



**HAL**  
open science

## Designing helper objectives in multi-objectivization

Shoichiro Tanaka, Arnaud Liefoghe, Keiki Takadama, Hiroyuki Sato

► **To cite this version:**

Shoichiro Tanaka, Arnaud Liefoghe, Keiki Takadama, Hiroyuki Sato. Designing helper objectives in multi-objectivization. CEC 2024 – IEEE Congress on Evolutionary Computation, Jun 2024, Yokohama, Japan. pp.1-8, 10.1109/CEC60901.2024.10612125 . hal-04692939

**HAL Id: hal-04692939**

**<https://hal.science/hal-04692939>**

Submitted on 10 Sep 2024

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Designing Helper Objectives in Multi-objectivization

Shoichiro Tanaka  
Graduate School of Informatics  
The University of Fukuchiyama  
Fukuchiyama, Kyoto, Japan  
tanaka-shoichiro@fukuchiyama.ac.jp

Arnaud Liefooghe  
LISIC  
Université du Littoral Côte d'Opale  
Calais, France  
arnaud.liefooghe@univ-littoral.fr

Keiki Takadama  
Graduate School of Information and Engineering Sciences  
The University of Electro-Communications  
Chofu, Tokyo, Japan  
keiki@inf.uec.ac.jp

Hiroyuki Sato  
Graduate School of Information and Engineering Sciences  
The University of Electro-Communications  
Chofu, Tokyo, Japan  
h.sato@uec.ac.jp

**Abstract**—Multi-objectivization transforms a single-objective optimization problem into a multi-objective one in order to facilitate the search for high-quality solutions with respect to the original target objective. This paper focuses on the multi-objectivization strategy of adding a helper objective. Depending on its definition, the helper objective might have a positive or negative impact on optimization. For multi-objectivization to work well, it is essential to select the helper objective with care, according to the nature of the target objective. However, the design of this helper objective remains unclear: should it be completely independent of the target objective or, by contrast, correlated in some respects? We propose and analyze different methods for generating helper objectives with varying degrees of correlation to the target objective, with the aim of guiding the setting of multi-objectivization. Inspired by existing works on multi-objective NK landscapes, we are particularly interested in the joint setting of the correlation between objective-values and the similarity of variable interactions on both objectives. We approximately decompose the target function into several sub-functions based on the Walsh transform. The proposed method combines these sub-functions to create helper objectives with different levels of correlation and heterogeneity. By analyzing bi-objective instances made of a target and of a helper objective under different definitions, we gain insights into the selection of helper objectives depending on the target objective. Our experimental findings suggest that a helper objective with a positive correlation and a smoother landscape is beneficial for multi-objectivization.

**Index Terms**—Multi-objective optimization, Combinatorial optimization, Multi-objectivization, pseudo-Boolean functions, Walsh transform.

## I. INTRODUCTION

Multi-objectivization [1] consists in re-formulating a single-objective optimization problem (SOP) into a multi-objective optimization problem (MOP). In SOP, local search is known to get stuck into local optima, resulting in premature convergence to lower quality solutions. However, in MOP, solutions are compared based on Pareto dominance instead of a single objective function value. The search is thus not necessarily

trapped into single-objective local optima. Relaxing the selection pressure through multi-objectivization allows the search process to escape from a local optimum and to potentially reach better solutions. Recent studies indicate that, in both continuous and discrete search spaces, multi-objectivization can connect local optima with MOP-specific solutions such as Pareto optimal and Pareto local optimal solutions [2]. Therefore, multi-objectivization fundamentally changes the landscape and significantly impacts search behavior.

Knowles et al. [1] were the first to introduce the multi-objectivization paradigm and demonstrated its effect on search performance. Jensen [3] later multi-objectivized a SOP by adding a *helper* objective that is different from the *target* (original) objective. Multi-objectivization was found to be successful for some multi-modal or neutral SOPs. Conversely, Brockhoff et al. [4] demonstrated that additional objectives might positively or negatively impact search performance. Therefore, we argue that the selection of an appropriate helper objective is crucially important in multi-objectivization.

To facilitate search exploration, we must thoughtfully design an appropriate helper objective for a given target objective. However, it remains unclear how to design this helper objective, and whether it should be completely independent of the target objective or, by contrast, somehow correlated. Insights into these issues would be beneficial to practitioners. Specifically, we investigate the correlation between helper and target objective values [5], as well as the heterogeneity of variable interactions [6]. In this paper, we examine how the correlation and the heterogeneity influence the characteristics of multi-objectivized problems in order to guide design of helper objectives. We approximately decompose the target objective into several sub-functions based on the Walsh transform [7]. This research draws inspiration from [8], where surrogates of pseudo-Boolean functions are designed based on the Walsh transform. Decomposing the target objective into several sub-functions that are optimized simultaneously is one of the main

approach in multi-objectivization [9], [10]. The focus is given to problems that can be intuitively decomposed into sub-problems, such as the traveling salesperson problem. Although we also start by decomposing the target objective into sub-functions, our method is inherently different as these sub-functions are further used to design a single helper objective. The Walsh transform allows us to decompose pseudo-Boolean black-box functions, where the problem structure is not known in advance. By employing the sub-functions, it becomes possible to generate a helper objective from a uni-modal function with homogeneous variable interactions. We propose a multi-objectivization method that generates helper objectives with a gradual correlation and heterogeneity from solutions sampled by single-objective search. We compare our method to the original SOP and explore the impact of multi-objectivization. Our analysis provides a better understanding of the nature of the resulting MOP and valuable insights on the design of helper objectives. Our contributions can be summarized as follows:

- 1) We propose a helper objective generator based on the Walsh transform that can adjust the correlation between objectives and the heterogeneity of variable interactions for black-box pseudo-Boolean SOPs ;
- 2) We investigate the conditions under which multi-objectivization improves search performance, depending on the level of correlation and heterogeneity ;
- 3) We predict the effect of multi-objectivisation from local landscape features that can be computed during the generation of the helper objective.

The paper is organized as follows. Section 2 describes single- and multi-objective optimization. Section 3 provides the basis for the Walsh transform. Section 4 introduces the proposed helper objective generator and the related multi-objectivization search approach. Section 5 gives the details of our experimental setup. Section 6 presents and discusses the results of our experiments. Section 7 concludes the paper and discusses future research.

