Pareto local optimal solutions networks with compression, enhanced visualization and expressiveness

Arnaud Liefooghe, Gabriela Ochoa, Sébastien Verel, Bilel Derbel

Pareto Local Optimal Solutions Networks with Compression, Enhanced Visualization and Expressiveness

Arnaud Liefooghe
arnaud.liefooghe@univ-lille.fr
Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL
F-59000 Lille, France

Gabriela Ochoa
gabriela.ochoa@stir.ac.uk
University of Stirling
Stirling, FK9 4LA, UK

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

Bilel Derbel
bilel.derbel@univ-lille.fr
Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL
F-59000 Lille, France

ABSTRACT

The structure of local optima in multi-objective combinatorial optimization and their impact on algorithm performance are not yet properly understood. In this paper, we are interested in the representation of multi-objective landscapes and their multi-modality. More specifically, we revise and extend the network of Pareto local optimal solutions (PLOS-net), inspired by the well-established local optima network from single-objective optimization. We first define a compressed PLOS-net which allows us to enhance its perception while preserving the important notion of connectedness between local optima. We then study an alternative visualization of the (compressed) PLOS-net that focuses on good-quality solutions, improves the distinction between connected components in the network, and generalizes well to landscapes with more than 2 objectives. We finally define a number of network metrics that characterize the PLOS-net, some of them being strongly correlated with search performance. We visualize and experiment with small-size multi-objective nk-landscapes, and we disclose the effect of PLOS-net metrics against well-established multi-objective local search and evolutionary algorithms.

CCS CONCEPTS

• **Theory of computation** → *Optimization with randomized search heuristics*; *Design and analysis of algorithms*; • **Applied computing** → *Multi-criterion optimization and decision-making*.

ACM Reference Format:
Arnaud Liefooghe, Gabriela Ochoa, Sébastien Verel, and Bilel Derbel. 2023. Pareto Local Optimal Solutions Networks with Compression, Enhanced Visualization and Expressiveness. In *Genetic and Evolutionary Computation Conference (GECCO '23), July 15–19, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3583131.3590474

Reproducibility
Relevant data is available at the following URL: https://doi.org/10.5281/zenodo.8010644

1 INTRODUCTION

Although evolutionary and local search algorithms are often methods of choice for single- and multi-objective (black-box) optimization, it remains difficult to recommend which algorithm to choose from the plethora of available methods in order to solve a given problem. Landscape analysis and visualization are becoming increasingly popular to gain a better understanding of optimization problems and of the challenges that search heuristics face while solving them [15, 21, 22, 26]. However, compared to the rich literature on single-objective landscapes, the study of multi-objective landscapes is more scarce. For visualization, the Pareto front or its approximations can be visualized by means of scatter plots for 2 or 3 objectives, or by means of dimensionality reduction for 3+ objectives [32]. Meanwhile, for representing the whole landscape, there exist some works on multi-objective continuous optimization, such as cost landscapes [10], gradient field heatmaps [14], local dominance landscapes [9], or the plot of landscapes with optimal trade-offs [27]. In the combinatorial case, the well-established local optima networks [23, 24] have been extended to multi-objective optimization with the Pareto local optimal solutions networks (PLOS-net) [20] and the Pareto local optima networks (PLON) [8]. These multi-objective extensions have proven to be useful tools to better understand multi-objective landscapes [9, 28].

In this paper, we focus on the PLOS-net model [20] and we extend its expressiveness by following three complementary axes:

(1) We introduce the compressed PLOS-net (C-PLOS-net). Our main motivations are that (i) uncompressed PLOS-nets often turn out to be too large to visualize, and (ii) the structure of their connected components were found to be strongly correlated with search performance [20], both in terms of runtime and approximation quality.
(2) We improve the visualization of PLOS-nets and C-PLOS-nets under the so-called *rank layout*, and we show that it provides a better representation of some network's structural properties, especially for more than 2 objectives.
(3) We propose and investigate new network metrics, some of which are found to strongly correlate with performance.

We also present and discuss the visualization of C-PLOS-nets for a wide range of small-size multi-objective combinatorial landscapes, and we correlate C-PLOS-net metrics with search, showing their strong effect on algorithm performance and on algorithm recommendation according to network metrics.

