Dominance, epsilon, and hypervolume local optimal sets in multi-objective optimization, and how to tell the difference

Arnaud Liefooghe, Manuel López-Ibáñez, Luís Paquete, Sébastien Verel

Dominance, Epsilon, and Hypervolume Local Optimal Sets in Multi-objective Optimization, and How to Tell the Difference

Arnaud Liefooghe
Univ. Lille, CNRS, UMR 9189 – CRIStAL
Inria Lille – Nord Europe
F-59000 Lille, France
arnaud.liefooghe@univ-lille1.fr

Manuel López-Ibáñez
Alliance Manchester Business School
University of Manchester
Manchester, UK
manuel.lopez-ibanez@manchester.ac.uk

Luís Paquete
University of Coimbra, DEI, CISUC
Coimbra, Portugal
paquete@dei.uc.pt

Sébastien Verel
Univ. Littoral Côte d'Opale, LISIC
F-62100 Calais, France
verel@lisic.univ-littoral.fr

ABSTRACT

Local search algorithms have shown good performance for several multi-objective combinatorial optimization problems. These approaches naturally stop at a local optimal set (LO-set) under given definitions of neighborhood and preference relation among subsets of solutions, such as set-based dominance relation, hypervolume or epsilon indicator. It is an open question how LO-sets under different set preference relations relate to each other. This paper reports an in-depth experimental analysis on multi-objective *nk*-landscapes. Our results reveal that, whatever the preference relation, the number of LO-sets typically increases with the problem non-linearity, and decreases with the number of objectives. We observe that strict LO-sets of bounded cardinality under set-dominance are LO-sets under both epsilon and hypervolume, and that LO-sets under hypervolume are LO-sets under set-dominance, whereas LO-sets under epsilon are not. Nonetheless, LO-sets under set-dominance are more similar to LO-sets under epsilon than under hypervolume. These findings have important implications for multi-objective local search. For instance, a dominance-based approach with bounded archive gets more easily trapped and might experience difficulty to identify an LO-set under epsilon or hypervolume. On the contrary, a hypervolume-based approach is expected to perform more steps before converging to better approximations.

CCS CONCEPTS

• **Theory of computation → Design and analysis of algorithms**; *Randomized local search*; • **Applied computing** → Multi-criterion optimization and decision-making;

KEYWORDS

Multi-objective combinatorial optimization, Local search, Local optima, Set-based multi-objective optimization, Quality Indicators.

ACM Reference Format:
Arnaud Liefooghe, Manuel López-Ibáñez, Luís Paquete, and Sébastien Verel. 2018. Dominance, Epsilon, and Hypervolume Local Optimal Sets in Multi-objective Optimization, and How to Tell the Difference. In *GECCO '18: Genetic and Evolutionary Computation Conference, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3205455.3205572

GECCO '18, July 15–19, 2018, Kyoto, Japan
2018. ACM ISBN 978-1-4503-5618-3/18/07.
https://doi.org/10.1145/3205455.3205572

1 INTRODUCTION

Local search methods operate over a search landscape defined by a triplet $(S, \preccurlyeq, N)$, where S denotes a finite, or countably infinite, set of solutions (the *search space*), $\preccurlyeq$ is a preorder on S (the *preference relation*) and N is a mapping $N\colon S \mapsto 2^S$ (the *neighborhood relation*). For any pair of solutions $s, s' \in S$, $s \preccurlyeq s'$ denotes that solution s is at least as preferred as solution s'. For a given solution $s \in S$, the set $N(s)$ is the *neighborhood* of s and an element $s' \in N(s)$ is a *neighbor* of s. The most basic local search algorithm, commonly known as *hill-climbing*, starts from an initial solution $s \in S$, and iteratively improves the current solution by exploring its neighborhood and moving to an improving neighboring solution. When no improving neighbor is available, the algorithm is trapped in a *local optimum*. Hence, based on the triplet $(S, \preccurlyeq, N)$, we can define the notions of local optimum (LO) and strict local optimum (sLO).

Although the concepts of (strict) LO are well-studied in single-objective optimization, their extension and properties in a multi-objective context are much less understood. The difficulties arise from the fact that the search space S is actually the set of all mutually nondominated sets of feasible solutions, possibly bounded in size, thus the neighborhood can be seen as operating on sets. Moreover, the preference relation is usually defined in terms of Pareto dominance, but it may also be any other quality indicator that induces a preorder, such as the hypervolume. The implications of these different aspects are still open to investigation and may guide the design of new multi-objective algorithms, since even global search methods, such as multi-objective evolutionary algorithms, can be seen as iteratively identifying better-quality local optima, without any guarantee of finding a *global optimum*. Indeed, a global optimum is also a local optimum for any neighborhood relation.

Paquete et al. [8] provided definitions of local optimality with respect to solution- and set-dominance, and related them to the convergence point of multi-objective local search. Verel et al. [9] introduced a set-based fitness landscape and measured ruggedness and non-linearity for fixed-size sets of solutions, using the hypervolume as the preference relation. That study was later extended to the quality of LO-sets and the convergence profile of hypervolume-based local search under different notions of set neighborhoods [2].

More recently, López-Ibáñez et al. [7] showed that the size of sets that are LO with respect to dominance (Pareto local optimum sets or PLO-sets) is exponentially correlated with the number of objectives or with their degree of conflict, while variable interactions have a minor effect. In addition, they showed that the estimated number of PLO-sets is also correlated with the number of objectives or their degree of conflict. It is also known that the number of PLO-solutions increases linearly with the problem non-linearity [10], but that the number of unbounded PLO-sets decreases [7]. Finally, the use of bounded size archives [6] does not change these trends, but increases the number of bounded PLO-sets significantly by a factor that depends on the size of the unbounded PLO-sets. However, so far, no work has examined how various definitions of local optimality relate to each other. In this paper, we extend previous work on LO-sets by considering various types of local optima, induced by different set preference relations (dominance, epsilon, hypervolume), and by analyzing their properties.

The paper is organized as follows. Section 2 defines the concept of (strict) LO-sets in a way the matches the usual definition of LO solutions in the single-objective case, but allows the use of set preference relations based on dominance or quality metrics. In addition, we describe an adaptive walk for sampling such LO-sets. Section 3 describes the multi-objective *nk*-landscapes and the experimental setup used for our experimental study. Section 4 describes the experiments carried out in this paper and the conclusions that can be extracted from them. Finally, we summarize our main findings and list remaining open questions in Section 5.

2 LOCAL OPTIMAL SETS

Multi-objective Optimization. Let us assume that we are given an optimization problem characterized by a pair (X, f), where X is the set of feasible solutions (the *decision space*) and f is the *objective function* $f : X \mapsto \mathbb{R}^m$, to be maximized. In multi-objective optimization ($m > 1$), one is often interested in finding more than one optimal solution. Given two solutions $x, x' \in X$, we say that x *weakly dominates* x' ($x \preccurlyeq_{\mathsf{dom}} x'$) if $f_i(x') \leqslant f_i(x)$ for all $i \in \{1, \ldots, m\}$. In terms of Pareto optimality, the goal is to find a set $X^\star \subseteq X$ for which there exists no solution $x \in X$ such that $x \preccurlyeq_{\mathsf{dom}} x^\star$ for all $x^\star \in X^\star$. The set $X^\star$ is the *Pareto set*, and its image in the objective space is the *Pareto front*.

