the degree of non-linearity, the number of objectives and the correlation degree between objective functions on the number of Pareto local optima.

## 1 Motivations

The aim of fitness landscape analysis is to understand the properties of a given combinatorial optimization problem in order to design efficient search algorithms. One of the main feature is related to the number of local optima, to their distribution over the search space and to the shape of their basins of attraction. For instance, in single-objective optimization, it has been shown that local optima tend to be clustered in a ‘central massif’ for numerous combinatorial problems, such as the family of *NK*-landscapes [1]. A lot of methods are designed to ‘escape’ from such local optima. However, very little is known in the frame of multiobjective combinatorial optimization (MoCO), where one of the most challenging question relies on the identification of the set of Pareto optimal solutions. A Pareto Local Optima (PLO) [2] is a solution that is not dominated by any of its neighbors. The description of PLO is one of the first fundamental step towards the description of the structural properties of a MoCO problem. Surprisingly, up to now, there is a lack of study on the number and on the distribution of PLO in MoCO.

Like in single-objective optimization, the PLO-related properties clearly have a strong impact on the landscape of the problem, and then on the efficiency of search algorithms. In particular, local search algorithms are designed in order to take them into account. For instance, the family of Pareto Local Search (PLS) [2] iteratively improves a set of solutions with respect to a given neighborhood

operator and to the Pareto dominance relation. The aim of PLS, like a number of other search algorithms, is to find a set of mutually non-dominated PLO. PLS has been proved to terminate on such a set, called a *Pareto local optimum set* [2]. Notice that a Pareto optimal solution is a PLO, and that the whole set of Pareto optimal solutions is a Pareto local optimum set. The behavior of multiobjective algorithms clearly depends on the properties related to the PLO. First, a Pareto local optimum set is always a subset of the whole set of PLO. Second, the dynamics of a PLS-like algorithm depends of the number of PLO found along the search process. The probability to improve an approximation set that contains a majority of PLO should be smaller than the probability to improve an approximation set with few PLO.

There exists a small amount of literature related to fitness landscape for MoCO. Borges and Hansen [3] study the distribution of local optima, in terms of scalarized functions, for the multiobjective traveling salesman problem (TSP). Another analysis of neighborhood-related properties for biobjective TSP instances of different structures is given in [4]. Knowles and Corne [5] lead a landscape analysis on the multiobjective quadratic assignment problem with a rough objective correlation. Next, the transposition of standard tools from fitness landscape analysis to MoCO are discussed by Garrett [6], and an experimental study is conducted with fitness distance correlation. But this measure requires the true Pareto optimal set to be known. In another study, the landscape of a MoCO problem is regarded as a neutral landscape, and divided into different fronts with the same dominance rank [7]. In such a case, a small search space needs to be enumerated. In previous works on multiobjective  $NK$ -landscapes by Aguirre and Tanaka [8], small enumerable fitness landscapes are studied according to the number of fronts, the number of solutions on each front, the probability to pass from one front to another, and the hypervolume of the Pareto front. However, the study of fronts simply allows to analyze small search spaces, and from the point of view of dominance rank only.

In this work, our attempt is to analyze the structure of large search space using the central notion of local optimum. For the design of a local search algorithm for MoCO, the following questions are under study in this paper: (i) What is the number of PLO in the whole search space? (ii) Is the number of PLO related to the number of Pareto optimal solutions? In particular we want to study such properties according to the correlation degree between objective functions. In order to study the problem structure, and in particular the PLO, we use the multiobjective  $NK$ -landscapes with objective correlation,  $\rho MNK$ -landscapes for short, recently proposed in [9]. The contributions of this work can be summarized as follows. First, we show the co-influence of objective correlation, objective space dimension and epistasis on the number of PLO. Next, we propose a method based on the length of a Pareto adaptive walk to estimate this number. At last, we study the number of PLO for large-size instances.

The paper is organized as follows. Section 2 deals with MoCO and local search algorithms. Section 3 is devoted to the definition of multiobjective  $NK$ -landscapes with objective correlation. In Section 4, we study the number of PLO

for enumerable instances and we propose a method to estimate it. Moreover, we analyze the correlation between the number of PLO and Pareto optimal solutions. In Section 5, the co-influence of objective space dimension, objective correlation and epistasis is studied for the PLO of large-size instances. The last section concludes the paper.