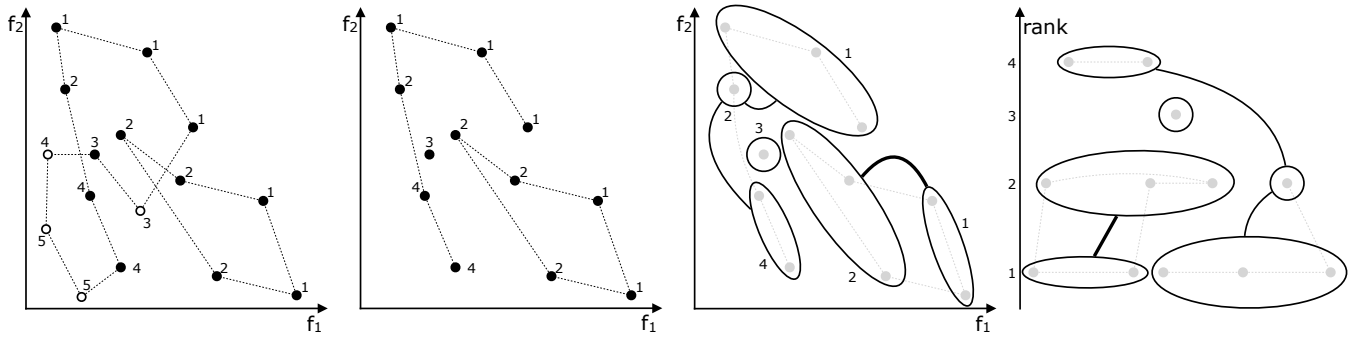

Figure 1: Construction of a compressed PLOS-net (C-PLOS-net) on a toy example: full landscape in the objective space (far-left), PLOS-net (middle-left), C-PLOS-net under the objective-space layout (middle-right), and under the rank layout (far-right).

The rest of the paper is organized as follows. In Section 2, we introduce the necessary background on multi-objective optimization and landscapes. In Section 3, we detail our contributions in terms of compressed network, network visualization and network metrics. In Section 4, we give the setup of our experimental analysis. In Section 5, we present and discuss the results of our experiments. In the last section, we conclude the paper and discuss further research.

2 BACKGROUND

This section provides definitions for multi-objective combinatorial optimization, multi-objective landscapes and PLOS-nets, and presents the considered benchmark problems.

2.1 Multi-objective Optimization

We assume an m-dimensional objective function vector $f: X \to Z$ is to be maximized, such that every solution from the (discrete) solution space $x \in X$ maps to a vector in the objective space $z \in Z$, with $z = f(x)$ and $Z \subseteq \mathbb{R}^m$. Given two objective vectors $z, z' \in Z$, z is dominated by z' if $z_i \leqslant z'_i$ for all $i \in \{1, \ldots, m\}$, and there is a $j \in \{1, \ldots, m\}$ such that $z_j < z'_j$. Likewise, a solution $x \in X$ is dominated by $x' \in X$ if $f(x)$ is dominated by $f(x')$. An objective vector $z^\star \in Z$ is non-dominated if there is no $z \in Z$ such that $z^\star$ is dominated by z. A solution $x^\star \in X$ is Pareto optimal if $f(x)$ is non-dominated. The set of Pareto optimal solutions is the Pareto set, and its mapping in the objective space is the Pareto front. Multi-objective evolutionary and local search algorithms aim at identifying a Pareto set approximation that is to be presented to the decision maker for further consideration [4, 6].

2.2 ρmnk-Landscapes

In terms of benchmark, we consider ρmnk-landscapes [33] as multi-objective multi-modal problems with objective correlation. They extend single-objective nk-landscapes [13] and multi-objective nk-landscapes with independent objectives [1, 16]. Candidate solutions are binary strings of size n. The objective function vector $f = (f_1, \ldots, f_i, \ldots, f_m)$ is defined as $f: \{0,1\}^n \to [0,1]^m$ such that each objective f_i is to be maximized. The objective 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. The contribution of x_j depends on its own value and on the values of $k < n$ variables other than x_j. The variable 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 [13]. By increasing k, landscapes can be gradually tuned from smooth to rugged. In ρmnk-landscapes, the contribution values follow a multivariate uniform distribution such that $\rho > \frac{-1}{m-1}$ defines the correlation among the objectives [33]. The positive (resp. negative) correlation ρ decreases (resp. increases) the degree of conflict between the objective values. Interestingly, ρmnk-landscapes exhibit different characteristics and degrees of difficulty for multi-objective algorithms; see, e.g., [5, 19].

