Local Optimal Sets and Bounded Archiving on Multi-objective NK-Landscapes with Correlated Objectives

Manuel López-Ibáñez[1], Arnaud Liefooghe[2], and Sébastien Verel[3]

[1] IRIDIA, Université Libre de Bruxelles (ULB), Brussels, Belgium
manuel.lopez-ibanez@ulb.ac.be
[2] Université Lille 1, LIFL, UMR CNRS 8022, Inria Lille-Nord Europe, France
arnaud.liefooghe@univ-lille1.fr
[3] Université du Littoral Côte d'Opale, LISIC, France
verel@lisic.univ-littoral.fr

Abstract. The properties of local optimal solutions in multi-objective combinatorial optimization problems are crucial for the effectiveness of local search algorithms, particularly when these algorithms are based on Pareto dominance. Such local search algorithms typically return a set of mutually nondominated Pareto local optimal (PLO) solutions, that is, a PLO-set. This paper investigates two aspects of PLO-sets by means of experiments with Pareto local search (PLS). First, we examine the impact of several problem characteristics on the properties of PLO-sets for multi-objective NK-landscapes with correlated objectives. In particular, we report that either increasing the number of objectives or decreasing the correlation between objectives leads to an exponential increment on the size of PLO-sets, whereas the variable correlation has only a minor effect. Second, we study the running time and the quality reached when using bounding archiving methods to limit the size of the archive handled by PLS, and thus, the maximum size of the PLO-set found. We argue that there is a clear relationship between the running time of PLS and the difficulty of a problem instance.

1 Introduction

Several state-of-the-art algorithms for multi-objective combinatorial optimization problems (MCOPs) are based on local search. These local search algorithms are either based on solving multiple scalarizations of the objective function vector, or they are based on Pareto dominance. The most successful local search algorithms combine both approaches [3,4,10,13]. These successes explain the increasing interest on understanding the role of local optimal solutions in the context of MCOPs. Previous works have focused on the properties of individual local optimal solutions [14]. However, algorithms for MCOPs typically return not a single solution, but a set of solutions that approximates the Pareto set. Thus, local search algorithms for MCOPs are typically concerned by (Pareto) local optimal *sets* (PLO-sets), that is, sets of (Pareto) local optimal solutions where solutions are mutually nondominated and are also local optimal with respect to the neighborhood of the other solutions in the set [12].

Experimental studies on PLO-sets have been so far limited to the study of some of their properties for the bi-objective traveling salesman problem, in particular the number of solutions in each PLO-set and the connectedness of PLO-sets [11]. This paper extends significantly this initial work in two aspects. First, we consider multi-objective NK-landscapes with correlated objectives (ρMNK-landscapes) [14], which allow us to examine the effect of various problem characteristics on the properties of PLO-sets. Second, we examine the effect of using bounded archiving methods [7,9] in order to limit the size of the PLO-set handled by the local search. Archiving methods are often used in both local search and evolutionary multi-objective algorithms, and there are a few recent works on their theoretical properties [2,8,9]. Several authors have mentioned as interesting future work the experimental study of the PLO-sets induced by bounded archiving. To the best of our knowledge, we present in this paper the first results of such a study. Our conclusions not only give support to previous theoretical results, but also give insights on how to improve local search algorithms for particularly difficult MCOPs. In the following, we recall the required background on problems and algorithms in Section 2; we provide the experimental setup of our study in Section 3; we discuss our experimental results in Section 4; and we conclude in the last section.

2 Background

Multi-objective Combinatorial Optimization. A *multi-objective combinatorial optimization problem* (MCOP) is defined by an objective function vector $f = (f_1, \ldots, f_m)$ with $m \geq 2$ objective functions, and a discrete set X of feasible solutions in the *solution space*. Let $Z = f(X) \subseteq \mathbb{R}^m$ be the set of feasible outcome vectors in the *objective space*. Each solution $x \in X$ is assigned an objective vector $z \in Z$ on the basis of the multidimensional function vector $f\colon X \rightarrow Z$ such that $z = f(x)$. When all objectives are to be maximized, a solution x weakly dominates another solution x' if $\forall i \in \{1, \ldots, m\}$, $f_i(x) \geq f_i(x')$. If, in addition, $\exists i \in \{1, \ldots, m\}$ such that $f_i(x) > f_i(x')$, then we say that x dominates x' ($x \prec x'$). When $x \nprec x' \wedge x' \nprec x$, we say that x and x' are mutually nondominated. A solution $x^\star \in X$ is *Pareto optimal* if $\nexists x \in X$ such that $x \prec x^\star$. The set of all Pareto optimal solutions is the *Pareto set*. Its mapping in the objective space is the *Pareto front*. One of the goals in multi-objective optimization is to identify the Pareto set, or a good approximation of it.