Set Preference Relations. In set-based multi-objective optimization [12], the search space can be defined as the collection of sets of feasible solutions (feasible sets): $\Sigma \subset 2^X$. We restrict to sets of mutually nondominated solutions and we consider that the cardinality of the sets is bounded by $\mu \in \mathbb{N}^+$, that is, we define $\Sigma := \{A \in 2^X : |A| \leqslant \mu \ \wedge \ \forall x, x' \in A, x \neq x' \Rightarrow \neg(x \preccurlyeq_{\mathsf{dom}} x')\}$. Notice, however, that if the Pareto set $X^\star$ is larger than μ, then $X^\star \notin \Sigma$.

The aforementioned dominance relation among solutions can naturally be extended to sets. Given two sets $A, B \in \Sigma$, A *weakly dominates* B ($A \preccurlyeq_{\mathsf{dom}} B$), if for all $b \in B$ there exists an $a \in A$ such that $a \preccurlyeq_{\mathsf{dom}} b$. The quality of a set $A \in \Sigma$ can also be measured as a single scalar value through a *unary quality indicator* $I : \Sigma \mapsto \mathbb{R}$. We consider the (additive) *epsilon indicator* (I_{eps}), to be minimized, and the *hypervolume indicator* (I_{hv}), to be maximized [11]. Interestingly, I_{eps}(resp. I_{hv}) is order-preserving (resp. strictly order-preserving) with respect to the weak-dominance set preference relation [11]:

$$A \preccurlyeq_{\mathsf{dom}} B \Rightarrow I_{\mathrm{eps}}(A) \leqslant I_{\mathrm{eps}}(B) \ , \tag{1}$$

$$(A \preccurlyeq_{\mathsf{dom}} B) \wedge \neg(B \preccurlyeq_{\mathsf{dom}} A) \Rightarrow I_{\mathrm{hv}}(\sigma) < I_{\mathrm{hv}}(\sigma') \ . \tag{2}$$

We also define the corresponding set preference relations:

$$A \preccurlyeq_{\mathsf{eps}} B \iff I_{\mathrm{eps}}(A) \leqslant I_{\mathrm{eps}}(B) \ , \tag{3}$$

$$A \preccurlyeq_{\mathsf{hv}} B \iff I_{\mathrm{hv}}(A) \geqslant I_{\mathrm{hv}}(B) \ . \tag{4}$$

Local Optimality. Let $A, B \in \Sigma$. We define the *strict* partial order $\prec$ of a given partial order $\preccurlyeq$ as:

$$A \prec B \iff \neg(B \preccurlyeq A) \wedge (A \preccurlyeq B) \ . \tag{5}$$

Sets A and B are *incomparable* if neither ($A \prec B$) nor ($B \prec A$) holds.

Given a collection of sets Σ, a preorder (preference relation) between sets $\preccurlyeq$, and a neighborhood relation between sets $\mathsf{N} : \Sigma \mapsto 2^\Sigma$, the definition of local optima can be adapted as follows.

Definition 2.1 (Local optimal set, LO-set $(\Sigma, \preccurlyeq, \mathsf{N})$*).* A set $A \in \Sigma$ is a *local optimal set* iff $\forall B \in \mathsf{N}(A) \setminus A, \neg(B \preccurlyeq A)$.

Definition 2.2 (Strict LO-set, sLO-set $(\Sigma, \preccurlyeq, \mathsf{N})$*).* A set $A \in \Sigma$ is a *strict local optimal set* iff $\forall B \in \mathsf{N}(A) \setminus A, A \prec B$.

Under the definitions above, a *Pareto local optimum set* [7, 8] is an LO-set where $\preccurlyeq$ is the set-dominance relation $\preccurlyeq_{\mathsf{dom}}$. It would be a *strict* LO-set under the same definitions if there is no $B \in \mathsf{N}(A)$ such that A and B are incomparable. As another example, a multi-objective local search based on hypervolume ($\preccurlyeq_{\mathsf{hv}}$) stops on an LO-set $A \in \Sigma$ if there exists no neighboring set $B \in \mathsf{N}(A)$ that has a larger hypervolume value. It stops on a strict LO-set if all neighboring sets have a (strictly) smaller hypervolume value than the current set. Therefore, the proposed definitions allow us to compare various types of LO-sets under a common terminology.

A Walk to Sample Local Optimal Sets. Following the definitions of strict and non-strict LO-sets, we define a set-based adaptive walk (Alg. 1), where the first improving neighboring set encountered during neighborhood exploration is accepted. This set-based local search is analogous to a classical single-objective first-improvement local search (or *hill-climber*). In Alg. 1, μ initial solutions are randomly generated and added to a nondominated archive A, which represents the current solution-set. Then, a main loop explores each neighboring solution x' of each element in A in a random order without replacement. If this neighbor x' is nondominated with respect to any solution in A and the cardinality of A is smaller than μ, then A can be trivially improved by adding solution x'. Otherwise, the algorithm explores all sets that are constructed by replacing one solution from A with x'. If the resulting set improves over A, it is accepted. In the case of a *neutral* walk, the solution is also accepted if the resulting set is incomparable with A. Otherwise, the procedure explores the next neighboring set. The main loop stops once all neighbors have been explored, returning a (possibly strict) LO-set, when a budget of solution evaluations has been consumed, or when there is a number of steps without any strict improvement. A step is here defined as a change in the current set, i.e., an iteration of the main loop (lines 5–19).

The proposed adaptive walk shares similarities with existing multi-objective local search methods. Compared against PLS [8] and

Algorithm 1: Adaptive Walk

Input : Set cardinality bound μ, neutral $\in \{\text{TRUE}, \text{FALSE}\}$, partial order $\preccurlyeq \in \{\preccurlyeq_{\text{dom}}, \preccurlyeq_{\text{eps}}, \preccurlyeq_{\text{hv}}\}$

Output : Nondominated set A

1 $A \leftarrow \emptyset$
2 **for** $i \leftarrow 1$ **to** μ **do**
3 $x \leftarrow$ RandomSolution()
4 $A \leftarrow$ FilterDominated($A \cup \{x\}$)
5 **repeat** // main loop
6 **for each** $x' \in \{N(x) \setminus A \mid x \in A\}$ **do** // random order
7 $A' \leftarrow$ FilterDominated($A \cup \{x'\}$)
8 **if** $|A'| \leqslant \mu$ **then**
9 $A \leftarrow A'$
10 **goto** line 19
11 **for each** $x'' \in A$ **do** // random order
12 $A' \leftarrow \{A \cup x'\} \setminus \{x''\}$
13 **if** $(A' \prec A)$ **then** // A' better than A
14 $A \leftarrow A'$
15 **goto** line 19
16 **else if** neutral $\wedge\ A' \preccurlyeq A$ **then**
17 $A \leftarrow A'$ // A' at least as good as A
18 **goto** line 19
19 **until** *A is a (s)LO-set* ***or*** *no budget left* ***or*** *cutoff reached*