## II. DEFINITIONS

This section covers the basics of single- and multi-objective optimization. It also introduces the benchmark problems considered for the target objective.

### A. Single-objective Optimization

We assume the maximization of an  $N$ -dimensional pseudo-Boolean function  $f : X = \{0, 1\}^N \rightarrow \mathbb{R}$  as the target objective. Candidate solutions  $\mathbf{x} = (x_1, \dots, x_N)$  are  $N$ -bit strings, and the size of the solution space  $|X|$  is  $2^N$ . The goal is to find the global optimum  $\mathbf{x}^*$  with the highest function value  $f(\mathbf{x})$  in the solution space  $X$ . SOPs often have local optima, i.e. solutions that have no improving neighbors. The neighborhood is defined in terms of Hamming distance. For given  $\mathbf{x} \in X$ , its neighbors  $\mathcal{N}(\mathbf{x})$  are solutions whose Hamming distance to  $\mathbf{x}$  is 1.

### B. NK Landscapes

We consider NK landscapes [11] for defining the target objective function. They consist of pseudo-Boolean functions. Their difficulty can be adjusted with two parameters:  $N$  is the number of variables, and  $K$  is the number of co-variables per variable. The objective function of NK landscapes is defined as follows:

$$f(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N g_i(\text{mask}_i(\mathbf{x})) \quad (1)$$

where  $g_i : \{0, 1\}^{K+1} \rightarrow [0, 1]$  is a sub-function that determines the contribution of the  $i$ -th variable  $x_i$  to the objective value.  $\text{mask}_i : \{0, 1\}^N \rightarrow \{0, 1\}^{K+1}$  extracts variable  $x_i$  and its  $K$  co-variables from solution  $\mathbf{x}$ . The  $K$  co-variables are randomly selected from the  $N - 1$  variables other than  $x_i$ . The maximum value of  $K$  is thus  $N - 1$ . The higher the value of  $K$ , the more local optima there are, which makes it harder to find a global optimum.

### C. Multi-objective Optimization

In multi-objectivization, the target SOP is re-formulated and explored as a MOP. We define  $\mathbf{f} = (f_{\text{target}}, f_{\text{helper}})$  as a bi-objective function  $\mathbf{f} : X \rightarrow Z$  that maps each solution  $\mathbf{x} \in X$  to a vector in the objective space  $\mathbf{z} = \mathbf{f}(\mathbf{x}) \in Z \subseteq \mathbb{R}^2$ . In a MOP, solutions are compared based on Pareto dominance, rather than a single objective value. Given two solutions  $\mathbf{x}, \mathbf{x}' \in X$ ,  $\mathbf{x}$  is dominated by  $\mathbf{x}'$  iff for all  $i \in \{1, \dots, M\}$   $f_i(\mathbf{x}) \leq f_i(\mathbf{x}')$  and there exists  $j \in \{1, \dots, M\}$  such that  $f_j(\mathbf{x}) < f_j(\mathbf{x}')$ . A solution  $\mathbf{x}$  is Pareto optimal if there does not exist any solution  $\mathbf{x}'$  such that  $\mathbf{x}'$  dominates  $\mathbf{x}$ . Multi-objective optimization aims to identify or approximate the set of Pareto optimal solutions, known as the Pareto set. The concept of local optima can be extended to MOPs: A solution  $\mathbf{x}$  is a Pareto local optimal solution if no other solution in its neighborhood dominates it [12].

## III. WALSH TRANSFORM

### A. Walsh Functions

A Walsh function  $\phi_i : [0, 1] \rightarrow \{-1, 1\}$  forms an orthonormal basis for the Hilbert space  $L^2([0, 1])$  [7]. Similar to trigonometric functions in Fourier transforms, Walsh functions can be used as a basis for decomposing pseudo-Boolean functions. They have been used in the field of evolutionary computation since the late 1970s [13]. For a pseudo-Boolean function  $f : \{0, 1\} \rightarrow \mathbb{R}$ , the Walsh function  $\phi_i : \{0, 1\}^N \rightarrow [-1, 1]$  is defined as follows [8]:

$$\phi_{\text{key}(\mathbf{x}')}(\mathbf{x}') = (-1)^{\sum_{j=1}^N x'_j \cdot x_j} \quad (2)$$

where  $\text{key}(\mathbf{x}') = \sum_{i=0}^{N-1} 2^i \times x_i$  is the decimal representation of  $\mathbf{x}'$ . We denote the order of the Walsh basis  $\phi_{\text{key}(\mathbf{x}')}$  as  $o(\mathbf{x}')$ . The order of the basis  $\phi_{\text{key}(\mathbf{x}')}$  is  $o(\mathbf{x}') = \sum_{i=1}^N x'_i$ . For example, the order of the basis  $\phi_{\text{key}(\mathbf{x}')=0}$  corresponding to the bit sequence  $\mathbf{x}' = \{0, 0, \dots, 0\}$  is zero. The Walsh basis  $\phi_{\text{key}(\mathbf{x}')=3}$  corresponds to  $\mathbf{x}' = \{0, \dots, 0, 1, 1\}$  and its order is 2.

## B. Walsh Decomposition

Any pseudo-Boolean function can be converted into a weighted linear sum of Walsh bases as follows:

$$f(\mathbf{x}) = \sum_{i=0}^{2^N-1} w_i \cdot \phi_i(\mathbf{x}), \quad w_i = \frac{1}{2^N} \sum_{\mathbf{x} \in X} f(\mathbf{x}) \cdot \phi_i(\mathbf{x}) \quad (3)$$

The Walsh basis  $\phi_i$  is uniquely defined and does not depend on the function  $f$  to be transformed. By contrast, the value of the coefficient  $w_i$  is dependent on the function. For a Walsh transform to be accurate,  $2^N$  coefficients and  $2^N$  Walsh bases are required. The coefficients of the Walsh basis, i.e. the Walsh coefficients, loosely relate to the interaction strength between a subset of variables in the function. For instance, the Walsh coefficient  $w_6$  and basis  $\phi_6$  of a function with  $N = 4$  and linking to a bit string  $(0, 1, 1, 0)$  and the interaction between variables  $x_2$  and  $x_3$ . A high absolute value of the coefficient  $w_6$  indicates a strong interaction between variables  $x_2$  and  $x_3$ . Conversely, a value near zero signifies no interaction. Therefore, the coefficients of a Walsh basis with order  $o \in \{0, \dots, N\}$  capture the interaction among the  $o$  variables in the function.