Pareto Local Optimal Sets. Set-based local search algorithms for MCOPs generally combine the use of a neighborhood operator with the management of an archive of mutually nondominated solutions found so far. A *neighborhood operator* is a mapping function $\mathcal{N}\colon X \rightarrow 2^X$ that assigns a set of solutions $\mathcal{N}(x) \subset X$ to any solution $x \in X$. $\mathcal{N}(x)$ is called the *neighborhood* of x, and a solution $x' \in \mathcal{N}(x)$ is called a *neighbor* of x. A solution $x \in X$ is a *Pareto* *local optimum (PLO)* with respect to a neighborhood structure $\mathcal{N}$ if there is no neighbor $x' \in \mathcal{N}(x)$ such that $x' \prec x$ [12]. A set $S \subseteq X$ is a *Pareto local optimal set (PLO-set)* with respect to $\mathcal{N}$ if, and only if, it contains only PLO-solutions with respect to $\mathcal{N}$ and all solutions are mutually nondominated [12]. In addition, a set $S \subseteq X$ is a *maximal PLO-set* with respect to $\mathcal{N}$ if, and only if, $\forall s' \in \mathcal{N}(S)$, $\exists s \in S$ such that $s \prec s' \vee f(s) = f(s')$, where $\mathcal{N}(S) = \bigcup_{s \in S} \mathcal{N}(s)$ [12]. In other words, any neighbor of any solution in a maximal PLO-set is weakly dominated by a solution in the set.

Pareto Local Search. A typical example of a multi-objective local search algorithm is Pareto Local Search (PLS) [12]. PLS is an extension of the conventional hill-climbing algorithm to the multi-objective case. An archive of nondominated solutions is initialized with at least one solution. At each iteration, one solution is chosen at random from the archive and all its neighbors are evaluated and compared against the archive. Each neighbor is added to the archive, and marked as *unvisited*, if it is not dominated by any other solution in the archive. Moreover, solutions in the archive dominated by this neighbor are removed. Once all neighbors have been evaluated, the current solution is marked as *visited*. The algorithm stops once all solutions from the archive are marked as *visited*.

PLS is known to be a well-performing algorithm, either as a stand-alone approach or as a hybrid component, for many MCOPs [3,4,10,13]. Moreover, independently of the initial archive, PLS always terminates and returns a maximal PLO-set [12]. However, the archive of (unvisited) solutions may grow exponentially with respect to the instance size and, in that case, PLS may require an exponential number of iterations. In such a situation, it would be more interesting to bound the size of the archive in order to prevent an exponential grow, but still return a (perhaps non-maximal) PLO-set.

Given an archive A and a maximum size μ, a bounded archiving method, or *archiver* for short, will return a new archive $A' \subseteq A$ such that $|A'| \leq \mu$ [7,9]. We can use an archiving method in PLS such that whenever a new solution x' is added to the archive A and $|A \cup \{x'\}| = \mu + 1$, then the archiving method will select one solution to be removed. This is equivalent to the $\mu + \lambda$ strategy [2], with $\lambda = 1$. The various archiving methods differ on how the solution to be removed is selected. Here we focus on two archiving methods, hypervolume archiver (HVA) [6] and multi-level grid archiver (MGA) [8], which are the only known archiving methods belonging to the class with the most desirable convergence properties [9]. Two of these properties are: *(i)* accepting solutions outside the objective space region dominating the current archive (*diversifies*) and *(ii)* a subsequent archive cannot be worse in terms of Pareto dominance than an earlier archive ($\lhd$-*monotone*).

When PLS uses either HVA or MGA as its archiving method, then a run of PLS will stop at a PLO-set, but not necessarily at a maximal PLO-set. Indeed, there may exist a solution $s' \in \mathcal{N}(A)$, such that $\nexists s \in A$ with $s \prec s'$, but the archiving method chose to discard s' in order to maintain $|A| \leq \mu$. Therefore, it is expected that when using an archiving method, PLS will converge faster but to a possibly worse PLO-set. Moreover, since the decision of which solutions are discarded are fundamentally different for HVA and MGA, we would expect that each method may converge to disjoint sets of PLO-sets. In the experiments presented here, we study the properties of the PLO-sets returned by the classical PLS (using an unbounded archive) and when PLS uses either HVA or MGA to bound the archive size.