SEMO [5], we consider mutually nondominated sets of *bounded* cardinality. The neighbors of a set are the same as those in SEMO [5]; i.e., given $A \in \Sigma$, then $B \in \mathsf{N}(A) \iff |B \setminus A| \leqslant 1 \wedge \forall b \in B \setminus A, \exists a \in A$ such that $b \in N(a)$ [2]. The main difference is that we explore neighboring sets *without* replacement, which allow us to detect when the walk falls into a (possibly strict) LO-set. As in single-objective local search, the proposed non-neutral adaptive walk always falls into an LO-set, whereas a neutral walk may either eventually fall into a strict LO-set, or terminate without reaching any type of LO-set. By using this adaptive walk, we can experimentally estimate the number, quality, and dissimilarity of various types of LO-sets, as shown below.

3 EXPERIMENTAL SETUP

Multi-objective *nk*-Landscapes. We consider *nk*-landscapes as a problem-independent model of multi-objective multimodal landscapes [1, 10]. Candidate solutions are binary strings of size n and the objective function vector $f = (f_1, \ldots, f_i, \ldots, f_m)$ is defined as $f\colon \{0,1\}^n \mapsto [0,1]^m$ such that each objective f_i is to be maximized. As in well-established single-objective *nk*-landscapes [3], each separate objective function value $f_i(x)$ of a solution $x = (x_1, \ldots, x_j, \ldots, x_n)$ is an average value of the individual contributions associated with each variable x_j. Given objective f_i, $i \in \{1, \ldots, m\}$, and variable x_j, $j \in \{1, \ldots, n\}$, a component function $f_{ij}\colon \{0,1\}^{k+1} \mapsto [0,1]$ assigns a real-valued contribution for every combination of x_j and its k *epistatic interactions* $\{x_{j_1}, \ldots, x_{j_k}\}$. These f_{ij}-values are uniformly distributed in $[0,1]$. Thus, the individual contribution of a variable x_j depends on its value and on the values of $k < n$ other variables $\{x_{j_1}, \ldots, x_{j_k}\}$. In this work, the epistatic interactions, i.e., the k variables that influence the contribution of x_j, are set uniformly at random among the $(n-1)$ variables other than x_j [3]. By increasing the number of epistatic interactions k from 0 to $(n-1)$, landscapes can be gradually tuned from smooth to rugged. We use the same epistatic degree and interactions for all the objectives. By construction, it is very unlikely that different solutions map to the same point in the objective space.

Parameter Settings. We generate 15 multi-objective *nk*-landscapes with the following settings. The problem size is set to $n = 16$, number of objectives $m \in \{2, 3, 5\}$, and problem non-linearity $k \in \{0, 1, 2, 4, 8\}$, that is, from linear to highly rugged landscapes. We generate one instance independently at random for each combination of instance settings. We run the adaptive walk (Alg. 1) with respect to the set preference relations $\{\preccurlyeq_{\text{dom}}, \preccurlyeq_{\text{hv}}, \preccurlyeq_{\text{eps}}\}$ and with various set cardinality bounds $\mu \in \{2, 4, 8, 16, 32\}$. The reference set for computing I_{eps} is the (exact) Pareto front. The reference point for computing I_{hv} is set to $(0, \ldots, 0)$. We experiment with both neutral and non-neutral walks. In order to ensure a reasonable runtime for neutral walks, we set a maximum budget of 10^7 evaluations and a cutoff of 30 consecutive iterations of the main loop without improvement. The neighborhood relation among solutions (N) is defined by the *1-bit-flip* operator; i.e., two solutions are neighbors if the Hamming distance between them is one. We replicate each experiment 30 times with different random seeds.

4 EXPERIMENTAL ANALYSIS

4.1 Number of Local Optimal Sets

As a first question, we investigate the number of LO-sets of each type, that is, for each set preference relation ($\preccurlyeq_{\text{dom}}$, $\preccurlyeq_{\text{hv}}$, and $\preccurlyeq_{\text{eps}}$) and either strict or non-strict definition (LO/sLO). Given previous results regarding Pareto local optimum sets [7], we expect the number of LO-sets to be affected by the number of objectives (m), the epistasis (k) and the cardinality bound (μ). However, we do not know how each type of LO-set is affected by these characteristics. Moreover, although we conjecture that some LO-sets of one type are also LO-sets of other types, their relative ratios are unknown. To answer these questions, after running the adaptive walk as described in the previous section, we simply count how many of the sets returned at the end of the runs satisfy the definition of each type of LO-set. Results are shown in Fig. 1 for selected settings. Results on other instances confirm the trends observed here.

The first observation is that non-neutral walks using a particular set preference relation always find a non-strict LO-set according to the same relation, in every run. That is, a walk based on $\preccurlyeq_{\text{dom}}$ (resp. $\preccurlyeq_{\text{eps}}$, $\preccurlyeq_{\text{hv}}$) always falls into an $\text{LO}_{\preccurlyeq_{\text{dom}}}$ (resp. $\text{LO}_{\preccurlyeq_{\text{eps}}}$, $\text{LO}_{\preccurlyeq_{\text{hv}}}$). Moreover, the LO-set where a given walk falls into might be the same at different executions, as observed, for instance $m = 2$, $k = 0$ with $\mu = 32$. This suggests that there is a single LO-set in this case, which is not a surprise because the corresponding *nk*-landscape is linear ($k = 0$) and its Pareto set cardinality is lower than $\mu = 32$.

We did not notice any difference between neutral or non-neutral walks with $\preccurlyeq_{\text{hv}}$, which suggests that neighboring LO-sets with the same hypervolume value are rare, thus there is no neutrality in the corresponding landscapes. Although we do not expect that real-world problems have many sets with the same hypervolume value, we cannot generalize this finding to any landscape since it