Unless most variables interact, the majority of coefficients will be close to zero. Although the number of possible Walsh bases increases exponentially with the number of variables  $N$ , the coefficient of the bases with high order are nearly zero and their effect is often negligible. Therefore, by setting an upper limit of order  $O$ , we can limit the number of bases. This low-order Walsh transform can be serve as a surrogate of the original objective function [8]. The low-order function  $f_O$  restricted by the upper bound of order  $O$  is defined as:

$$f_O(\mathbf{x}) = \sum_{\mathbf{x}': o(\mathbf{x}') \leq O} w_{\text{key}(\mathbf{x}')} \cdot \phi_{\text{key}(\mathbf{x}')}(\mathbf{x}) \quad (4)$$

## C. Approximate Walsh Transform

While low-order coefficients allow us to reduce the number of bases, computing exact Walsh coefficients require a full-enumeration of the solution space. Based on previous study [8], we estimate approximate coefficients  $\tilde{w}_{\text{key}(\mathbf{x}')}$  by means of a sample of solutions together with their objective function values. Eq. (3) can be formulated as a linear regression problem with Walsh coefficient  $w_i$  as independent variables and function values  $f(\mathbf{x})$  as dependant variables. We refer to this low-order function with estimated coefficients as the Walsh approximate function. It is defined as follows:

$$f_{O,A}(\mathbf{x}) = \sum_{\mathbf{x}': o(\mathbf{x}') \leq O} \tilde{w}_{\text{key}(\mathbf{x}')} \cdot \phi_{\text{key}(\mathbf{x}')}(\mathbf{x}) \quad (5)$$

where  $A$  is an archive of sampled solutions along with their function values, which serve as the input data for the regression. Following previous studies [8], we employ Lasso regression [14] to estimate the Walsh coefficients, aiming to minimize the number of non-zero coefficients.

## IV. WALSH-BASED MULTI-OBJECTIVIZATION

We propose an approach that does not only include a helper objective, but also an optimizer that multi-objectivize a difficult SOP for solving it. Our method involves two phases: the first phase generates a helper objective from a sample of solutions, while the second phase solves the multi-objectivized problem. This section describes the helper objective based on the Walsh transform and the multi-objectivization optimizer.

### A. Walsh-based Helper Objectives

The Walsh approximate function can be adjusted from a constant function  $f_{O=0,A}(\mathbf{x}) = \tilde{w}_0$  to the target objective function itself, depending on the upper limit of order  $O$  and the size of the archive  $A$ . When  $O = N$  and  $A = \{(\mathbf{x}, f(\mathbf{x})) \mid \mathbf{x} \in X\}$ , the Walsh approximation function perfectly matches the original function. However, under this definition, existing Walsh approximate functions cannot have a negative correlation with the original function. Yet, negative correlations between the target and helper objectives lead to the acquisition of solution diversity [15]. As such, we extend the Walsh approximation function as follows:

$$f'_{O',A}(\mathbf{x}) = \begin{cases} f_{O',A}(\mathbf{x}) & \text{if } O' > 0, \\ 2\tilde{w}_0 - f_{-O',A}(\mathbf{x}) & \text{otherwise.} \end{cases} \quad (6)$$

$\tilde{w}_0$  is the Walsh coefficient corresponding to the bit sequence  $\mathbf{x}' = \{0, 0, \dots, 0\}$ .  $\phi_0(\mathbf{x})$  is 1, regardless of  $\mathbf{x}$ . The correlation can then positive or negative, depending on whether the parameter  $O' \in \{-N, \dots, N\}$  is positive or negative. We propose this extended Walsh approximate function as a new helper function.

### B. Multi-objectivization Search

We consider a single-objective local search as a sampling algorithm to generate the helper objective, while Pareto local search [12] is used as an optimizer for multi-objectivization.

1) *Single-objective Local Search*: First, we describe the local search. The local search maintains the best solution as the current search point  $\mathbf{x}$ . It evaluates its neighbors and updates it with the most improved neighbor. The termination criteria are finding a global optimum  $\mathbf{x}^*$  or reaching a maximum number of evaluations  $FE_{max}$ . Given that local search usually get stuck in a local optimum, we simply restart from a random solution until the stopping condition is satisfied. A pseudo-code of the *multi-start* local search is given in **Alg. 1**. It takes the objective function to be optimized, i.e. the target objective function  $f$ , and a maximum number of evaluations  $FE_{max}$  as input. We store all evaluated solutions in an archive  $A$  such that its size corresponds to the number of evaluations performed. To avoid revisiting a solution that has already been chosen as a search point, we record these points as  $X_{searched}$ . The archive  $A$  is then used to estimate Walsh coefficients and generate a Walsh-based helper objective function.

---

**Algorithm 1: Multi-start local search**

---

```
Procedure Multi-start LS( $f, FE_{max}$ ):  
  Initialize archive and searched points  $A, X_{searched} \leftarrow \emptyset$   
  while  $|A| < FE_{max}$  or  $\mathbf{x}^* \notin X_{searched}$  do  
     $\mathbf{x}$  is randomly selected from  $X$ .  
     $A, X_{searched} \leftarrow \text{LS}(\mathbf{x}, f, FE_{max}, A, X_{searched})$   
  return  $A, X_{searched}$   
  
Procedure LS( $\mathbf{x}, f, FE_{max}, A, X_{searched}$ ):  
   $A \leftarrow A \cup \{(\mathbf{x}, f(\mathbf{x}))\}$   
  while  $\mathbf{x} \notin X_{searched}$  or  $|A| < FE_{max}$  do  
     $X_{searched} \leftarrow X_{searched} \cup \{\mathbf{x}\}$   
    foreach neighborhood  $\mathbf{x}' \in \mathcal{N}(\mathbf{x})$  do  
       $A \leftarrow A \cup \{(\mathbf{x}', f(\mathbf{x}'))\}$   
     $\mathbf{x} \leftarrow \arg \max_{\mathbf{x} \in \{\mathbf{x}\} \cup \mathcal{N}(\mathbf{x})} f(\mathbf{x})$   
  return  $A, X_{searched}$ 
```