Multi-Objective NK-landscapes with Correlated Objectives. We study the effect of various characteristics of MCOPs on the properties of PLO-sets by means of ρMNK-landscapes, which are artificial multi-objective multimodal problems with objective correlation [14]. They extend both single-objective NK-landscapes [5] and multi-objective NK-landscapes with independent objective functions [1]. Feasible solutions are binary strings of size n. The parameter k refers to the number of variables that influence a particular position from the bit-string (the epistatic interactions). The objective function vector is defined as $f\colon \{0,1\}^n \rightarrow [0,1)^m$. Each objective function is to be maximized, and can be formalized as follows: $f_i(x) = \frac{1}{n}\sum_{j=1}^{n} c_j^i(x_j, x_{j_1}, \ldots, x_{j_k})$, $i \in \{1,\ldots,m\}$, where $c_j^i\colon \{0,1\}^{k+1} \rightarrow [0,1)$ defines the component function associated with each variable x_j, $j \in \{1,\ldots,n\}$, for objective f_i, and where $k < n$. By increasing the number of variable interactions k from 0 to $(n-1)$, ρMNK-landscapes can be gradually tuned from smooth to rugged.

We generate an instance of a ρMNK-landscape by randomly setting the position of these epistatic interactions, following a uniform distribution. The same epistatic degree k and the same epistatic interactions are used for all the objectives. Component function values are sampled within the range $[0,1)$ following a multivariate uniform distribution of dimension m with a correlation coefficient ρ. A positive (resp. negative) correlation coefficient decreases (resp. increases) the degree of conflict between the objective function values.

3 Experimental Setup

In the following, we investigate ρMNK-landscapes with a problem size $n \in \{8,16\}$, an epistatic degree $k \in \{1,2,4,8\}$ such that $k < n$, an objective space dimension $m \in \{2,3,5\}$, and an objective correlation $\rho \in \{-0.7,-0.2,0.0,0.2,0.7\}$ such that $\rho > \frac{-1}{m-1}$, because the corresponding correlation matrix is symmetric positive-definite [14]. The investigated problem sizes allow us to enumerate the solution space exhaustively, and then to solve all the instances to optimality. One independent random instance is considered for each parameter combination: $\langle \rho, m, n, k \rangle$. This leads to a total of 91 problem instances.

In our implementation of PLS for ρMNK-landscapes, the neighborhood structure is taken as the *1-bit-flip*. The archive is initialized with one random solution from the solution space. At each iteration, the neighborhood of the selected solution is explored exhaustively in a random order. This order has an impact on the dynamics of PLS since bounded archiving methods treat incoming solutions sequentially. The cost of each iteration of PLS is exactly n evaluations, corresponding to the neighborhood size. We experiment with three variants of PLS. PLS_{UNB} corresponds to the classical PLS with an unbounded archive. PLS_{HVA} and PLS_{MGA} use HVA and MGA, respectively, to bound the size of the archive to a maximum of μ solutions, where $\mu \in \{10, 20, 40, 80\}$. We consider 25 different seeds for the random number generator used in PLS, and we run each PLS variant on each problem instance using each random seed. This leads to a total of 20 475 runs.

4 Experimental Analysis

4.1 Cardinality of Pareto Local Optimal Sets

Figure 1 shows the size of the PLO-sets identified by PLS_{UNB} with respect to different instance characteristics. Each point gives the mean value over the 25 random seeds on the same instance, and the error bars indicate the standard deviation. PLO-set sizes are plotted in logarithmic scale.