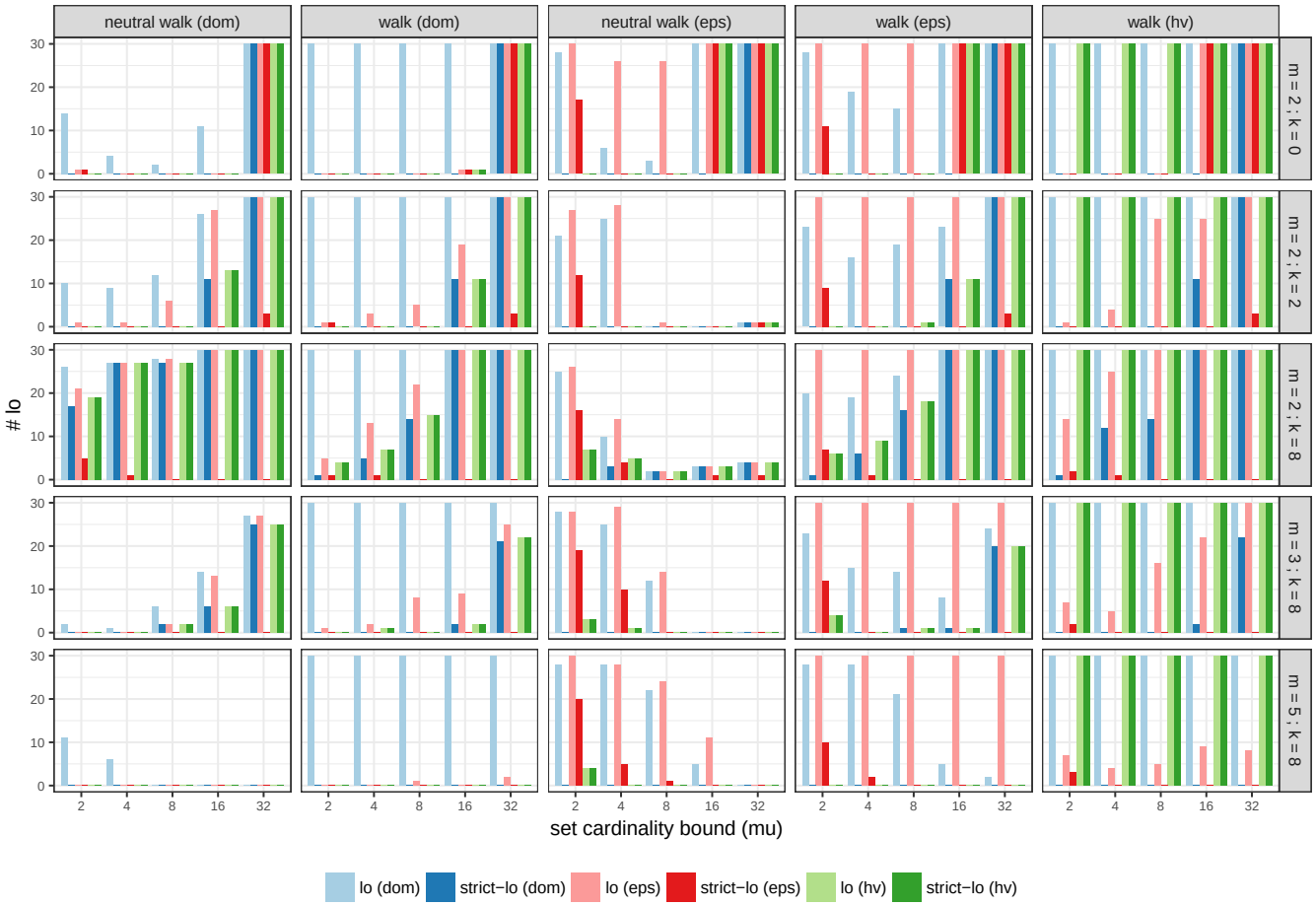

Figure 1: Number of each type of LO-set found by each type of LS walk (by column) for different instances (by row), depending on the set cardinality bound (μ). Results for neutral walk (hv) is not reported because they are the same as for walk (hv).

is easy to think of artificial examples where two neighboring sets have equal hypervolume values. By contrast, we observe a large neutrality for $\preccurlyeq_{\text{dom}}$ and $\preccurlyeq_{\text{eps}}$, as shown by the large differences between neutral and non-neutral walks in such cases. In fact, the neutral walk using $\preccurlyeq_{\text{dom}}$ is only able to find a $\text{sLO}_{\preccurlyeq_{\text{dom}}}$ when μ is large, and when there are few objectives, e.g., for $m = 2$ and $k = 8$. We attribute this to a large non-linearity in the objective values in such cases, with many incomparable neighboring sets, which seem to increase the number of strict LO-sets. By contrast, the neutral-walk using $\preccurlyeq_{\text{eps}}$ is only able to find a $\text{sLO}_{\preccurlyeq_{\text{eps}}}$ when μ is small and/or when non-linearity is small ($k = 0$), whereas the neutral-walk using $\preccurlyeq_{\text{hv}}$ is always able to find a $\text{sLO}_{\preccurlyeq_{\text{hv}}}$, as already reported above. The probable reason is that there are more neighboring sets with the same epsilon value and/or that the hypervolume gradient is easier to optimize that of epsilon. Interestingly, when there are many objectives, and when μ is especially small relative to the size of the exact Pareto set, it appears to be difficult to obtain a $\text{LO}_{\preccurlyeq_{\text{hv}}}$ with any method besides a walk based on $\preccurlyeq_{\text{hv}}$.

To summarize, by comparing LO-sets under $\preccurlyeq_{\text{dom}}$ and $\preccurlyeq_{\text{hv}}$, we conjecture that: $\text{sLO}_{\preccurlyeq_{\text{dom}}} \Rightarrow \text{sLO}_{\preccurlyeq_{\text{hv}}} \Rightarrow \text{LO}_{\preccurlyeq_{\text{hv}}} \Rightarrow \text{LO}_{\preccurlyeq_{\text{dom}}}$. Although we do not obtain any clear trend by comparing LO-sets under $\preccurlyeq_{\text{dom}}$ and $\preccurlyeq_{\text{hv}}$ with LO-sets under $\preccurlyeq_{\text{eps}}$, we conjecture that $\text{sLO}_{\preccurlyeq_{\text{dom}}} \Rightarrow \text{LO}_{\preccurlyeq_{\text{eps}}}$. We also suspect that there are slightly more $\text{LO}_{\preccurlyeq_{\text{eps}}}$ than $\text{LO}_{\preccurlyeq_{\text{hv}}}$, given that the walk based on $\preccurlyeq_{\text{dom}}$ consistently finds more of those, but the difference seems to be rather small.

Finally, a general observation worth mentioning is that, whatever the set preference relation, the adaptive walk gets more easily trapped into an LO-set as the problem non-linearity k increases, and as the number of objectives m decreases.