## 2 Local Search for Multiobjective Combinatorial Optimization

### 2.1 Multiobjective Combinatorial Optimization

A multiobjective combinatorial optimization (MoCO) problem can be defined by a set of  $M \geq 2$  objective functions  $(f_1, f_2, \dots, f_M)$ , and a (discrete) set  $X$  of feasible solutions in the *decision space*. Let  $Z = f(X) \subseteq \mathbb{R}^M$  be the set of feasible outcome vectors in the *objective space*. In a maximization context, a solution  $x' \in X$  is dominated by a solution  $x \in X$ , denoted by  $x' \prec x$ , iff  $\forall i \in \{1, 2, \dots, M\}$ ,  $f_i(x') \leq f_i(x)$  and  $\exists j \in \{1, 2, \dots, M\}$  such that  $f_j(x') < f_j(x)$ . A solution  $x \in X$  is said to be *Pareto optimal* (or *efficient*, *non-dominated*), if there does not exist any other solution  $x' \in X$  such that  $x'$  dominates  $x$ . The set of all Pareto optimal solutions is called the *Pareto optimal set* (or the *efficient set*), denoted by  $X_E$ , and its mapping in the objective space is called the *Pareto front*. A possible approach in MoCO is to identify the minimal complete Pareto optimal set, *i.e.* one solution mapping to each point of the Pareto front. However, the overall goal is often to identify a good Pareto set approximation. To this end, metaheuristics in general, and evolutionary algorithms in particular, have received a growing interest since the late eighties. Multiobjective metaheuristics still constitute an active research area [10].

### 2.2 Local Search

A *neighborhood structure* is a function  $\mathcal{N} : X \rightarrow 2^X$  that assigns a set of solutions  $\mathcal{N}(x) \subset X$  to any solution  $x \in X$ . The set  $\mathcal{N}(x)$  is called the *neighborhood* of  $x$ , and a solution  $x' \in \mathcal{N}(x)$  is called a *neighbor* of  $x$ . In single-objective combinatorial optimization, a fitness landscape can be defined by the triplet  $(X, \mathcal{N}, h)$ , where  $h : X \rightarrow \mathbb{R}$  represents the fitness function, that can be pictured as the *height* of the corresponding solutions. Each peak of the landscape corresponds to a local optimum. In a single-objective maximization context, a *local optimum* is a solution  $x^*$  such that  $\forall x \in \mathcal{N}(x^*)$ ,  $f(x) \leq f(x^*)$ . The ability of local search algorithms has been shown to be related to the number of local optima for the problem under study, and to their distribution over the landscapes [11].

In MoCO, given that Pareto optimal solutions are to be found, the notion of local optimum has to be defined in terms of Pareto optimality. Let us define the concepts of Pareto local optimum and of Pareto local optimum set. For more details, refer to [2]. A solution  $x \in X$  is a *Pareto local optimum* (PLO) with respect to a neighborhood structure  $\mathcal{N}$  if there does not exist any neighboring solution

$x' \in \mathcal{N}(x)$  such that  $x \prec x'$ . A *Pareto local optimum set*  $X_{PLO} \in X$  with respect to a neighborhood structure  $\mathcal{N}$  is a set of mutually non-dominated solutions such that  $\forall x \in X_{PLO}$ , there does not exist any solution  $x' \in \mathcal{N}(X_{PLO})$  such that  $x \prec x'$ . In other words, a locally Pareto optimal set cannot be improved, in terms of Pareto optimality, by adding solutions from its neighborhood.

Recently, local search algorithms have been successfully applied to MoCO problems. Such methods seem to take advantage of some properties of the landscape in order to explore the search space in an effective way. Two main classes of local search for MoCO can be distinguished. The first ones, known as *scalar approaches*, are based on multiple scalarized aggregations of the objective functions. The second ones, known as *Pareto-based approaches*, directly or indirectly focus the search on the Pareto dominance relation (or a slight modification of it). One of them is the Pareto Local Search (PLS) [2]. It combines the use of a neighborhood structure with the management of an archive (or population) of mutually non-dominated solutions found so far. The basic idea is to iteratively improve this archive by exploring the neighborhood of its own content until no further improvement is possible, *i.e.* the archive falls in a Pareto local optimum set [2].

### 3 $\rho MNK$ -Landscapes: Multiobjective $NK$ -Landscapes with Objective Correlation

In single-objective optimization, the family of  $NK$ -landscapes constitutes an interesting model to study the influence of non-linearity on the number of local optima. In this section, we present the  $\rho MNK$ -landscapes proposed in [9]. They are based on the  $MNK$ -landscapes [8]. In this multiobjective model, the correlation between objective functions can be precisely tuned by a correlation parameter value.