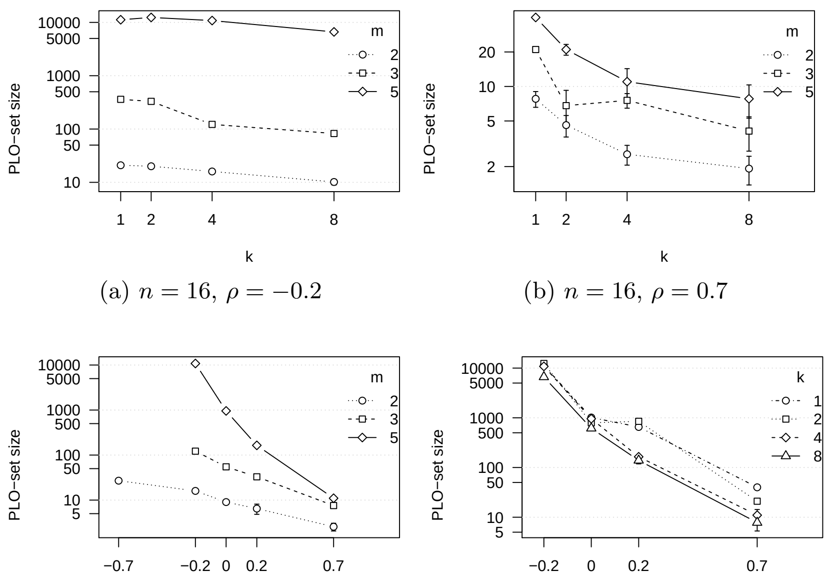
In general, the cardinality of PLO-sets increases exponentially with the number of objectives m, as shown in Figs. 1a–1c. The correlation between objective values ρ is also a crucial factor: the cardinality of the PLO-sets increases exponentially with the linear decrease of ρ (Fig 1c). The cardinality of PLO-sets also increases with lower k, but much less noticeably (Fig. 1a). Moreover, the small error bars in Figs. 1a–1d indicate that, for a given instance, maximal PLO-sets consistently have roughly the same size.

Given the above results, restricting the size of the PLO-sets by using bounding archiving methods must have a stronger effect as the correlation decreases and the number of objectives increases. In the next section, we examine this effect for each archiving method.

4.2 Quality of Local Optimal Sets

We examine the quality of the PLO-sets found by PLS_{UNB}, PLS_{HVA} and PLS_{MGA} in terms of two unary quality measures, namely, the hypervolume and the multiplicative epsilon [15]. In order to compare the hypervolume value for instances with very different characteristics, we compute the hypervolume relative difference as $hvr(A) = (hv(P) - hv(A))/hv(P)$, where $A \subseteq Z$ is the image of a PLO-set in the objective space and P is the exact Pareto front for the instance under consideration. The reference point is set to the origin. The epsilon measure gives the minimum multiplicative factor by which a PLO-set has to be shifted in the objective space to weakly dominate the exact Pareto front. Thus, for both measures, a lower value is preferred. Results are reported in Fig. 2 for different parameter settings.

As conjectured above, a first observation is that, as the size of maximal PLO-sets increases, the quality of PLO-sets obtained by bounded archiving decreases. The increase in quality is almost logarithmic with respect to the archive size



(a) $n = 16$, $\rho = -0.2$

(b) $n = 16$, $\rho = 0.7$

(c) $n = 16$, $k = 4$

(d) $n = 16$, $m = 5$

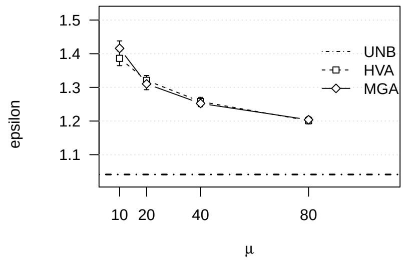
Fig. 1. Mean size of the PLO-sets returned by PLS_{UNB}. Error bars give the standard deviation.

limit (μ), as shown in Figs. 2a–2d. Here we show only results for $m = 5$, but the trends are similar for a lower number of objectives, although less pronounced. This trend appears independently of whether we measure quality in terms of epsilon or hypervolume. Moreover, as in the single-objective case, the average quality of local optima decreases with k [5], as shown in Figs. 2e–2f.

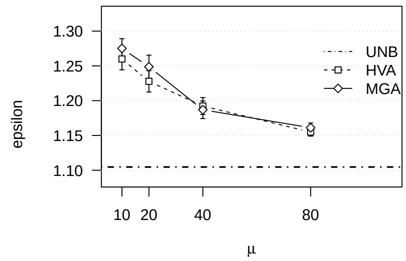
Lastly, there is not much difference in terms of quality between PLS_{HVA} and PLS_{MGA}. When differences occur, the PLO-sets returned by PLS_{HVA} have a better hypervolume (lower hvr value) than those returned by PLS_{MGA} (Figs. 2c–2d), however, there is no clear winner in terms of epsilon value (Fig. 2a–2b).