2.3 Multi-objective Landscape

A multi-objective landscape is a triplet $(X, \mathcal{N}, f)$ such that X is the solution space, $\mathcal{N}: X \mapsto 2^X$ is a neighborhood relation, $f: X \to Z$ is an objective function vector. An example of a multi-objective landscape pictured in the objective space is shown in Fig. 1 (far-left). For ρmnk-landscapes, we define the solution space $X = \{0,1\}^n$ as the set of binary strings of length n. The neighborhood $\mathcal{N}$ is based on the well-established *1-bit-flip* operator: two solutions are neighbors if the Hamming distance between them is equal to one. Finally, the objective function vector f is as described in Section 2.2.

A solution $x \in X$ is a *Pareto local optimal solution* (PLOS) if it is not dominated by any of its neighbors [25]; i.e. $\forall x' \in \mathcal{N}(x)$, x' does not dominate x. For $m = 1$, we remark that this is equivalent to the conventional definition of a single-objective local optimum. PLOS are colored in black in Fig. 1 (far-left), whereas non-PLOS appear as empty nodes.

2.4 PLOS-net

Given a multi-objective landscape $(X, \mathcal{N}, f)$, the Pareto local optimal solutions network (PLOS-net) proposed in [20] is constructed as an unweighted graph $G = (N, E)$, such that:

- The set of nodes N are the PLOS,
- There is an edge $e_{ij} \in E$ from x^i to x^j if $x^j \in \mathcal{N}(x^i)$.

An example of a PLOS-net is shown in Fig. 1 (middle-left). We remark that two connected solutions (nodes) in the PLOS-net are mutually non-dominated. In addition, the Pareto (global) optimal solutions are particular nodes of the PLOS-net.

3 COMPRESSION, VISUALIZATION, METRICS

In this section, we detail our contributions in terms of *compressed* PLOS-nets, network visualization and network metrics.

3.1 Compressed PLOS-nets

For constructing the compressed model, we first add a numerical attribute to the PLOS-net nodes, which gives the rank of the corresponding solution in the landscape. More specifically, all solutions from the search space are organized into different layers of mutually non-dominated solutions, following the non-dominated sorting procedure [12] used, e.g., in NSGA-II [7]. The rank of a solution corresponds to the layer it belongs to, such that a lower rank is better and that a Pareto optimal solution has a rank of 1; see Fig. 1.

Based on this augmented PLOS-net, we construct the compressed PLOS-net (C-PLOS-net for short) by compressing the nodes (1) that are connected and (2) that have the same rank. This resembles the compressed LON in single-objective optimization, where nodes with the same fitness are compressed [24]. A C-PLOS-net is thus a weighted graph $G' = (N', E')$ such that:

- The set of nodes N' are connected components of each PLOS-net's sub-graph induced by the nodes with the same rank r, with $r \in \{1, 2, \ldots\}$.
- There is an edge $e'_{ij} \in E'$ if a PLOS within the compressed node i has a neighbor in the compressed node j.

We finally add the following attributes to the so-constructed graph. The *width* of a C-PLOS-net node is the number of uncompressed nodes it contains. The *weight* of a C-PLOS-net edge is the normalized number of uncompressed edges it contains. The edge *distance* is simply the multiplicative inverse of its weight. An example of a C-PLOS-net is shown in Fig. 1 (middle-right and far-right) under two different visualization layouts that are discussed below.

3.2 Network Visualization

Visualizing networks offers an intuitive insight into their structure, which is not easily captured with statistical metrics only. The most familiar forms of network visualization are node-edge diagrams, which assign nodes to points in the 2-D Euclidean space and connect adjacent nodes by straight lines or curves. Relevant network features can be highlighted with nodes and edges decorations such as size and color. Graphs are mathematical objects, they do not have a unique visual representation. The graph-layout accounts for the positions of nodes in the 2-D Euclidean space, and it is key for appreciating the network structure.

We propose using two alternative layouts as shown in Fig. 2. For two-objective problems, the objective space offers a natural graph layout that allows appreciating the configuration of PLOS. This layout, however, does not scale to many objectives. Another disadvantage of the objective-space layout is that it occludes the presence of disconnected components in the graph. The number of connected components, as well as the size of the largest component were found to have a strong predictive power according to the work proposing PLOS-nets [20].