---

2) *Pareto Local Search*: The single-objective local search holds a single best solution, in contrast to Pareto Local Search (PLS) that maintains an unbounded set of candidate search points [12]. The candidate search points are the set of non-dominated solutions found so far during the search. In its standard implementation, PLS selects a solution randomly from the candidates at each iteration [12]. However, in a multi-objectivization scenario, the goal is to find a global optimum for the target objective. As such, we apply a simple modification by selecting the solution with the highest target objective value as the current search point. All its neighbors are evaluated, and dominated solutions are discarded from the set of candidates. The search point  $x$  is then added to  $X_{searched}$ . If there are no remaining candidate search points outside of  $X_{searched}$ , PLS has converged. If some budget remains, we restart the process with a different solution until the stopping condition is met. This restart carries over both the archive  $A$  and the searched point set  $X_{searched}$ . The pseudo-code of multi-objectivization PLS is given in **Alg. 2**.

The proposed optimizer takes the target objective function  $f$ , a maximum number of evaluations  $FE_{max}$ , and two parameters  $\alpha \in ]0, 1]$  and  $O \in \{-N, \dots, N\}$  as input. The parameter  $\alpha$  adjusts the proportion of evaluations devoted to the sampling phase. When  $\alpha$  is 1, the optimizer simply works as a local search, as all evaluations are used for sampling. If a global optimum is found during sampling, the optimizer terminates without performing multi-objectivization PLS. The lower  $\alpha$ , the less budget is devoted to sampling and the more budget is devoted to multi-objectivization PLS. The parameter  $O'$  adjusts the upper bound of the order of the Walsh basis used for the helper objective function and controls the correlation between the helper and the target objectives.

After sampling, a helper objective is generated from the archive  $A$ , as described in Eq. (6). Though we could continuously update the helper function as the archive expands, within this paper we set it at the end of sampling. The function values of the solutions in the archive are updated from  $f(\mathbf{x})$  to  $\mathbf{f}(\mathbf{x}) = (f(\mathbf{x}), f_{O',A}(\mathbf{x}))$ . Lastly, the multi-objectivized function  $\mathbf{f}$  is solved by means of PLS, the number of evaluations for PLS being  $(1 - \alpha) \times FE_{max}$ .

---

**Algorithm 2: Multi-objectivization Pareto local search**

---

```
Procedure Multi-objectivization  
  PLS( $\mathbf{f}, FE_{max}, O', \alpha$ ):  
    Sampling phase:  
     $A, X_{searched} \leftarrow \text{Multi-start LS}(f, FE_{max} \times \alpha)$   
  
    Multi-objectivization phase:  
    if global optimum  $\mathbf{x}^* \notin X_{searched}$  then  
      Generate helper objective  $f'_{O',A}$ . (Eq. (6))  
      Multi-objectivize:  $\mathbf{f}(\mathbf{x}) = (f(\mathbf{x}), f'_{O',A}(\mathbf{x}))$   
      Update archive  $A$ :  
       $A, X_{searched} \leftarrow \text{Multi-start PLS}(f, FE_{max}, A, X_{searched})$   
    return  $A, X_{searched}$   
  
Procedure Multi-start PLS( $\mathbf{f}, FE_{max}, A, X_{searched}$ ):  
  while  $|A| < FE_{max}$  or  $\mathbf{x}^* \notin X_{searched}$  do  
     $\mathbf{x}$  is randomly selected from  $X$ .  
     $A, X_{searched} \leftarrow \text{PLS}(\mathbf{x}, \mathbf{f}, FE_{max}, A, X_{searched})$   
  return  $A, X_{searched}$   
  
Procedure PLS( $\mathbf{x}, \mathbf{f}, FE_{max}, A, X_{searched}$ ):  
  Initialize candidate of search points  $X_c \leftarrow \{\mathbf{x}\}$   
   $X_{c-s} \leftarrow X_c \setminus X_{searched}$   
  while  $X_{c-s} \neq \emptyset$  or  $|A| < FE_{max}$  or  $\mathbf{x}^* \notin X_{searched}$  do  
     $\mathbf{x} \leftarrow \arg \max_{\mathbf{x} \in X_{c-s}} f(\mathbf{x})$   
     $X_{searched} \leftarrow X_{searched} \cup \{\mathbf{x}\}$   
    foreach neighborhood  $\mathbf{x}' \in \mathcal{N}(\mathbf{x})$  do  
       $A \leftarrow A \cup \{(\mathbf{x}', \mathbf{f}(\mathbf{x}'))\}$   
     $X_c \leftarrow \text{non-dominated}(X_c \cup \mathcal{N}_1(\mathbf{x}))$   
     $X_{c-s} \leftarrow X_c \setminus X_{searched}$   
  return  $A, X_{searched}$ 
```

---

## V. EXPERIMENTAL SETUP

We provide the experimental setup of our analysis, which include the benchmark problem setting and the parameters of the proposed optimizer. We conduct two sets of experiments. The first investigates the properties of Walsh-based helper objective functions, while the second benchmarks the performance of the whole multi-objectivization optimizer.

### A. Properties of Helper Objectives

We investigate the properties of Walsh-based helper objectives on small instances. We consider NK landscapes with  $N = 16$  and  $K \in \{2, 3, 4, 6\}$  as target objective functions. For each setting, we generate 50 independent instances, resulting in a total of 200 small-size instances. We also generate a helper objective function  $f'_{O',A}$  with different  $A$  and different parameters  $O'$  for each target objective. The parameters are given in **Table I**. The solution space is fully enumerated for all generated helper objectives in order to examine their correlation with the target objectives and the resulting number of local optima.

### B. Benchmarking of Multi-objectivization Search

We evaluate the impact of the Walsh-based helper objective on relatively small instances with  $N = 20$  and  $K \in \{2, 3, 4\}$ . For each  $K$ , we generate 25 independent instances, that is 75 instances in total. Parameters are given in **Table II**. For benchmarking purposes, we set the parameters more practicality.