#### 3.1 $NK$ - and $MNK$ -Landscapes

The family of  $NK$ -landscapes [1] is a problem-independent model used for constructing multimodal landscapes.  $N$  refers to the number of (binary) genes in the genotype (*i.e.* the string length) and  $K$  to the number of genes that influence a particular gene from the string (the epistatic interactions). By increasing the value of  $K$  from 0 to  $(N - 1)$ ,  $NK$ -landscapes can be gradually tuned from smooth to rugged. The fitness function (to be maximized) of a  $NK$ -landscape  $f_{NK} : \{0, 1\}^N \rightarrow [0, 1)$  is defined on binary strings of size  $N$ . An ‘atom’ with fixed epistasis level is represented by a fitness component  $f_i : \{0, 1\}^{K+1} \rightarrow [0, 1)$  associated to each bit  $i \in N$ . Its value depends on the allele at bit  $i$  and also on the alleles at  $K$  other epistatic positions ( $K$  must fall between 0 and  $N - 1$ ). The fitness  $f_{NK}(x)$  of a solution  $x \in \{0, 1\}^N$  corresponds to the mean value of its  $N$  fitness components  $f_i$ :  $f_{NK}(x) = \frac{1}{N} \sum_{i=1}^N f_i(x_i, x_{i_1}, \dots, x_{i_K})$ , where  $\{i_1, \dots, i_K\} \subset \{1, \dots, i - 1, i + 1, \dots, N\}$ . In this work, we set the  $K$  bits randomly on the bit string of size  $N$ . Each fitness component  $f_i$  is specified by extension, *i.e.* a number  $y_{x_i, x_{i_1}, \dots, x_{i_K}}^i$  from  $[0, 1)$  is associated with each element

$(x_i, x_{i_1}, \dots, x_{i_K})$  from  $\{0, 1\}^{K+1}$ . Those numbers are uniformly distributed in the range  $[0, 1)$ .

More recently, a multiobjective variant of  $NK$ -landscapes (namely  $MNK$ -landscapes) [8] has been defined with a set of  $M$  fitness functions:

$$\forall m \in [1, M], f_{NK_m}(x) = \frac{1}{N} \sum_{i=1}^N f_{m,i}(x_i, x_{i_{m,1}}, \dots, x_{i_{m,K_m}})$$

The numbers of epistasis links  $K_m$  can theoretically be different for each fitness function. But in practice, the same epistasis degree  $K_m = K$  for all  $m \in [1, M]$  is used. Each fitness component  $f_{m,i}$  is specified by extension with the numbers  $y_{x_i, x_{i_{m,1}}, \dots, x_{i_{m,K_m}}}^{m,i}$ . In the original  $MNK$ -landscapes [8], these numbers are randomly and independently drawn from  $[0, 1)$ . As a consequence, it is very unlikely that two different solutions map to the same point in the objective space.

### 3.2 $\rho MNK$ -Landscapes

In [9],  $CMNK$ -landscapes have been proposed. The epistasis structure is identical for all the objectives:  $\forall m \in [1, M], K_m = K$  and  $\forall m \in [1, M], \forall j \in [1, K], i_{m,j} = i_j$ . The fitness components are not defined independently. The numbers  $(y_{x_i, x_{i_1}, \dots, x_{i_K}}^{1,i}, \dots, y_{x_i, x_{i_1}, \dots, x_{i_K}}^{M,i})$  follow a multivariate uniform law of dimension  $M$ , defined by a correlation matrix  $C$ . Thus, the  $y$ 's follow a multidimensional law with uniform marginals and the correlations between  $y_{\dots}^{m,i}$ 's are defined by the matrix  $C$ . So, the four parameters of the family of  $CMNK$ -landscapes are (i) the number of objective functions  $M$ , (ii) the length of the bit string  $N$ , (iii) the number of epistatic links  $K$ , and (iv) the correlation matrix  $C$ .