4.2 Length of Adaptive Walks

As in single-objective optimization, the length of the adaptive walk provides an estimation of the number of LO-sets. The number of steps performed by the algorithm defines the length of the adaptive walk. This length is an estimator of the diameter of local optima's basins of attraction. Roughly speaking and assuming isotropy in the search space, the longer the walk, the larger the basins size, and the lower the number of local optima [3]. Fig. 2 reports the number of steps performed by each type of adaptive walk. In our experiments, the solution space has the same size for all instances; i.e. $|X| = 2^n = 2^{16}$, whatever k and m. However, the number of candidate sets depends on the set cardinality bound μ and on the

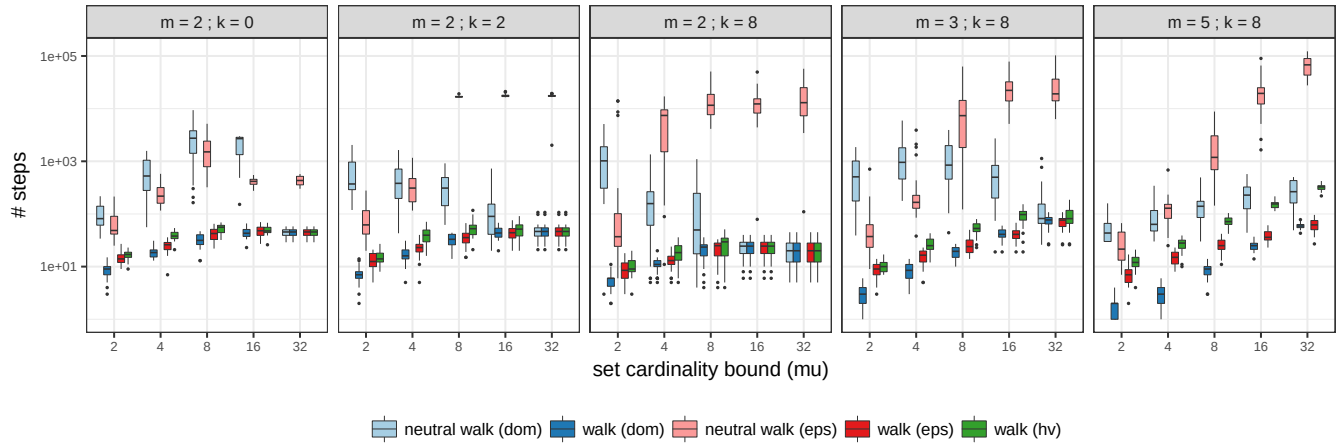

Figure 2: Number of steps performed by each walk (colors) for different instances (non-linearity k and number of objectives m, by column), depending on the set cardinality bound.

dominance relations between solutions. For a given μ, the number of candidate sets is bounded by $\sum_{i=1}^{\mu} \binom{|2^X|}{i} = \sum_{i=1}^{\mu} \binom{2^{(2^{16})}-1}{i}$. Therefore, for a given instance, a larger value of μ induces an exponentially larger number of candidate sets. Of course, depending on the dominance relations between solutions, many candidate sets might be equivalent once dominated solutions are discarded.

We observe that the length of adaptive walks typically increases with μ. Therefore, a local search is more easily stuck when μ is small. This means that the absolute number of LO-sets decreases with μ. When relating that to the number of candidate sets, we argue that the proportional number of LO-sets is larger when μ is small, whatever the set preference relation. As expected, the length of adaptive walks decreases with the problem non-linearity k. As for single-objective nk-landscapes, the larger k, the larger the number of local optima [3]. When considering the number of objectives m, we observe that adaptive walks runs longer as m increases, especially when μ is relatively large. A local search has a larger probability of getting stuck for two-objective landscapes than for three- and five-objective landscapes. This suggests that the number of LO-sets decreases with the objective space dimension.

When comparing neutral and non-neutral walks for a given set preference relation, we observe that neutral walks typically run longer. There are two potential explanations for this result: (i) a neutral walk may fail to identify a strict LO-set simply because it does not exist, explaining why the number of steps reaches the overall budget limit for some settings; and (ii) every strict LO-set under a given set preference relation is also an LO-set under the same relation, thus the number of sLO-sets is smaller or equal than the number of LO-sets. This is the case for $\preccurlyeq_{\mathsf{dom}}$ and $\preccurlyeq_{\mathsf{eps}}$. Interestingly, the gap between the number of strict and non-strict LO-sets seems to decrease with μ for $\preccurlyeq_{\mathsf{dom}}$, whereas it increases for $\preccurlyeq_{\mathsf{eps}}$. In fact, for large sets, almost all $\mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}}$ are $\mathrm{sLO}_{\preccurlyeq_{\mathsf{dom}}}$, as also noticed in Fig. 1. We attribute this to the fact that it is more unlikely to come across a neighboring solution that is not dominated by the set when this set is larger. As already mentioned above, in the case of $\preccurlyeq_{\mathsf{hv}}$, there is no distinction between neutral and non-neutral walks for the considered instances, an $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$ is always a $\mathrm{sLO}_{\preccurlyeq_{\mathsf{hv}}}$.

Let us now compare non-neutral walks and non-strict LO-sets for different set preference relations. For $m = 2$ and $\mu \in \{16, 32\}$, the length of the adaptive walk is roughly the same for all relations. They are the sole settings where the cardinality of LO-sets is actually smaller than the bound μ (not reported here due to space restriction), which is explained by the fact that μ is larger than the Pareto set in those cases. This suggests that there is no distinction between LO-sets under the different set preference relations when μ has the same order of magnitude than the Pareto set, as also observed in Fig. 1. By contrast, for other instances, the length of the adaptive walk for $\preccurlyeq_{\mathsf{dom}}$ is typically smaller than for $\preccurlyeq_{\mathsf{eps}}$, which is itself typically smaller than for $\preccurlyeq_{\mathsf{hv}}$. This gives us more evidence that, when μ is smaller than the Pareto set, we have more $\mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}}$ than $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$, and more $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$ than $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$. A multi-objective local search with bounded archive is then expected to get more easily trapped when comparing sets in terms of dominance rather than in terms of epsilon or hypervolume. On the contrary, a hypervolume-based local search is expected to perform more steps before being stuck.

4.3 Quality of Local Optimal Sets

In Fig. 3, we report the quality of the final set obtained by each walk on a selected subset of instances. The quality of the resulting sets is evaluated both in terms of the additive epsilon indicator (I_{eps}) and of the relative hypervolume deviation (I_{hvr}). The relative hypervolume deviation is computed as $I_{\mathrm{hvr}}(A) := (I_{\mathrm{hv}}(R) - I_{\mathrm{hv}}(A))/I_{\mathrm{hv}}(R)$, where R is the Pareto set. Although a walk based on $\preccurlyeq_{\mathsf{hv}}$ (resp. $\preccurlyeq_{\mathsf{eps}}$) always outputs the best hypervolume (resp. epsilon) value, the indicator value of the *final* set is not necessarily the best one observed during the search process for walks that are based on a different set preference relation. In the next section, we will analyze the best indicator value obtained at different time steps.