TABLE I: Problem parameters for analyzing helper objectives.

description	values
number of variables	$N = 16$
number of interactions	$K \in \{2, 3, 4, 6\}$
parameter of order	$O' \in \{-7, -6, \dots, 7\}$ s.t. $O' \leq K + 1$
archive size	$ A  \in \{3276(5\%), 6553(10\%), 65536(100\%)\}$

TABLE II: Problems and algorithms parameters for benchmarking.

description	values
number of variables	$N = 20$
number of interactions	$K \in \{2, 3, 4\}$
parameter of order	$O' \in \{-3, -2, -1, 0, 1, 2, 3\}$
archive size	$FE_{max} = 5000$
proportion of sampling	$\alpha \in \{0.25, 0.5, 0.75\}$

TABLE III: Considered local features from [16].

features	description
inf.avg	average proportion of dominated neighbors
sup.avg	average proportion of dominating neighbors
inc.avg	average proportion of incomparable neighbors

For instance, generating helper objectives with a high order  $|O'|$  requires regression of a tremendous number of Walsh coefficients, thus the maximum order  $|O'|$  should be 2 or 3. With this setup, we use  $FE_{max} \times \alpha \in \{1250, 2500, 3750\}$  sampled solutions to generate the helper objectives. This is about 0.12-0.36% of the solution space. When  $O' = 0$ , the helper objective function is a constant function  $f_{O'=0,A} = \tilde{w}_0$ . As a result, it has no effect on search. In this case, the proposed optimizer is functionally identical to a multi-start local search.

We conduct 31 independent runs per instance for each parameter combination  $(O', \alpha)$ . We consider the success rate (succ) as a performance measure, defined as the proportion of trials in which a global optimum is found. We note that this global optimum may be found before the multi-objectivization process, i.e. during the sampling phase by local search. We therefore also evaluate the success rate in trials in which multi-objectivization was performed (succ2).

After multi-objectivization, during the archive update step, the features shown in Table 3 are additionally computed from the sampled solution  $X_{searched}$  and the archive  $A$ . The features, proposed in [16], are derived from trajectories of adaptive and random walks. The objective values of each search point's neighbors are already stored in the archive, so no extra evaluations are required to compute these features. They are only computed on trials where multi-objectivization is performed. The average value of features is then assigned to a combination of a target objective  $f$  and parameters  $O'$  and  $\alpha$ .

## VI. RESULTS AND DISCUSSION

### A. Characteristics of Helper Objectives

1) *Correlation between the Target and the Helper Objectives:* The correlation between the target objective  $f$  and

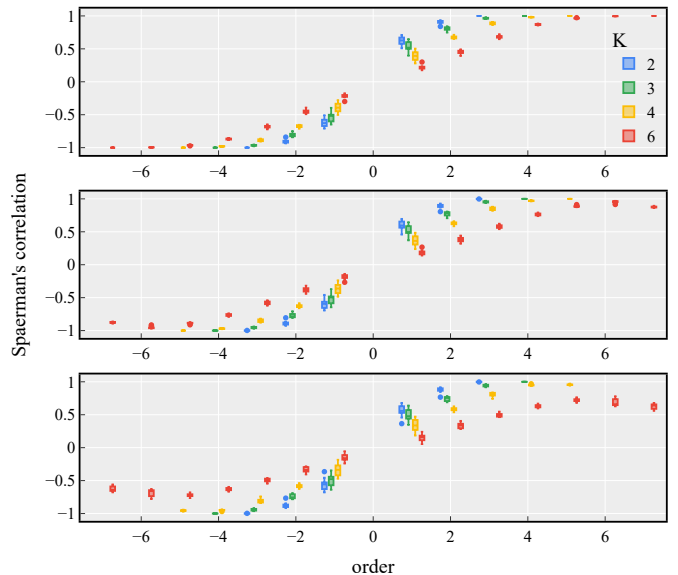


Fig. 1: Correlation between target objectives  $f$  and helper objectives  $f_{O',A}$ . The archive size  $|A|$  is 100% (top), 10% (middle), and 5% (bottom) of the solution space.

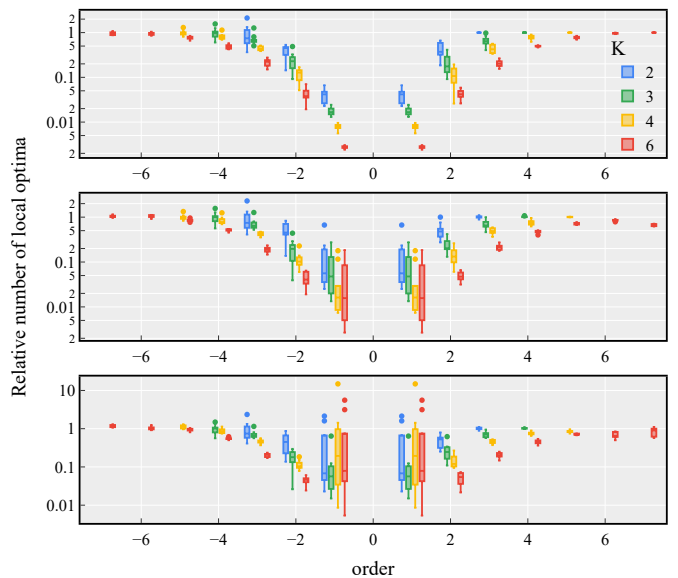


Fig. 2: Relative number of local optima of the helper objective  $f_{O',A}$  to the target objective  $f$ . The archive size  $|A|$  is 100% (top), 10% (middle), and 5% (bottom) of the solution space.

the Walsh-based helper objective  $f'_{O',A}$  is reported in **Fig. 1**. The  $x$ -axis represents the parameter  $O'$ , which determines the upper bound of the Walsh order and whether the correlation is positive or negative. The Spearman's correlation coefficient between the target and helper objective function values for all solutions in  $X$  is shown on the  $y$ -axis. The color of each box indicates the number of co-variables of the NK landscape (target objective). First, **Fig. 1** (top) is for a helper objective with Walsh coefficients obtained from 100% of solutions. The correlation strength increases with the absolute value of the