In the  $\rho MNK$ -landscapes, a matrix  $C_\rho = (c_{np})$  is considered, with the same correlation between all the objectives:  $c_{nn} = 1$  for all  $n$ , and  $c_{np} = \rho$  for all  $n \neq p$ . However, it is not possible to have the matrix  $C_\rho$  for all  $\rho$  between  $[-1, 1]$ :  $\rho$  must be greater than  $\frac{-1}{M-1}$ , see [9]. To generate random variables with uniform marginals and a specified correlation matrix  $C$ , we follow the work of Hotelling and Pabst [12]. The construction of  $CMNK$ -landscapes defines correlation between the  $y$ 's but not directly between the objectives. In [9], it is proven by algebra that the correlation between objectives is tuned by the matrix  $C$ :  $E(\text{cor}(f_n, f_p)) = c_{np}$ . In  $\rho MNK$ -landscape, the parameter  $\rho$  allows to tune very precisely the correlation between all pairs of objectives.

## 4 Study of Pareto Local Optima

In this section, we first study the number of Pareto local optima (PLO) according to the objective correlation, the number of objectives and the epistasis of  $\rho MNK$ -landscapes. Then, we analyze its relation with the size of the Pareto optimal set. At last, we propose an adaptive walk that is able to estimate the number of PLO very precisely. We conduct an empirical study for  $N = 18$  so that we can enumerate all the PLO exhaustively. In order to minimize the influence of

**Table 1.** Parameters used in the paper for the experimental analysis.

Parameter	Values
$N$	{18} (Section 4) , {18, 32, 64, 128} (Section 5)
$M$	{2, 3, 5}
$K$	{2, 4, 6, 8, 10}
$\rho$	{-0.9, -0.7, -0.4, -0.2, 0.0, 0.2, 0.4, 0.7, 0.9} such that $\rho \geq \frac{-1}{M-1}$

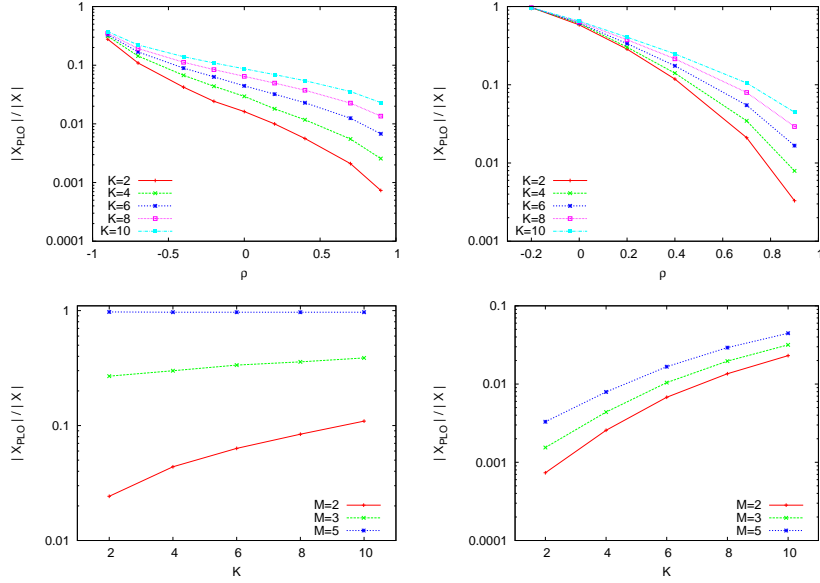
the random creation of landscapes, we considered 30 different and independent instances for each parameter combinations:  $\rho$ ,  $M$ , and  $K$ . The measures reported are the average over these 30 landscapes. The parameters under investigation in this study are given in Table 1.

#### 4.1 Number of Pareto Local Optima

Fig. 1 shows the average number of PLO to the size of the search space ( $|X| = 2^{18}$ ) for different  $\rho$ MNK-landscapes parameter settings. As the well-known result from single-objective  $NK$ -landscapes [1], the number of PLO increases with the epistasis degree. For instance, with an objective space dimension  $M = 2$  and an objective correlation  $\rho = 0.9$ , the average number of PLO increases more than 30 times: from 192 for  $K = 2$  to 6048 for  $K = 10$ . However, the range of PLO is larger with respect to objective correlation. For the same epistatic degree and number of objectives, the number of PLO decreases exponentially (Fig. 1, top). Indeed, for an objective space dimension  $M = 2$  and an epistasis degree  $K = 4$ , the average number of PLO decreases more than 120 times: from 82,093 for negative correlation ( $\rho = -0.9$ ) to 672 for positive correlation ( $\rho = 0.9$ ).