Unsurprisingly, the quality of LO-sets always improves with higher set cardinality bound μ, for both hypervolume and epsilon. In fact, as suspected above, for all variants, the set found by each walk always maps to the Pareto set for the linear two-objective instance ($k = 0$, $m = 2$), as long as the cardinality bound is larger than the number of Pareto optimal solutions ($\mu = 32$). This means

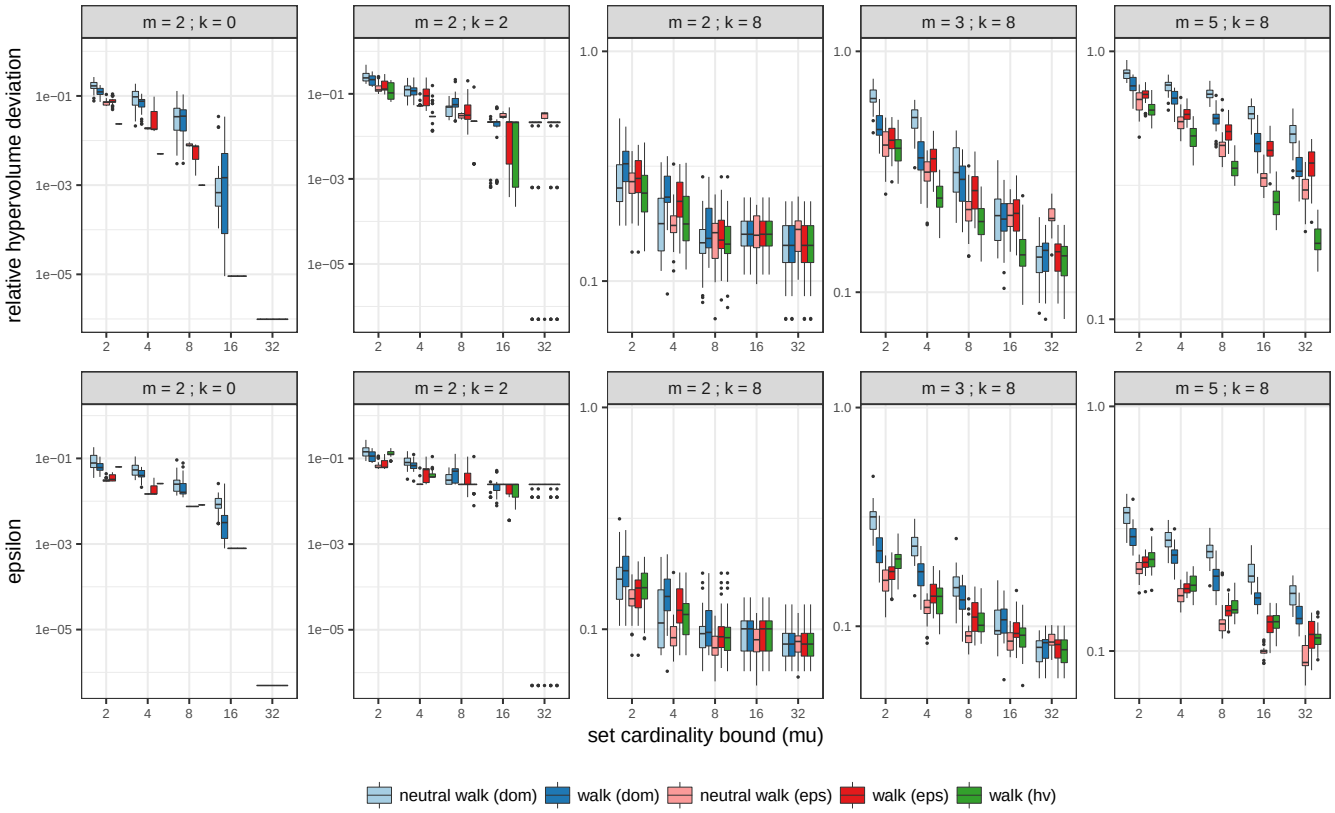

Figure 3: Final relative hypervolume deviation and epsilon value obtained by each walk (colors) for different instances (non-linearity k and number of objectives m, by column), depending on the set cardinality bound.

that, whatever the set preference relation, there is only one LO-set for this setting: the Pareto optimal set.

For all instances, the walk based on $\preccurlyeq_{\text{hv}}$ consistently converges to better LO-sets in terms of hypervolume. When analyzing LO-sets in terms of epsilon values, it is more difficult to distinguish between the walk based on $\preccurlyeq_{\text{hv}}$ and the (non-neutral) walk based on $\preccurlyeq_{\text{eps}}$. However, a *neutral* walk based on $\preccurlyeq_{\text{eps}}$ often leads to better epsilon values for the most difficult instances (with large m and k), although in much more steps as depicted in Fig. 2. This once again emphasizes the high neutral degree induced by $\preccurlyeq_{\text{eps}}$. As such, we argue that an epsilon-based local search will not necessarily converge to better epsilon values than a hypervolume-based local search, unless it explicitly handles equivalent sets in terms of epsilon. At last, the walks based on $\preccurlyeq_{\text{dom}}$ seem to converge to lower-quality LO-sets, in terms of both indicators, as the number of objectives increases. This confirms that dominance is probably not the best option to distinguish between candidate sets in many-objective local search.

4.4 Convergence Profile of Adaptive Walks

In order to better appreciate the anytime behavior of the walks under different settings, we report in Fig. 4 the convergence of the *best-found* indicator value for different budgets, measured in terms of a number of evaluations. This is different from Fig. 3, where only the quality of the final set was analyzed. It is worth noticing that we do not consider any restart mechanisms in our algorithm, and once an algorithm stops at a given iteration, it is assumed that the quality remains the same for subsequent ones.

First, we observe that a (non-neutral) walk under $\preccurlyeq_{\text{dom}}$ is only efficient when $\mu = 32$ and $m = 2$, that is, when μ is larger than the Pareto set. Otherwise, the performance of such a walk is always very low, both in terms of epsilon and hypervolume. A walk based on $\preccurlyeq_{\text{eps}}$ follows the same trend according to hypervolume, although it is always slightly better. By contrast, it performs much better in terms of epsilon, although it is mostly outperformed by the walk under $\preccurlyeq_{\text{hv}}$, except when μ is particularly small. The latter is actually never outperformed in terms of hypervolume, except for a few settings with a small μ and small m. It is also often the second-best approach in terms of epsilon, even consistently better than the former when $\mu \geqslant 8$.

Regarding neutral walks, the one based on $\preccurlyeq_{\text{dom}}$ performs nicely for small μ values, but not so good for $m = 5$, which once again might explain the low performance of dominance-based search in many-objective optimization. Showing the opposite behavior, the neutral walk based on $\preccurlyeq_{\text{eps}}$ is often the best-performing approach in terms of epsilon, and second-best approach in terms of hypervolume, except for large μ and small m. It performs particularly well for $m = 5$, whatever the set cardinality bound.