These limitations of the objective-space layout prompted us to design an alternative layout based on the solutions ranks. The idea is to use the logarithm of the ranks as the y coordinate, so the best solutions are at the bottom of the figure (see the bottom plots in Fig. 2). We find the logarithm scale preferable as it more clearly differentiates the solutions with low ranks. For the x coordinate, we rely on the existing graph-visualization body of knowledge. The most popular graph layout algorithms define an energy or cost function based on a simulated physical model of the graph. Minimizing this function produces an optimal drawing. We used a stress majorization layout algorithm [11], and a color gradient to highlight the solutions ranks. In the compressed model (right plots in Fig. 2), the size of nodes and the darkness of edges are proportional to how many solutions and transitions they aggregate, respectively. This rank layout can be extended to many objectives and allows appreciating the graph disconnected components. Our graph visualizations were produced using the igraph and ggraph packages of the R programming language.

3.3 Network Metrics

Because of their roots in graph theory and complex networks, the PLOS-net and C-PLOS-net models allow us to define informative structural metrics and statistics. Looking at the network as a mathematical object, we define and analyze a number of network metrics inspired by previous studies from single-objective LONs [23, 24] while taking into account the dominance relation when necessary, and hence accommodating the multi-objective nature of the considered landscapes. These network metrics are presented in Table 1.

Table 1: Description of network metrics.

	metric	description
uncompressed and compressed networks	node_n	proportion of nodes
	node_pareto_n	proportion of Pareto nodes (nodes with rank 1)
	node_adj_pareto_n	proportion of nodes adjacent to a Pareto node
	node_rank_worst	maximum (worst) node rank
	degree_avg	average degree of nodes
	rank_degree_cor	node rank-vs-degree correlation
	isolated_n	proportion of isolated nodes
	pareto_isolated_n	proportion of Pareto nodes that are isolated
	isolated_rank_avg	average rank of isolated nodes
	edge_density	density of edges
	assort_degree	assortativity by degree
	cc_n	proportion of connected components (cc)
	cc_max	size of largest cc
	cc_avg	average size of cc
	cc_max_pareto	size of largest cc that contains a Pareto node
	cc_pareto_max	(average) size of cc with most Pareto nodes
	cc_pareto_avg	average number of Pareto nodes per cc
	cc_rank_avg_avg	mean of average rank per cc
	cc_rank_best_avg	mean of best rank per cc
	path_length_avg	average path length
	path_length_max	longest path length (diameter)
	path_pareto_exist	number of nodes connected to a Pareto node
	path_pareto_avg	avg. nb. of Pareto nodes a node is connected to
	path_length_pareto_avg	avg. (existing) path length to a Pareto node
compressed networks	node_width_avg	average node width
	node_cmpr	compression rate over nodes
	strength_avg	average node strength
	strength_pareto	sum of strengths of Pareto nodes
	rank_strength_cor	node rank-vs-strength correlation
	edge_weight_avg	average edge weight
	edge_cmpr	compression rate over edges
	dist_avg	average distance
	dist_max	longest distance
	dist_pareto_avg	avg. dist. to Pareto nodes (existing paths)