This result can be interpreted as follows. Let us consider an arbitrary solution  $x$ , and two different objective functions  $f_i$  and  $f_j$ . When the objective correlation is high, there is a high probability that  $f_i(x)$  is close to  $f_j(x)$ . In the same way, the fitness values  $f_i(x')$  and  $f_j(x')$  of a given neighbor  $x' \in \mathcal{N}(x)$  are probably close. So, for a given solution  $x$  such that it exists a neighbor  $x' \in \mathcal{N}(x)$  with a better  $f_i$ -value, the probability is high that  $f_j(x')$  is better than  $f_j(x)$ . More formally, the probability  $\mathbb{P}(f_j(x') > f_j(x) \mid f_i(x') > f_i(x))$ , with  $x' \in \mathcal{N}(x)$ , increases with the objective correlation. Then, a solution  $x$  has a higher probability of being dominated when the objective correlation is high. Under this hypothesis, the probability that a solution dominates all its neighbors decreases with the number of objectives. Fig. 1 (bottom) corroborates this hypothesis. When the objective correlation is negative ( $\rho = -0.2$ ), the number of PLO changes in an order of magnitude from  $M = 2$  to  $M = 3$ , and from  $M = 3$  to  $M = 5$ . This range is smaller when the correlation is positive. When the number of objective is large and the objective correlation is negative, almost all solutions are PLO.

Assuming that the difficulty for Pareto-based search approaches gets higher when the number of PLO is large, the difficulty of  $\rho MNK$ -landscapes increases when: (i) the epistasis increases, (ii) the number of objective functions increases,



**Fig. 1.** Average number of PLO to the size of the search space ( $|X| = 2^{18}$ ) according to parameter  $\rho$  (top left  $M = 2$ , right  $M = 5$ ), and to parameter  $K$  (bottom left  $\rho = -0.2$ , right  $\rho = 0.9$ ). The problem size is  $N = 18$ .

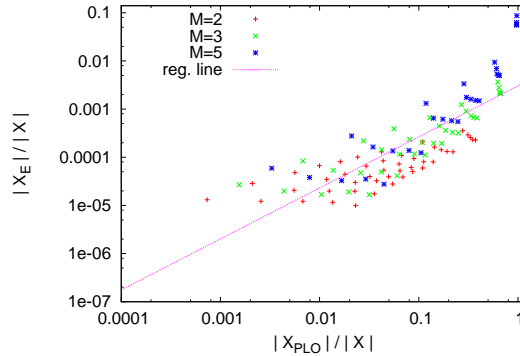
(iii) the objective correlation is negative, and its absolute value increases. Section 5 will precise the relative difficulty related to those parameters for large-size problem instances.

## 4.2 Estimating the Cardinality of the Pareto Optimal Set?

When the number of Pareto optimal solutions is too large, it becomes impossible to enumerate them all. A metaheuristic should then manipulate a limited-size solution set during the search. In this case, we have to design specific strategies to limit the size of the approximation set [13]. Hence, the cardinality of the Pareto optimal set also plays a major role in the design of multiobjective metaheuristics.

In order to design such an approach, it would be convenient to approximate the size of the Pareto optimal set from the number of PLO. Fig. 2 shows the scatter plot of the average size of the Pareto optimal set *vs.* the average number of PLO in log-scales. Points are scattered over the regression line with the Spearson correlation coefficient of 0.82, and the regression line equation is  $\log(y) = a \log(x) + b$  with  $a = 1.059$  and  $b = -6.536$ . For such a log-log scale, the correlation is low. It is only possible to estimate the cardinality of the Pareto optimal set from the number of PLO with a factor 10. Nevertheless, the number of Pareto optimal solutions clearly increases when the number of PLO increases.





**Fig. 2.** Scatter plot of the average size of the Pareto optimal set (to the size of the search) *vs.* the average number of PLO (to the size of the search) for the 110 possible combinations of parameters. The problem size is  $N = 18$ . The correlation coefficient is 0.82. Notice the log-scales.

### 4.3 Adaptive Walk

In single-objective optimization, the length of adaptive walks, performed with a hill-climber, allows to estimate the diameter of the local optima basins of attraction. Then, the number of local optima can be estimated when the whole search space cannot be enumerated exhaustively. In this section, we define a multiobjective hill-climber, and we show that the length of the corresponding adaptive walk is correlated to the number of PLO. We define a very basic single solution-based *Pareto Hill-Climbing* (PHC) for multiobjective optimization. A pseudo-code is given in Algorithm 1. At each iteration of the PHC algorithm, the current solution is replaced by one random neighbor solution which dominates it. So, the PHC stops on a PLO. The number of iterations, or steps, of the PHC algorithm is the length of the Pareto adaptive walk.