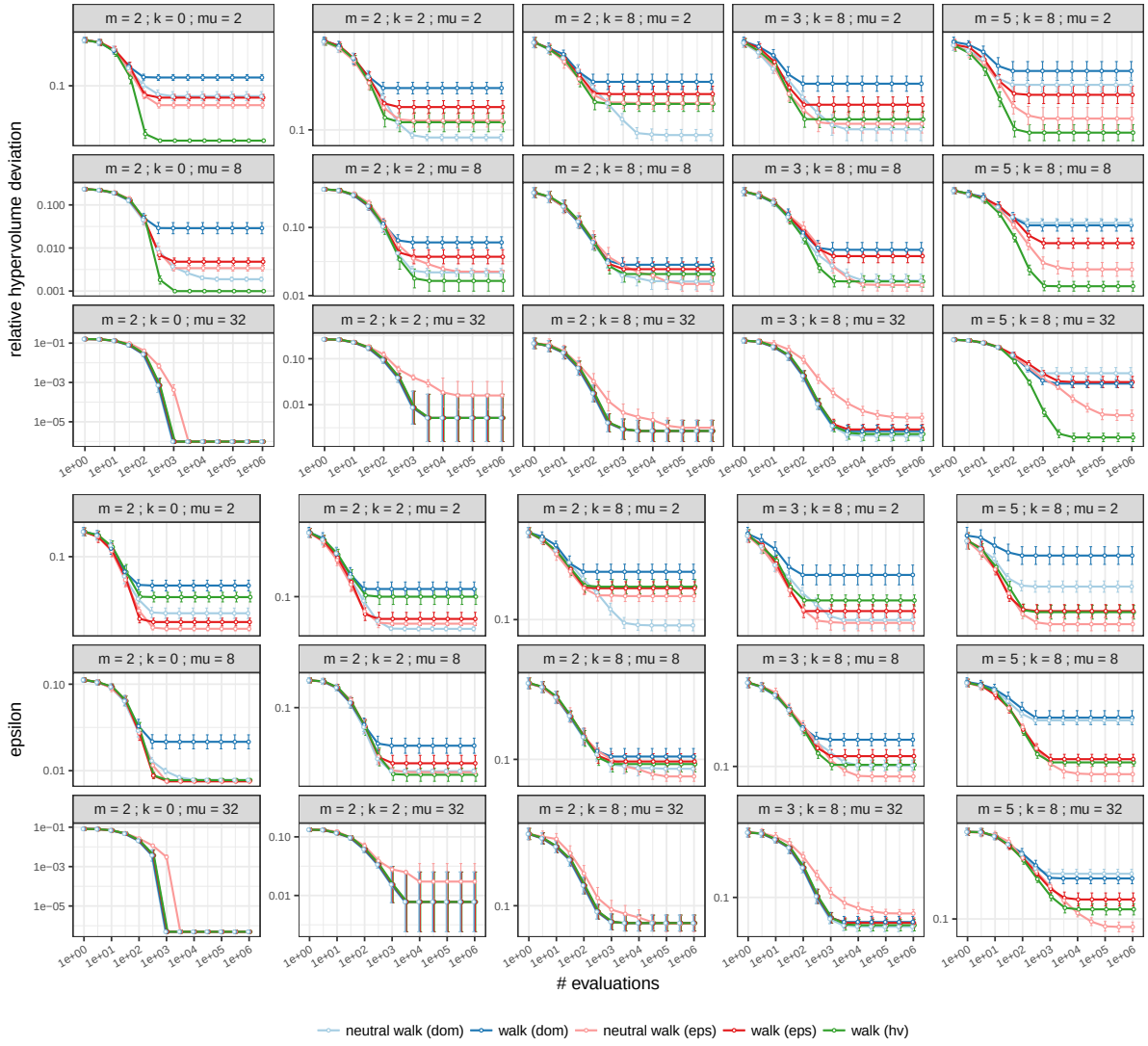

Figure 4: Convergence profile of hypervolume and epsilon for each walk (colors), for different instances (non-linearity k and number of objectives m, by column) and different set cardinality bound (μ, by row), depending on the number of evaluations.

4.5 Distance between Local Optimal Sets

As a last question, we go deeper into the comparison of LO-sets under different set preference relations by investigating their dissimilarity in the space of sets. In particular, we want to know how much different is a $\text{LO}_{\preccurlyeq_{\text{eps}}}$ or a $\text{LO}_{\preccurlyeq_{\text{hv}}}$ from a $\text{LO}_{\preccurlyeq_{\text{dom}}}$. We do not consider strict LO-sets in this section, since they do not necessarily exist for all settings, and since the success rate for the corresponding neutral walks is typically lower than 1.

Let us define the *distance* between a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ and a $\text{LO}_{\preccurlyeq_{\text{eps}}}$ as the length (number of steps) required by a walk based on $\preccurlyeq_{\text{eps}}$ to reach a $\text{LO}_{\preccurlyeq_{\text{eps}}}$, while starting from a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ as an initial set. To do so, (i) we simply start by running a walk under $\preccurlyeq_{\text{dom}}$ until it falls into a $\text{LO}_{\preccurlyeq_{\text{dom}}}$, and then (ii) we run a walk under $\preccurlyeq_{\text{eps}}$ starting from the obtained $\text{LO}_{\preccurlyeq_{\text{dom}}}$. Only the steps performed in the second phase are taken into account to measure the distance. The distance between a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ and a $\text{LO}_{\preccurlyeq_{\text{hv}}}$ follows the same reasoning, but using a walk under $\preccurlyeq_{\text{hv}}$. This notion of distance gives how many *1-bit-flips*, performed on any solution from the initial set, separates a set A from a set B. Thus, if $dist(A, B) = d$, then A may differ from B in d solutions, all connected at Hamming distance 1, or they may differ in a single solution with Hamming distance d.

The obtained distances are reported in Fig. 5. When compared against the walks that start from a random set, as reported in Fig. 2, the number of steps performed from a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ is lower by an order of magnitude. This means that a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ is much closer to a $\text{LO}_{\preccurlyeq_{\text{eps}}}$ or a $\text{LO}_{\preccurlyeq_{\text{hv}}}$ than a random set is with any of the three. The distance between a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ and a $\text{LO}_{\preccurlyeq_{\text{eps}}}$ is often larger for medium μ values ($\mu \in \{4, 8\}$) than for small and large values ($\mu \in \{2, 16, 32\}$).

When considering the hypervolume, the distances from a $\text{LO}_{\preccurlyeq_{\text{dom}}}$ to a $\text{LO}_{\preccurlyeq_{\text{hv}}}$ are always larger than to a $\text{LO}_{\preccurlyeq_{\text{eps}}}$. When $m = 2$, these distances roughly follow the same trend as for $\text{LO}_{\preccurlyeq_{\text{eps}}}$, however,