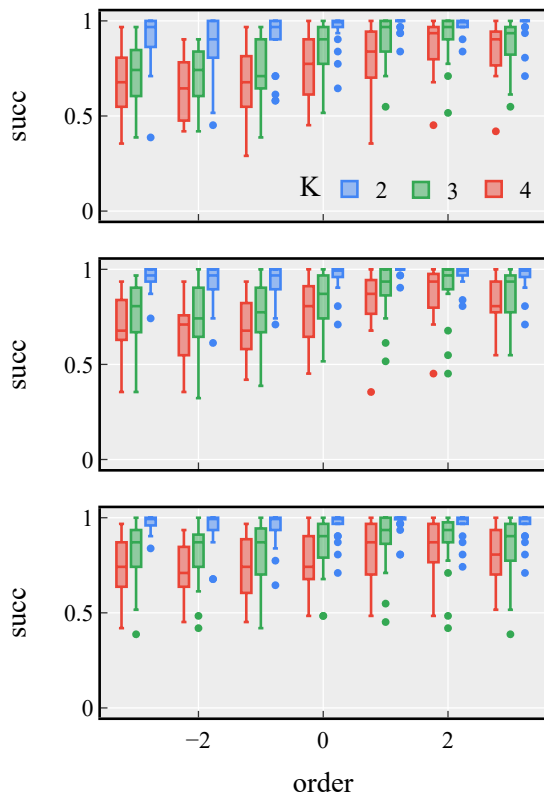


Fig. 3: Success rate of Multi-objectivization PLS with different parameters  $O'$  and  $\alpha$ . The ratio of evaluations devoted to sampling  $\alpha$  is 0.25 (top), 0.5 (middle), and 0.75 (bottom).

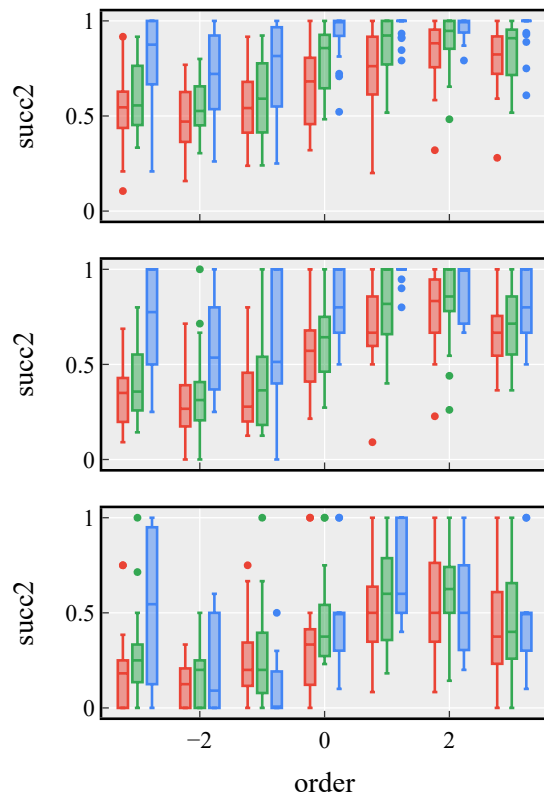


Fig. 4: Percentage of runs for which a global optimum is found after multi-objectivization.  $\alpha$  is 0.25 (top), 0.5 (middle), and 0.75 (bottom).

order  $O'$ . The number of orders required to achieve a correlation of 1 or  $-1$  depends on the value of  $K$ . The condition for fully positive or negative correlation is  $|O'| = K + 1$ , both empirically and theoretically. In NK landscapes, each variable interacts with  $K$  co-variables, meaning there is interaction between the  $K + 1$  variables. The minimal order that can represent an interaction between  $K + 1$  variables is  $K + 1$ , as confirmed by **Fig. 1** (top).

Next, the correlation between the target and the Walsh-based helper objective, estimated using 10% of solutions is shown in **Fig. 1** (middle). For the most part, trends similar to the fully-enumerated case are visible. However, for instances with  $K = 6$  (red box), the correlation peaks at  $|O'| = K = 6$  instead of  $|O'| = K + 1 = 7$ , and does not reach 1 or  $-1$ . As  $K$  increases, the accuracy of estimating the coefficients and approximating the target objective decreases. This is clearer in the results of the helper objective estimated by 5% solutions, shown in **Fig. 1** (bottom). For instances with  $K = 6$ , the maximum absolute value of the correlation coefficient is around 0.6, and even instances with  $K = 4$  do not correlate perfectly with  $|O'| = K + 1$ . These results suggest that approximating a target objective with many interacting variables from a sample of solutions is challenging. However, such a complex objective function is unusual in the real world, and the approximation accuracy does not necessarily reflect the effectiveness of the helper objective for multi-objectivization search.

The number of local optima in the Walsh-based helper objective is reported in **Fig. 2**. The  $y$ -axis gives the proportion of local optima of the helper objective, that is the number of local optima of the helper objective divided by those of the target objective. A proportion of 1 indicates that the helper and target objectives have the same number of local optima. A lower value implies fewer local optima in the helper objective. **Fig. 2** (top) is for a helper objective obtained by 100% of solutions. As the absolute value of the order  $|O'|$  increases, the value approaches 1, and as it decreases, the value approaches 0. This is because the number of local optima in the target objective remains unaffected by the order, leading to a decrease in local optima in the helper objective as parameter  $|O'|$  decreases. This decrease is attributed to the reduction of the basis, cutting off some variable interactions. Weaker interactions between variables typically results in fewer local optima. The low-order Walsh transform relaxes the degree of variable dependency of the target objective.