We performed  $10^3$  independent PHC executions for each problem instance. Fig. 3 shows the average length of the Pareto adaptive walks for different landscapes according to the set of parameters given in Table 1. The variation of the average length follows the opposite variation of the number of PLO. In order to

---

#### Algorithm 1 Pareto Hill-Climbing (PHC)

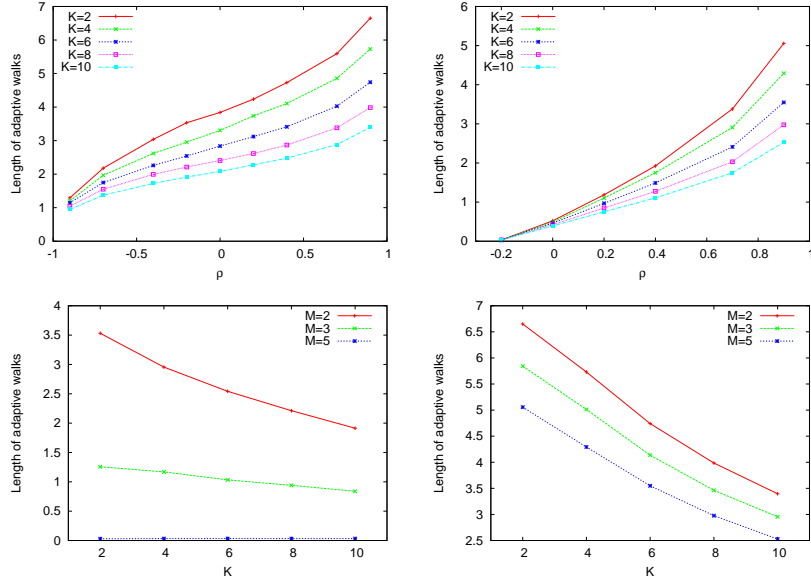
---

```

start with a random solution  $x \in X$ 
step  $\leftarrow 0$ 
while  $x$  is not a Pareto Local optimum do
    randomly choose  $x'$  from  $\{y \in \mathcal{N}(x) | x \prec y\}$ 
     $x \leftarrow x'$ 
    step  $\leftarrow$  step + 1
end while

```

---

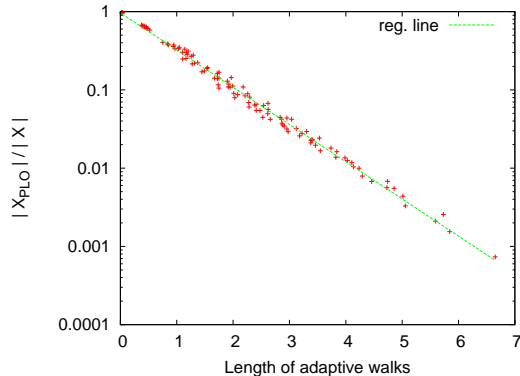


**Fig. 3.** Average length of the Pareto adaptive walk according to parameter  $\rho$  (top left  $M = 2$ , right  $M = 5$ ), and according to parameter  $K$  (bottom left  $\rho = -0.2$ , right  $\rho = 0.9$ ). The problem size is  $N = 18$ .

show the link with the number of PLO more clearly, Fig. 4 gives the scatter-plot of the average Pareto adaptive length *vs.* the logarithm of the average number of PLO. The correlation is strong ( $r = 0.997$ ), and the regression line equation is:  $\log(y) = ax + b$ , with  $a = -1.095$  and  $b = 12.443$ . For bit-string of length  $N = 18$ , the average length of the Pareto adaptive walks can then give a precise estimation of the average number of PLO. When the adaptive length is short, the diameter of the basin of attraction associated with a PLO is short. This means that the distance between PLO decreases. Moreover, assuming that the volume of this basin is proportional to a power of its diameter, the number of PLO increases exponentially when the adaptive length decreases. This corroborates known results from single-objective optimization. Of course, for larger bit-string length, the coefficients are probably different.

## 5 Properties *vs.* Multi-modality for Large-size Problems

In this section, we study the number of PLO for large-size  $\rho MNK$ -landscapes using the length of the adaptive walk proposed in the previous section. First, we analyze this number according to the problem dimension ( $N$ ). Then, we precise the difficulty, in terms of PLO, with respect to objective space dimension ( $M$ ) and objective correlation ( $\rho$ ).