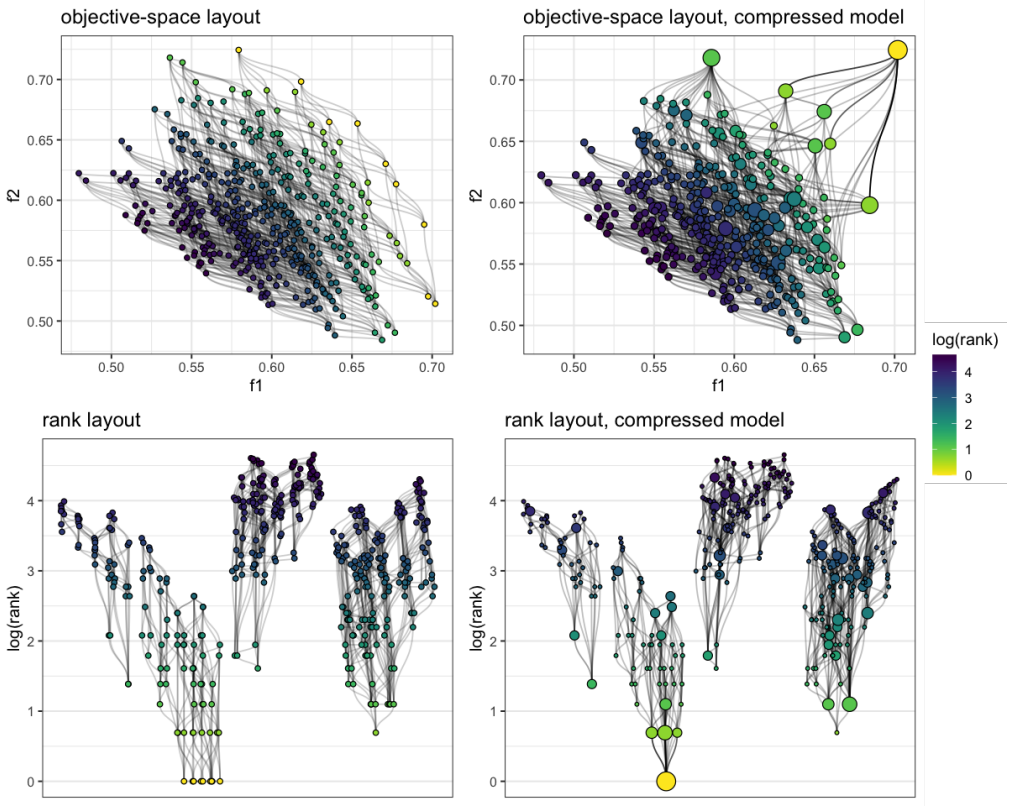

Figure 2: PLOS-net visualization for an example ρmnk-landscape with $\rho = 0.4$, $m = 2$, $n = 16$, and $k = 1$. The compressed model (right plots) is contrasted with the standard PLOS-net model. Two alternative graph layouts are also contrasted. The top plots use the objective values as the x and y coordinates. The bottom plots use the logarithm of the solution ranks as the y coordinates, with the x coordinate following a force-directed graph layout. In the compressed model, the size of nodes and the darkness of edges are proportional to how many solutions and transitions they aggregate.

The top part of the table lists metrics for both model types, while the bottom part lists those for compressed networks only. Node-related metrics include the proportional number of nodes, their relationship with Pareto optimal solutions and dominance ranks, their degree, and whether they are *isolated* (i.e. nodes with no outgoing edges). Edge-related metrics include the density of edges, the tendency for nodes to be connected to nodes with a similar degree (assortativity), together with a number of statistics accounting for the graph connectedness and connected components (starting with cc_) and existing paths between nodes (starting with path_). At last, metrics for compressed networks only are based on the C-PLOS-net attributes. They include the average width of nodes (i.e. the number of solutions in a compressed node), the compression rate of nodes (i.e. the average ratio of solutions in a compressed node) and of edges, the strength of nodes (i.e. the sum of weights of edges adjacent to nodes) and of Pareto nodes, the average weight of edges, and finally path distance metrics among nodes and to Pareto nodes. The network metrics were computed using the igraph R package. We remark that a subset of metrics are considered in the PLOS-net original paper [20] as well, that is node_n, degree_avg, cc_n, cc_max, assort_degree, path_length_avg, path_pareto_exist, and path_length_pareto_avg.

4 EXPERIMENTAL SETUP

This section describes the experimental setup of our analysis, covering the considered problems, algorithms, and their parameters.

4.1 Benchmark Problems

In terms of benchmarks, we generate 520 ρmnk-landscapes (Section 2.2) following the parameters listed in Table 2. This allows us to investigate relatively smooth to relatively rugged landscapes, with two and three objectives, and with conflicting, uncorrelated or correlated objectives. We consider small landscapes so that the they can be exhaustively enumerated, and thus avoid sampling bias.

Table 2: Benchmark parameters (10 instances are randomly generated for each parameter combination).

description	values
number of variables	$n = 16$
number of interactions	$k \in \{0, 1, 2, 4\}$
number of objectives	$m \in \{2, 3\}$
objective correlation	$\rho \in \{-0.7, -0.4, -0.2, 0.0, 0.2, 0.4, 0.7\}$ s.t. $\rho > \frac{-1}{m-1}$

4.2 Algorithms

We consider the following multi-objective algorithms, that are often used for ρmnk-landscapes [5, 19].

4.2.1 PLS. Pareto local search (PLS) [25] is a multi-objective local search that maintains an unbounded archive A of mutually non-dominated solutions. The archive is initialized with a random solution. At each iteration, one solution is selected at random from the archive $x \in A$, and its neighbors are evaluated. For ρmnk-landscapes, the neighbors $\mathcal{N}(x)$ are solutions located at a Hamming distance 1. Dominated solutions are filtered and non-dominated solutions from $A \cup \mathcal{N}(x)$ are stored in the archive. The current solution x is thus tagged as *visited* in order to avoid re-evaluating its neighbors in subsequent iterations. This process naturally stops once all solutions from the archive are tagged as *visited*.