2) *Multi-modality of the Helper Objective*: The proportion of local optima in the helper objective estimated by 10% of solutions is reported in **Fig. 2** (middle). Here as well, a trend similar to the fully enumerated case can be observed. However, in the helper objective with  $O' \in \{-1, 1\}$ , there are significant variations in the number of local optima. When  $|O'| = 1$ , there is typically a single local optimum, as there is no interaction between variables, similar to NK landscapes with  $K = 0$  or the

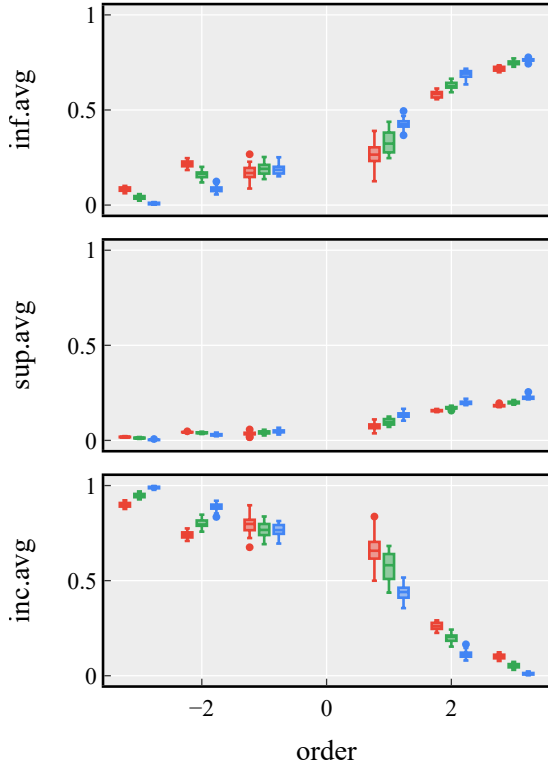


Fig. 5: The value of considered local landscape features with different parameter  $O'$ .

OneMax problem. The increase in the number of local optima is due to the existence of plateaus on the helper objective, which are sub-spaces composed of solutions with equal objective values. We recall that we estimated the Walsh coefficients using Lasso regression, which minimizes the number of non-zero coefficients. If there is a Walsh basis with  $O' = 1$  and zero-valued coefficients  $\tilde{w} = 0$ , the corresponding variables to the basis have no effect on the objective function value. This is more evident in the results for the helper objective estimated by 5% of solutions, shown in **Fig. 2** (bottom). In the case of  $|O'| = 1$ , the number of local optima in the helper objective exceeds that of the target objective. These findings imply that, when estimating coefficients with Lasso regression, a low order not only reduces the number of local optima, but also causes the existence of plateaus. However, this issue could be mitigated by adding noise to the coefficients to ensure they are non-zero.

### B. Benchmarking of Multi-objectivization Search

We now analyze the performance of the entire multi-objectivization search process. The success rate of the multi-objectivization PLS optimizer with varying parameters  $O'$  and  $\alpha$  is depicted in **Fig. 3**. When the order is zero, the helper objective has no influence on search, making it equivalent to the multi-start local search. Therefore, we can assess the effect of the helper objective and multi-objectivization by comparing the results to the case of order  $O' = 0$ . In **Fig. 3** (left and middle), a positive order  $O' > 0$  outperforms the setting of

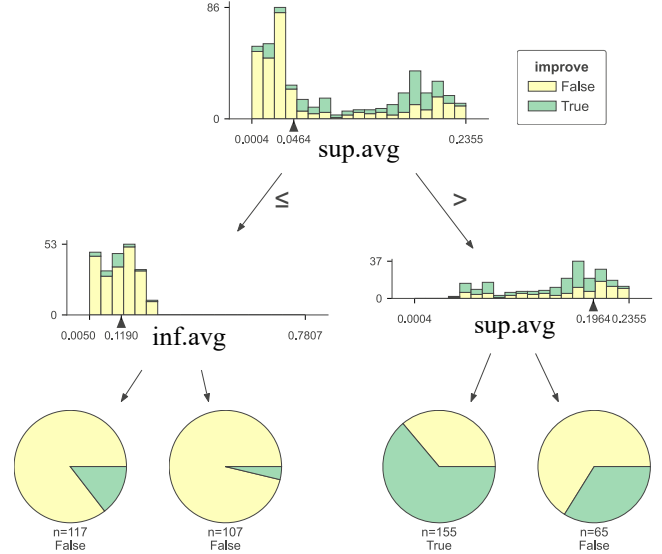


Fig. 6: Recommendations for multi-objectivization based on local landscape features.

$O' = 0$ . Conversely, for  $\alpha = 0.75$ , the increase in success rate due to multi-objectivization, even in positive orders, is tiny compared to the settings of  $\alpha = 0.25$  and  $0.5$ . This is because 75% of the budget of evaluations are devoted to sampling. The performance decay due to multi-objectivization with negative order can be attributed to the growth in the number of solutions the algorithm can move towards. A negative correlation between the objectives weakens the selection pressure and increases solution diversity, but necessitate a large number of evaluations to converge. Success rates that only focus on trials where multi-objectivization was performed are displayed in **Fig. 4**. These results also confirm that helper objectives with a positive correlation achieve a higher success rate. Interestingly, The performance does not peak when  $O' = 3$ , suggesting that a positive correlation is not the sole factor of success. As  $O'$  increases, so does the number of local optima, revealing a trade-off between multi-modality and correlation.

In summary, our experiments reveal that helper objectives with a positive correlation and relaxed variable interactions greatly improve search performance. A small number of evaluations were spent on generating the helper objective: 1250 ( $\approx 0.1\%$ ), 2500 ( $\approx 0.2\%$ ) and 3750 ( $\approx 0.3\%$ ). This suggests that it is possible to generate helper objectives during the optimization process.

### C. Recommendations for the Helper Objective

Finally, we analyse the local features derived from the sample of solutions  $X_{\text{searched}}$ . The ratios of dominating, dominated and incomparable neighbors are reported in **Fig. 5**. These features are significantly influenced by the correlation between the objectives, causing feature values to change based on the order that determines correlation. When there is a strong negative correlation between the objectives, most solutions



are Pareto optimal. These solutions neither dominate nor are dominated, making them incomparable one another. The ratios of dominated neighbors (`inf.avg`) and incomparable neighbors (`inc.avg`) are contrasted. There are more dominating neighbors with increasing  $O'$ , thus facilitating the early discovery of a global optimum.