4.3 Difficulty of Identifying Local Optimal Sets

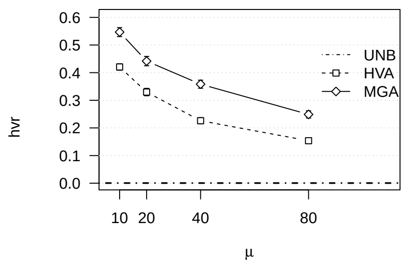
For single-objective NK-landscapes, the number of iterations, or steps, of a conventional hill-climbing algorithm provides an estimation of the average diameter of the basins of attraction of local optima [5]. This diameter characterizes a problem instance in terms of multimodality: The larger the length, the larger the basin diameter and the lower the number of local optima. Conversely, the smaller the length, the smaller the basin diameter and the higher the number of local optima. Multimodality characterizes an important aspect of instance difficulty (i.e., the number of local optima).



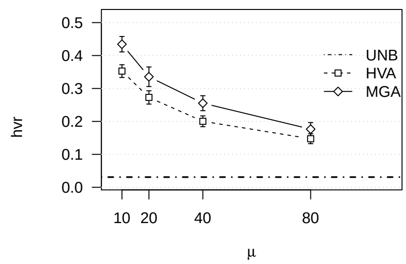
(a) $n = 16,\ m = 5,\ k = 8,\ \rho = -0.2$



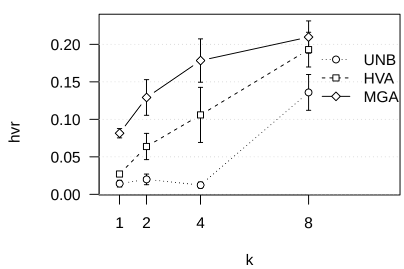
(b) $n = 16,\ m = 5,\ k = 8,\ \rho = 0$



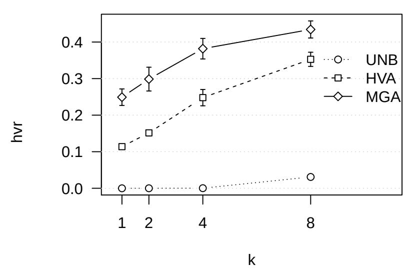
(c) $n = 16,\ m = 5,\ k = 8,\ \rho = -0.2$



(d) $n = 16,\ m = 5,\ k = 8,\ \rho = 0$



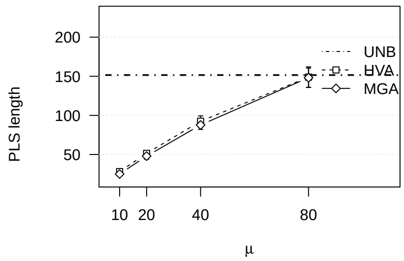
(e) $n = 16,\ m = 3,\ \rho = 0.0, \mu = 10$



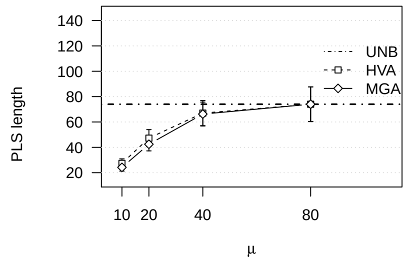
(f) $n = 16,\ m = 5,\ \rho = 0.0, \mu = 10$

Fig. 2. Mean quality of the PLO-sets found by PLS_{UNB}, PLS_{HVA} and PLS_{MGA} measured in terms of multiplicative epsilon (*epsilon*) and hypervolume relative difference (*hvr*). Error bars give the standard deviation.