4.2.2 G-SEMO. The global simple evolutionary multi-objective optimizer (G-SEMO) is a an elitist steady-state multi-objective evolutionary algorithm [17]. As PLS, it maintains an unbounded archive A and selects one solution $x \in A$ at random at each iteration. Each binary variable from x is thus independently flipped with a rate of $1/n$ in order to produce an offspring x'. The archive is then updated by keeping the non-dominated solutions from $A \cup \{x'\}$. This process is iterated until a stopping condition is satisfied. While PLS is based on the exploration of the whole 1-bit-flip neighborhood, G-SEMO uses a stochastic bit-flip mutation. As such, every iteration has a non-zero probability to reach any solution from the search space.

4.2.3 NSGA-II. The non-dominated sorting genetic algorithm II (NSGA-II) is an elitist dominance-based multi-objective evolutionary algorithm using dominance for selection [7]. At a given iteration t, the current population P_t is merged with its offspring Q_t, and is divided into non-dominated fronts $\{F_1, F_2, \dots\}$ based on the non-dominated sorting procedure [12]. The front in which a given solution belongs to gives its rank within the population. This is precisely the ranking used as an attribute of nodes in the considered C-PLOS-net. In NSGA-II, crowding distance is also calculated within each front. Selection is based on dominance ranking, and crowding distance is used as a tie breaker. Survival selection consists in filling the new population P_{t+1} with solutions having the best ranks. In case a front F_i overfills the population size, the required number of solutions from F_i are chosen based on their crowding distance. Parent selection for reproduction consists of binary tournaments between randomly-chosen solutions. We employ one-point crossover and stochastic bit-flip mutation for variation. In addition, we record all non-dominated solutions found so far in an unbounded archive.

4.2.4 Parameter Setting. We perform 30 independent runs of each algorithm per instance. For PLS, we are interested in the total number of evaluations performed by the algorithm before it naturally falls into a Pareto local optimum set [25]. For G-SEMO and NSGA-II, the stopping condition is set to 10 000 evaluations (about 15% of the search space). NSGA-II uses a population of size 100. For all algorithms, we are interested in the quality of the final archive, measured in terms of the *Pareto resolution* (reso), i.e. the proportion of Pareto optimal solutions identified by the considered algorithm. In addition, we report the hypervolume [34], and more particularly the relative hypervolume (hv) covered by the final archive with respect to the exact Pareto front. A higher hv value is thus better, and hv = 1 actually means that the exact Pareto front was found. The hypervolume reference point is set to the origin.

5 RESULTS AND DISCUSSION

In this section, we visualize the C-PLOS-net for selected instances, and we investigate the correlation between network metrics and algorithm performance, together with their joint effect on performance prediction and algorithm selection.

5.1 Network Visualization

We start by visually inspecting the C-PLOS-net structure for a set of landscapes with different parameters, using the rank layout introduced in Section 3.2. Fig. 3 shows the C-PLOS-net for selected ρmnk-landscapes with two objectives ($m = 2$), while Fig. 4 does so for instances with three objectives ($m = 3$). In both figures, the sub-plots show instances with $n = 16$, $\rho \in \{-0.4, 0.0, -0.4\}$ (columns 1, 2, and 3, respectively), and $k \in \{1, 4\}$ (top and bottom row, respectively).

The C-PLOS-nets in Fig. 3 show the complete set of (compressed) nodes; this was not possible for the C-PLOS-nets in Fig. 4 with three objectives, due to their large size. Therefore, for three objectives, a sub-graph of the full C-PLOS-net is illustrated. The sub-graphs contain all the nodes with rank 1 and 2, as well as a sample of the remaining nodes selected uniformly at random. The size of the sample is a proportion of the total number of nodes, where the proportions used are $p \in \{0.3, 0.6, 0.8\}$ for $\rho \in \{-0.4, 0.0, -0.4\}$, respectively. The shape of C-PLOS-nets varies depending on problem parameters, and the following general observations can be made from Figs. 3 and 4:

- For both values of m and k, the number of nodes decreases with increasing ρ (objective correlation).
- For both values of m and k, the number of separated connected components increases with increasing ρ.
- For $m = 2$ and a fixed ρ, the number of nodes increases as k gets larger.
- For all m and ρ values, isolated nodes (visualized as triangles) appear only for $k = 4$.

These observations are consistent with the properties of ρmnk-landscapes [20, 33].