We have demonstrated the potential of generating a helper objective from sampled solutions. However, the critical decision is to determine whether multi-objectivization is actually necessary. For the 75 SOP instances considered in our experiments, helper objectives were generated from archives of three different sizes and six different orders, excluding  $O' = 0$ . We focus on  $75 \times 6 = 450$  combinations of instances and parameters  $O'$  with  $\alpha = 0.5$ . For each combination, we have the three features `sup.avg`, `inf.avg`, and `inc.avg`. We train a simple CART decision tree [17] to predict whether multi-objectivization would improve performance compared to standard single-objective optimization (case with  $O' = 0$ ). An example of the decision tree using local features is shown in **Fig. 6**. After 30 runs of 10-fold cross-validation, the average classification accuracy is 76%. The decision tree visualization indicates that multi-objectivization is not effective in combinations with low `sup.avg`. These results make sense considering that the parameters  $O'$  and `sup.avg` are correlated. A low `sup.avg` value implies a negative  $O'$ , indicating a negative correlation between the objectives. Similar to the relationship between  $O'$  and the success rate, multi-objectivization does not always prove effective in combinations with high `sup.avg`.

## VII. CONCLUSIONS

In order to clarify the design of a helper objective in multi-objectivization, we proposed a helper objective generator whose correlation with the target objective and the heterogeneity of variable interactions can be adjusted. Our suggested helper objective is grounded on the Walsh transform. From a sample of solutions, the target objective is decomposed into sub-functions, which are then combined to create a helper objective. This Walsh-based helper objective can adjust the correlation and the number of local optima by relaxing the variable interactions of the target objective. We also proposed an optimizer based on Pareto local search for multi-objectivization. Our experimental results suggest that multi-objectivization under a helper objective that is positively correlated and has relaxed variable dependencies facilitates the search for a global optimum. We also computed landscape features from the sample of solutions, predicting with 76% accuracy whether multi-objectivization improves performance.

Future research will focus on adaptive parameter tuning and advanced analysis of landscape features. In particular, there is a need for complementary informative multi-objective landscape features to recommend a suitable helper objective. The features considered in this study are mutually correlated and do not sufficiently explain the successful cases of multi-objectivization. We are considering the Pareto local optimal solutions network [18] to aid in generating the helper objective and to analyze it using a wider set of landscape features.

*Acknowledgements:* This research was partially supported by the University of Lille under the MOBILLEX program, and the Japan Student Service Organization (JASSO) scholarship program. The authors are grateful to Sébastien Verel for his valuable suggestions on the Walsh transform.

## REFERENCES

- [1] J. D. Knowles, R. A. Watson, and D. W. Corne, "Reducing local optima in single-objective problems by multi-objectivization," in *International conference on evolutionary multi-criterion optimization*. Springer, 2001, pp. 269–283.
- [2] V. Steinhoff, P. Kerschke, P. Aspar, H. Trautmann, and C. Grimme, "Multiobjectivization of local search: single-objective optimization benefits from multi-objective gradient descent," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2020, pp. 2445–2452.
- [3] M. T. Jensen, "Helper-objectives: Using multi-objective evolutionary algorithms for single-objective optimisation," *Journal of Mathematical Modelling and Algorithms*, vol. 3, no. 4, pp. 323–347, 2004.
- [4] D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler, "Do additional objectives make a problem harder?" in *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, 2007, pp. 765–772.
- [5] S. Verel, A. Liefvooghe, L. Jourdan, and C. Dhaenens, "On the structure of multiobjective combinatorial search space: Mnk-landscapes with correlated objectives," *European Journal of Operational Research*, vol. 227, no. 2, pp. 331–342, 2013.
- [6] R. Cosson, R. Santana, B. Derbel, and A. Liefvooghe, "Multi-objective NK landscapes with heterogeneous objectives," in *Genetic and Evolutionary Computation Conference (GECCO 2022)*. Boston, MA, USA: ACM, 2022, pp. 502–510.
- [7] J. L. Walsh, "A closed set of normal orthogonal functions," *American Journal of Mathematics*, vol. 45, no. 1, pp. 5–24, 1923.
- [8] S. Verel, B. Derbel, A. Liefvooghe, H. Aguirre, and K. Tanaka, "A surrogate model based on walsh decomposition for pseudo-boolean functions," in *International conference on parallel problem solving from nature*. Springer, 2018, pp. 181–193.
- [9] J. Handl, S. C. Lovell, and J. Knowles, "Multiobjectivization by decomposition of scalar cost functions," in *Parallel Problem Solving from Nature (PPSN X)*. Dortmund, Germany: Springer, 2008, pp. 31–40.
- [10] X. Ma, Z. Huang, X. Li, Y. Qi, L. Wang, and Z. Zhu, "Multiobjectivization of single-objective optimization in evolutionary computation: a survey," *IEEE Transactions on Cybernetics*, 2021.
- [11] S. Kauffman and S. Levin, "Towards a general theory of adaptive walks on rugged landscapes," *Journal of theoretical Biology*, vol. 128, no. 1, pp. 11–45, 1987.
- [12] L. Paquete, T. Schiavinotto, and T. Stützle, "On local optima in multiobjective combinatorial optimization problems," *Annals of Operations Research*, vol. 156, no. 1, pp. 83–97, 2007.
- [13] A. D. Bethke, *Genetic algorithms as function optimizers*. University of Michigan, 1980.
- [14] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 58, no. 1, pp. 267–288, 1996.
- [15] K. Deb and A. Saha, "Finding multiple solutions for multimodal optimization problems using a multi-objective evolutionary approach," in *Proceedings of the 12th annual conference on genetic and evolutionary computation*, 2010, pp. 447–454.
- [16] A. Liefvooghe, F. Daolio, S. Verel, B. Derbel, H. Aguirre, and K. Tanaka, "Landscape-aware performance prediction for evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 6, pp. 1063–1077, 2019.
- [17] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. Andover, UK: Taylor & Francis, 1984.
- [18] A. Liefvooghe, B. Derbel, S. Verel, M. López-Ibáñez, H. Aguirre, and K. Tanaka, "On pareto local optimal solutions networks," in *Parallel Problem Solving from Nature—PPSN XV: 15th International Conference, Coimbra, Portugal, September 8–12, 2018, Proceedings, Part II 15*. Springer, 2018, pp. 232–244.