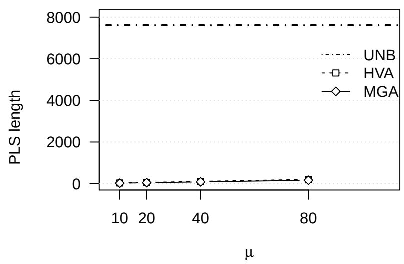
As in the single-objective case, the construction of ρMNK-landscapes implies that they are isotropic, that the neighborhood has the same properties in every direction of the objective space, and that the basins of attraction have a ball-like shape. In this section, we report the length of PLS for different archiving strategies. This allows us to study the running time, in terms of number of iterations, required by PLS to identify a PLO-set. Moreover, when the PLS length is smaller under one setting than another, we can reasonably assume that there exist more PLO-sets in the corresponding landscape since the search process got stuck more easily on a local optima. Figure 3 shows the PLS length



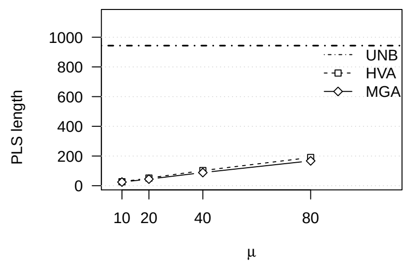
(a) $n = 16,\ m = 3,\ k = 8,\ \rho = -0.2$



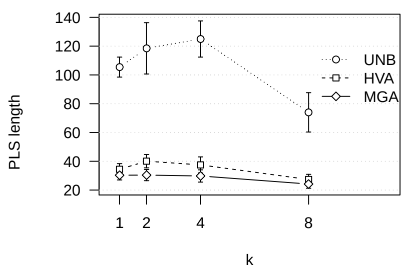
(b) $n = 16,\ m = 3,\ k = 8,\ \rho = 0$



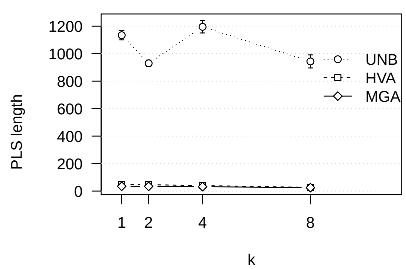
(c) $n = 16,\ m = 5,\ k = 8,\ \rho = -0.2$



(d) $n = 16,\ m = 5,\ k = 8,\ \rho = 0$



(e) $n = 16,\ m = 3,\ \rho = 0.0, \mu = 10$



(f) $n = 16,\ m = 5,\ \rho = 0.0, \mu = 10$

Fig. 3. Mean length of PLS_{UNB}, PLS_{HVA} and PLS_{MGA}. Error bars give the standard deviation.

for the same problem instances shown in Fig. 2. The comparison between each pair of plots shows clearly a relationship between the length of PLS (Fig. 3) and the quality of the results (Fig. 2): Larger length corresponds to better quality in general. Interestingly, bounding the archive size substantially reduces both the quality of the obtained approximation as well as the running time of PLS. The smaller the archive size, the larger the difference with PLS_{UNB}. However, this also has the effect of increasing the number of PLO-sets. Indeed, when the archive size is small, the PLS length is small, which suggests that the average distance between a local optimum and the solutions from the corresponding basin of attraction is also small and, hence, the number of PLO-sets is large.

Surprisingly, when the archive is unbounded as reported in Fig. 1, the PLS length increases from $k = 1$ to $k = 4$ (despite we know that the number of PLO-solutions increases linearly with k [14]) and only becomes smaller for $k = 8$ (Fig. 3). This increase in PLS length contradicts known results from single-objective optimization [5]. In fact, in the case of PLO-sets, both the number of PLO-solutions and the average size of the neighborhood of a PLO-set influence the number of PLO-sets. First, when the number of PLO-solutions increases, it is expected that the number of PLO-sets also increases, which will decrease the PLS length. By contrast, when a PLO-set is larger, the number of neighbor solutions in this set is larger as well. This potentially reduces the number of PLO-sets, making PLS run longer. The balance between these two effects could explain the PLS length; i.e., the PLS length starts to decrease when the number of PLO-solutions has a larger impact on the number of PLO-sets than the size of PLO-sets. Overall, these results suggest that the length of PLS could provide an estimation of the number of PLO-sets, and thus a measure of difficulty for archive-based local search.

5 Conclusions

In this paper, we analyzed the characteristics of local optima in set-based multi-objective local search when applied to multi-objective NK-landscapes with correlated objectives. First, the main factors affecting the cardinality of the maximal PLO-sets returned by PLS_{UNB} are the number of objectives and the correlation between them. By changing these two factors, the PLO-set size can vary from a few tens to tens of thousands. Our results confirm trends already noticed for other MCOP problems. In particular, the exponential increase in the PLO-set size with lower objective correlation has already been reported for the bi-objective QAP [13]. Another interesting observation is that, given a particular instance, the variability of PLO-set sizes is usually a very small fraction of the average size, that is, most maximal PLO-sets for a given instance have roughly the same size. This is not an obvious conjecture to make, and we currently do not know if it is also the case for other MCOPs. Our experiments also strongly indicate that the relationship between local search length and number of local optima, which is well-studied in the single-objective case [5] and in the case of PLO-solutions [14], also applies to PLS and PLO-sets. Our results clearly show that shorter PLS lengths typically correspond to lower quality results (and hence, more difficult instances). A precise estimation of this relationship in the case of PLO-sets would require to determine the exact number of PLO-sets for a given instance.