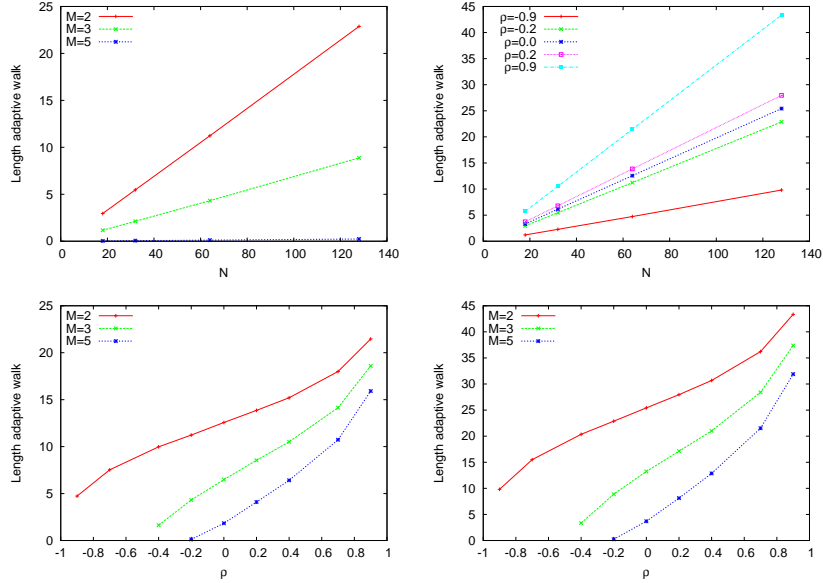
**Fig. 4.** Scatter plot of the average number of PLO to the size of the search space ( $|X| = 2^{18}$ ) vs. the average length of the Pareto adaptive walk. The problem size is  $N = 18$ .

We performed  $10^3$  independent PHC executions for each problem instance. Fig. 5 shows the average length of the Pareto adaptive walks for different landscapes according to the set of parameters given in Table 1. Whatever the objective space dimension and correlation, the length of the adaptive walks increases linearly with the search space dimension  $N$ . According to the results from the previous section, the number of PLO increases exponentially. We can then reasonably conclude that the size of the Pareto optimal set grows exponentially as well, to an order of magnitude (Section 4.2). However, the slope of the Pareto adaptive length increase is related to the objective space dimension ( $M$ ) and correlation ( $\rho$ ). The higher the number of objective functions, the smaller the slope. As well, the higher the objective correlation, the smaller the slope.

Fig. 5 (bottom) allows us to give a qualitative comparison for given problem sizes ( $N = 64$  and  $N = 128$ ). Indeed, let us consider an arbitrary adaptive walk of length 10. For  $\rho MNK$ -landscapes with  $N = 64$  and  $K = 4$ , this length corresponds approximately to parameters  $(\rho = -0.4, M = 2)$ ,  $(\rho = 0.3, M = 3)$ , and  $(\rho = 0.7, M = 5)$  at the same time. For  $N = 128$ , we have  $(\rho = -0.9, M = 2)$ ,  $(\rho = -0.1, M = 3)$ , and  $(\rho = 0.3, M = 5)$ . Still assuming that a problem difficulty is closely related to the number of PLO, an instance with a small objective space dimension and a negative objective correlation can be more difficult to solve than with many correlated objectives.

## 6 Discussion

This paper gives a fitness landscape analysis for multiobjective combinatorial optimization based on the local optima of multiobjective  $NK$ -landscapes with objective correlation. We first focused on small-size problems with a study of the number of local optima by complete enumeration. Like in single-objective optimization, the number of local optima increases with the degree of non-linearity



**Fig. 5.** Average length of the Pareto adaptive walk according to problem size ( $N$ ) for  $K = 4$  and  $\rho = -0.2$  (top-left) and for  $K = 4$  and  $M = 2$  (top-right). Average length of the Pareto adaptive walk according to objective correlation ( $\rho$ ) for  $K = 4$  and  $N = 64$  (bottom-left) and for  $K = 4$  and  $N = 128$  (bottom-right).

of the problem (epistasis). However, the number of objective functions and the objective correlation have a stronger influence. Furthermore, our results show that the cardinality of the Pareto optimal set clearly increases with the number of local optima. We proposed a Pareto adaptive walk, associated with a Pareto hill-climber, to estimate the number of local optima for a given problem size. Next, for large-size instances, the length of such Pareto adaptive walk can give a measure related to the difficulty of a multiobjective combinatorial optimization problem. We show that this measure increases exponentially with the problem size. A problem with a small number of negatively correlated objectives gives the same degree of multi-modality, in terms of Pareto dominance, than another problem with a high objective space dimension and a positive correlation.