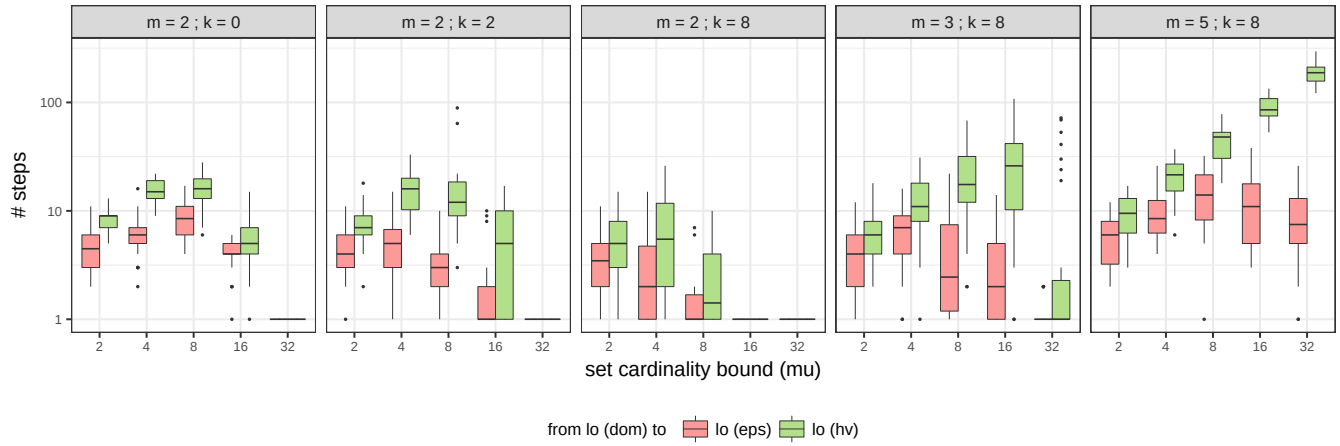

Figure 5: Number of steps performed by the walk to go from a $\mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}}$ to a $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$ and $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$ for different instances (non-linearity k and number of objectives m, by column), depending on the set cardinality bound.

when $m \geqslant 3$, they seem to increase with μ, with the exception of $\mu = 32$ and $m = 3$ where the distance is close to 1. For $m = 5$, the gap relative to the distances corresponding to $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$ increases by several orders of magnitude.

5 CONCLUSIONS

In this paper, we empirically studied the properties of various types of LO-sets. Our results confirm previous findings and observations in multi-objective optimization, such as the fact that algorithms relying solely on dominance tend to perform worse for more than three objectives. We also observed that the number of LO-sets of any type increases with increasing ruggedness of the landscape, and with decreasing number of objectives and decreasing cardinality bound. Similar results were previously known for (bounded size) PLO-sets [7] and here we show that they are true for other types of LO-sets. In addition, we advance several hypotheses based on our experimental results. In particular, we conjecture that:

$$\mathrm{sLO}_{\preccurlyeq_{\mathsf{dom}}} \Rightarrow \mathrm{sLO}_{\preccurlyeq_{\mathsf{hv}}} \Rightarrow \mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}} \Rightarrow \mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}} \Rightarrow \mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$$

which means that there are more $\mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}}$ than $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$, but more $\mathrm{sLO}_{\preccurlyeq_{\mathsf{hv}}}$ than $\mathrm{sLO}_{\preccurlyeq_{\mathsf{dom}}}$ (except for the trivial case when the cardinality bound is larger than the actual size of the Pareto set, in such case there is no distinction between the various LO-sets). In addition to the implications above, we also observed that there are many more $\mathrm{LO}_{\preccurlyeq_{\mathsf{dom}}}$ than $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$, and slightly more $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$ than $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$, the latter being perhaps the most surprising conclusion.

We conjecture that our findings regarding $\mathrm{LO}_{\preccurlyeq_{\mathsf{eps}}}$ (resp. $\mathrm{LO}_{\preccurlyeq_{\mathsf{hv}}}$) generalize to other LO-sets under any order-preserving (resp. strictly order-preserving) indicators. Our analysis should also guide the design of new multi-objective optimizers. For instance, we conclude that an epsilon-based local search does not necessarily converge to better epsilon values than a hypervolume-based local search, unless it explicitly handles sets that are equivalent in terms of epsilon.

Our conclusions provide at least two directions for further work. One direction should attempt to formally prove some of our conjectures, thus increasing our theoretical understanding of multi-objective landscapes. A second direction should try to extend our experimental analysis to additional problems, different neighborhoods and other order-preserving indicators, to corroborate that our conjectures indeed generalize as expected. Of particular interest is the extension of our work to LO-sets for continuous problems [4]. Furthermore, there are other factors that were not considered here, such as the size of the search space and the correlation between objectives.

Acknowledgments. This work benefited from the COST action "Improving applicability of nature-inspired optimization by joining theory and practice" (ImAppNIO, CA15140, European Cooperation in Science and Technology).

REFERENCES

[1] H. E. Aguirre and K. Tanaka. 2007. Working principles, behavior, and performance of MOEAs on MNK-landscapes. *Eur. J. Oper. Res.* 181, 3 (2007), 1670–1690.
[2] M. Basseur, A. Goëffon, A. Liefooghe, and S. Verel. 2013. On Set-based Local Search for Multiobjective Combinatorial Optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2013*, C. Blum and E. Alba (Eds.). ACM Press, New York, NY, 471–478.
[3] S. A. Kauffman. 1993. *The Origins of Order.* Oxford University Press.
[4] P. Kerschke, H. Wang, M. Preuss, C. Grimme, A. H. Deutz, H. Trautmann, and M. Emmerich. 2016. Towards Analyzing Multimodality of Continuous Multiobjective Landscapes. In *Parallel Problem Solving from Nature - PPSN XIV*, J. Handl et al. (Eds.). LNCS, Vol. 9921. Springer, 962–972.
[5] M. Laumanns, L. Thiele, and E. Zitzler. 2004. Running time analysis of evolutionary algorithms on a simplified multiobjective knapsack problem. *Natural Computing* 3, 1 (2004), 37–51.
[6] M. López-Ibáñez, J. D. Knowles, and M. Laumanns. 2011. On Sequential Online Archiving of Objective Vectors. In *EMO*, R. H. C. Takahashi et al. (Eds.). LNCS, Vol. 6576. Springer, 46–60.
[7] M. López-Ibáñez, A. Liefooghe, and S. Verel. 2014. Local Optimal Sets and Bounded Archiving on Multi-objective NK-Landscapes with Correlated Objectives. In *PPSN 2014*, T. Bartz-Beielstein et al. (Eds.). LNCS, Vol. 8672. Springer, 621–630.
[8] L. Paquete, T. Schiavinotto, and T. Stützle. 2007. On Local Optima in Multiobjective Combinatorial Optimization Problems. *Annals of Operations Research* 156 (2007), 83–97.
[9] S. Verel, A. Liefooghe, and C. Dhaenens. 2011. Set-based Multiobjective Fitness Landscapes: A Preliminary Study. In *GECCO*, N. Krasnogor and P. L. Lanzi (Eds.). ACM Press, New York, NY, 769–776.
[10] S. Verel, A. Liefooghe, L. Jourdan, and C. Dhaenens. 2013. On the Structure of Multiobjective Combinatorial Search Space: MNK-landscapes with Correlated Objectives. *Eur. J. Oper. Res.* 227, 2 (2013), 331–342.
[11] E. Zitzler, J. D. Knowles, and L. Thiele. 2008. Quality Assessment of Pareto Set Approximations. In *Multi-objective Optimization: Interactive and Evolutionary Approaches*, J. Branke et al. (Eds.). LNCS, Vol. 5252. Springer, 373–404.
[12] E. Zitzler, L. Thiele, and J. Bader. 2010. On Set-Based Multiobjective Optimization. *IEEE Trans. Evol. Comput.* 14, 1 (2010), 58–79.