From an algorithm design point of view, this work helps to better capture the relation between running time and approximation quality according to the problem instance characteristics. From a theoretical point of view, it would be interesting to understand precisely the relationships between the PLO-sets obtained by each archiving method. Moreover, it is clear to us that there is a direct relationship between the PLO-solutions of a problem, and the number and size of PLO-sets, however, a precise formulation remains to be described. Finally, we left for future work discussing the implications of the results reported here with respect to theoretical bounds reported in the literature [2]. Finally, complementary studies on other MCOPs and larger problem instances would allow us to better understand the structure of PLO-sets for different archiving techniques.

Acknowledgments. Manuel López-Ibáñez acknowledges support from the Belgian F.R.S.-FNRS, of which he is a postdoctoral researcher.

References

1. Aguirre, H.E., Tanaka, K.: Working principles, behavior, and performance of MOEAs on MNK-landscapes. Eur. J. Oper. Res. 181(3), 1670–1690 (2007)
2. Bringmann, K., Friedrich, T.: Convergence of hypervolume-based archiving algorithms I: Effectiveness. In: Krasnogor, N., et al. (eds.) Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2011), pp. 745–752. ACM Press, New York, NY (2011)
3. Drugan, M.M., Thierens, D.: Stochastic Pareto local search: Pareto neighbourhood exploration and perturbation strategies. J. Heuristics 18(5), 727–766 (2012)
4. Dubois-Lacoste, J., López-Ibáñez, M., Stützle, T.: A hybrid TP+PLS algorithm for bi-objective flow-shop scheduling problems. Comput. Oper. Res. 38(8), 1219–1236 (2011)
5. Kauffman, S.A.: The Origins of Order. Oxford University Press (1993)
6. Knowles, J.D.: Local-Search and Hybrid Evolutionary Algorithms for Pareto Optimization. Ph.D. thesis, University of Reading, UK (2002)
7. Knowles, J.D., Corne, D.: Bounded Pareto archiving: Theory and practice. In: Gandibleux, X., et al. (eds.) Metaheuristics for Multiobjective Optimisation, pp. 39–64. LNEMS, Springer, Berlin, Germany (2004)
8. Laumanns, M., Zenklusen, R.: Stochastic convergence of random search methods to fixed size Pareto front approximations. Eur. J. Oper. Res. 213(2), 414–421 (2011)
9. López-Ibáñez, M., Knowles, J.D., Laumanns, M.: On sequential online archiving of objective vectors. In: Takahashi, R.H.C., et al. (eds.) Evolutionary Multi-criterion Optimization (EMO 2011), LNCS, vol. 6576, pp. 46–60. Springer (2011)
10. Lust, T., Teghem, J.: Two-phase Pareto local search for the biobjective traveling salesman problem. J. Heuristics 16(3), 475–510 (2010)
11. Paquete, L., Chiarandini, M., Stützle, T.: Pareto local optimum sets in the biobjective traveling salesman problem: An experimental study. In: Gandibleux, X., et al. (eds.) Metaheuristics for Multiobjective Optimisation, pp. 177–200. LNEMS, Springer, Berlin, Germany (2004)
12. Paquete, L., Schiavinotto, T., Stützle, T.: On local optima in multiobjective combinatorial optimization problems. Annals of Operations Research 156, 83–97 (2007)
13. Paquete, L., Stützle, T.: A study of stochastic local search algorithms for the biobjective QAP with correlated flow matrices. Eur. J. Oper. Res. 169(3), 943–959 (2006)
14. Verel, S., Liefooghe, A., Jourdan, L., Dhaenens, C.: On the structure of multiobjective combinatorial search space: MNK-landscapes with correlated objectives. Eur. J. Oper. Res. 227(2), 331–342 (2013)
15. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V.: Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans. Evol. Comput. 7(2), 117–132 (2003)