5.2 Network Metrics vs. Performance

We pursue by measuring the impact of network metrics on algorithm performance. Fig. 5 gives the Spearman's rank correlation between each metric and algorithm performance. We consider the number of evaluations (eval) performed by PLS, and the relative hypervolume (hv) reached by PLS, G-SEMO and NSGA-II. Hypervolume values are to be maximized, and thus a positive correlation means that the metric has a positive effect on approximation quality. Due to space restriction, results for the Pareto resolution are omitted, but the trends are similar to those of hypervolume.

Apart from a few minor exceptions, we remark that the metrics for uncompressed and compressed networks follow the same trend overall. The network metrics most (negatively or positively) correlated with algorithm performance are as follows. Firstly, the number of evaluations of PLS increases with the number of nodes and with the degree of connectedness of both the uncompressed and

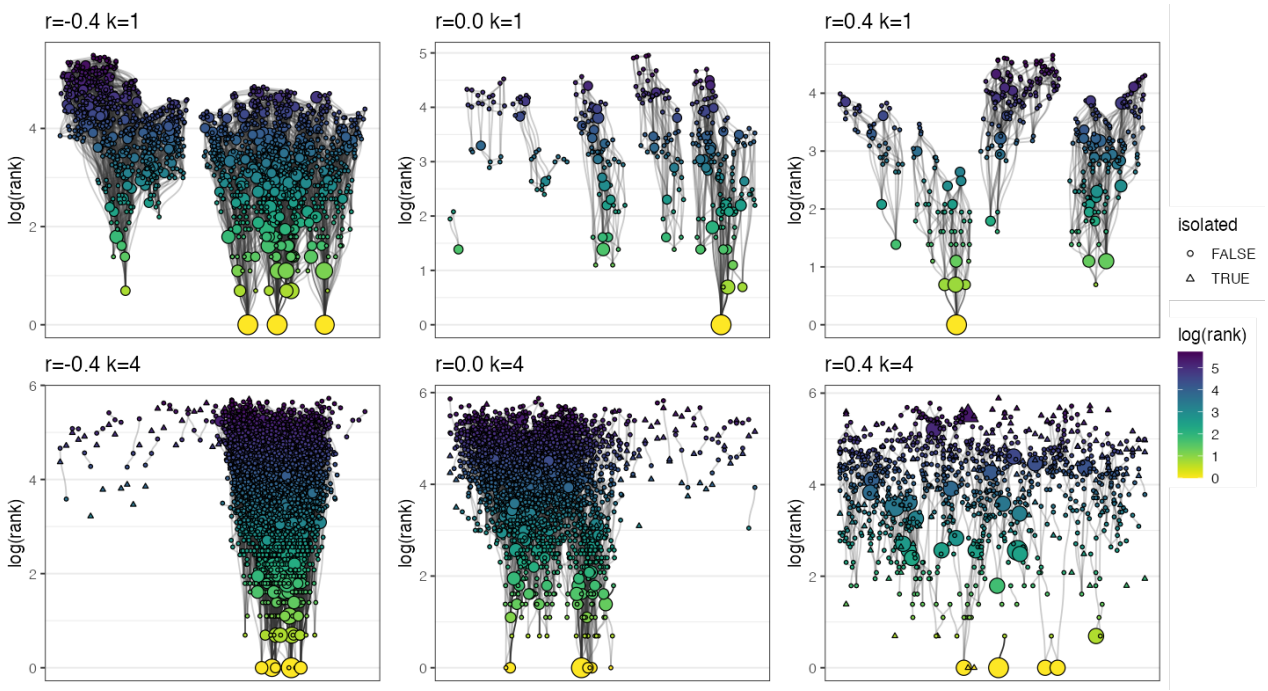

Figure 3: C-PLOS-net visualization for ρmnk-landscapes with two objectives ($m = 2$), $n = 16$, $\rho \in \{-0.4, 0.0, -0.4\}$ and $k \in \{1, 4\}$. The compressed model with the rank layout is used. Node sizes and edge darkness reflect how many solutions and transitions they aggregate.