A similar analysis would allow to better understand the structure of the landscape for other multiobjective combinatorial optimization problems. However, an appropriate model to estimate the number of local optima for any problem size still needs to be properly defined. A possible path is to generalize the approach from [14] for the multiobjective case. For a more practical purpose, our results should also be put in relation with the type of the problem under study, in particular on how to compute or estimate the problem-related measures reported in this paper. Moreover, we mainly focused our work on the number of local optima. The next step is to analyze their distribution by means of a local optima net-

work [15]. At last, we already know that the number and the distribution of local optima have a strong impact on the performance of multiobjective metaheuristics, but it is not yet clear how they exactly affect the search. This open issue constitutes one of the main challenge in the field of fitness landscape analysis for multiobjective combinatorial optimization.

## References

1. Kauffman, S.A.: *The Origins of Order*. Oxford University Press, New York, USA (1993)
2. Paquete, L., Schiavinotto, T., Stützle, T.: On local optima in multiobjective combinatorial optimization problems. *Ann Oper Res* **156**(1) (2007) 83–97
3. Borges, P., Hansen, M.: A basis for future successes in multiobjective combinatorial optimization. Technical Report IMM-REP-1998-8, Institute of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark (1998)
4. Paquete, L., Stützle, T.: Clusters of non-dominated solutions in multiobjective combinatorial optimization: An experimental analysis. In: *Multiobjective Programming and Goal Programming*. Volume 618 of *Lecture Notes in Economics and Mathematical Systems*. Springer (2009) 69–77
5. Knowles, J., Corne, D.: Towards landscape analyses to inform the design of a hybrid local search for the multiobjective quadratic assignment problem. In: *Soft Computing Systems: Design, Management and Applications*. (2002) 271–279
6. Garrett, D., Dasgupta, D.: Multiobjective landscape analysis and the generalized assignment problem. In: *Learning and Intelligent Optimization (LION 2)*. Volume 5313 of *Lecture Notes in Computer Science*. Springer, Trento, Italy (2007) 110–124
7. Garrett, D., Dasgupta, D.: Plateau connection structure and multiobjective metaheuristic performance. In: *Congress on Evolutionary Computation (CEC 2009)*, IEEE (2009) 1281–1288
8. Aguirre, H.E., Tanaka, K.: Working principles, behavior, and performance of MOEAs on MNK-landscapes. *Eur J Oper Res* **181**(3) (2007) 1670–1690
9. Verel, S., Liefoghe, A., Jourdan, L., Dhaenens, C.: Analyzing the effect of objective correlation on the efficient set of MNK-landscapes. In: *Learning and Intelligent Optimization (LION 5)*. *Lecture Notes in Computer Science*. Springer, Rome, Italy (2011) to appear.
10. Coello Coello, C.A., Dhaenens, C., Jourdan, L., eds.: *Advances in Multi-Objective Nature Inspired Computing*. Volume 272 of *Studies in Computational Intelligence*. Springer (2010)
11. Merz, P.: Advanced fitness landscape analysis and the performance of memetic algorithms. *Evol Comput* **12**(3) (2004) 303–325
12. Hotelling, H., Pabst, M.R.: Rank correlation and tests of significance involving no assumptions of normality. *Ann Math Stat* **7** (1936) 29–43
13. Knowles, J., Corne, D.: Bounded Pareto archiving: Theory and practice. In: *Metaheuristics for Multiobjective Optimisation*. Volume 535 of *Lecture Notes in Economics and Mathematical Systems*. Springer (2004) 39–64
14. Eremeev, A.V., Reeves, C.R.: On confidence intervals for the number of local optima. In: *Applications of Evolutionary Computing (EvoWorkshop 2003)*. Volume 2611 of *Lecture Notes in Computer Science*. Springer, Essex, UK (2003) 224–235
15. Daolio, F., Verel, S., Ochoa, G., Tomassini, M.: Local optima networks of the quadratic assignment problem. In: *Congress on Evolutionary Computation (CEC 2010)*, IEEE (2010) 1–8