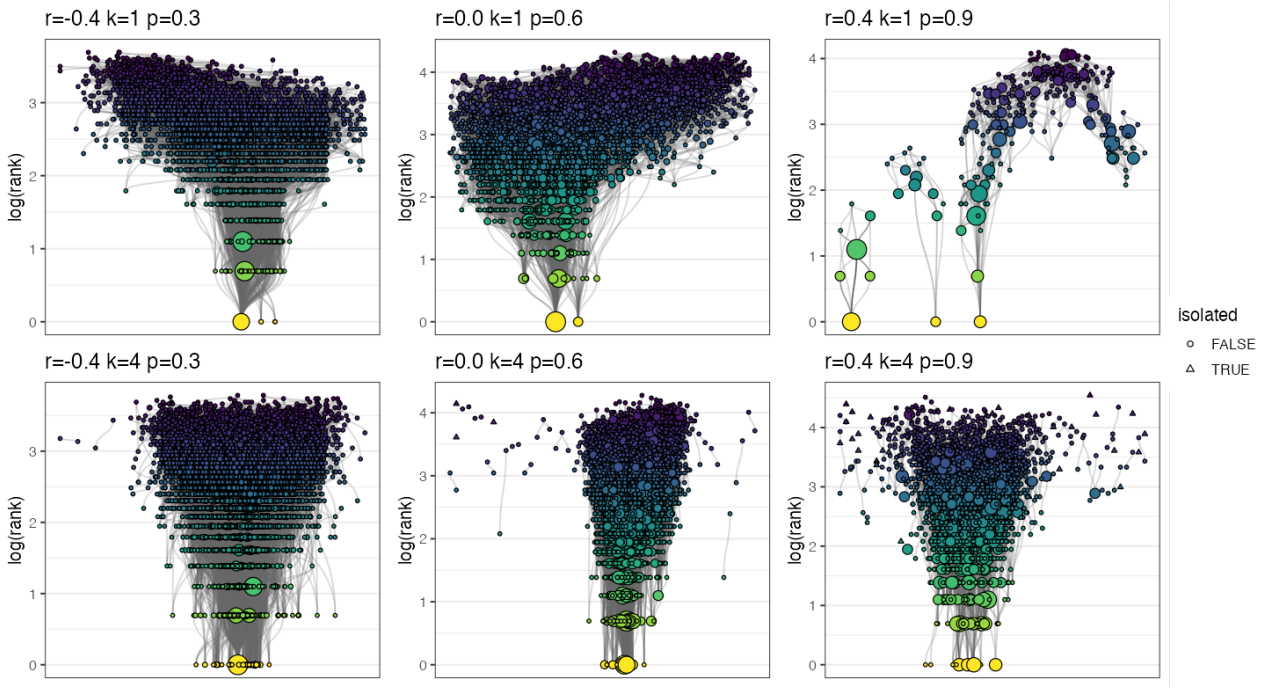

Figure 4: C-PLOS-net visualization for ρmnk-landscapes with three objectives ($m = 3$), $n = 16$, $\rho \in \{-0.4, 0.0, -0.4\}$ and $k \in \{1, 4\}$. The compressed model with the rank layout is used. Since the fully enumerated networks are too large, they were pruned to retain all nodes of rank 1 and 2, completed with a sample of the remaining nodes. The sample size is a proportion p of the complete number of nodes, with $p \in \{0.3, 0.6, 0.9\}$, for $\rho \in \{-0.4.0.0, 0.4\}$, respectively.

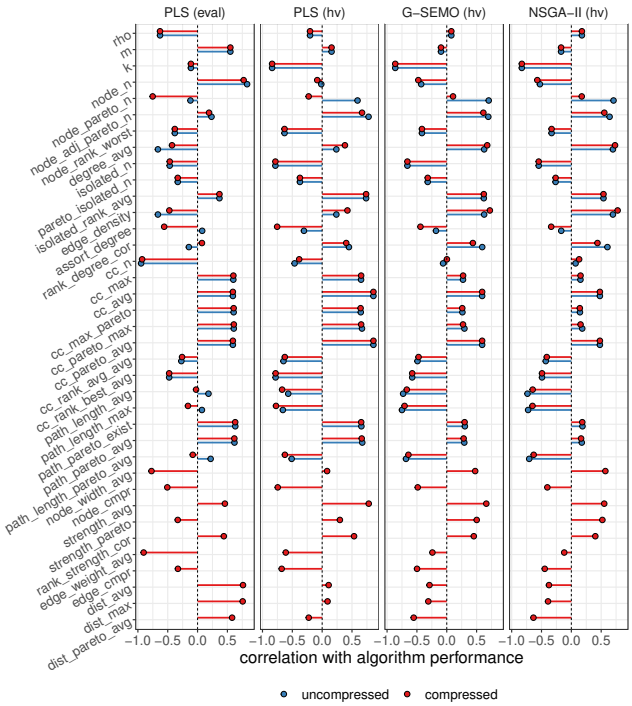

Figure 5: Spearman correlation between network metrics and algorithm performance (number of evaluations for PLS, and hypervolume for all algorithms).

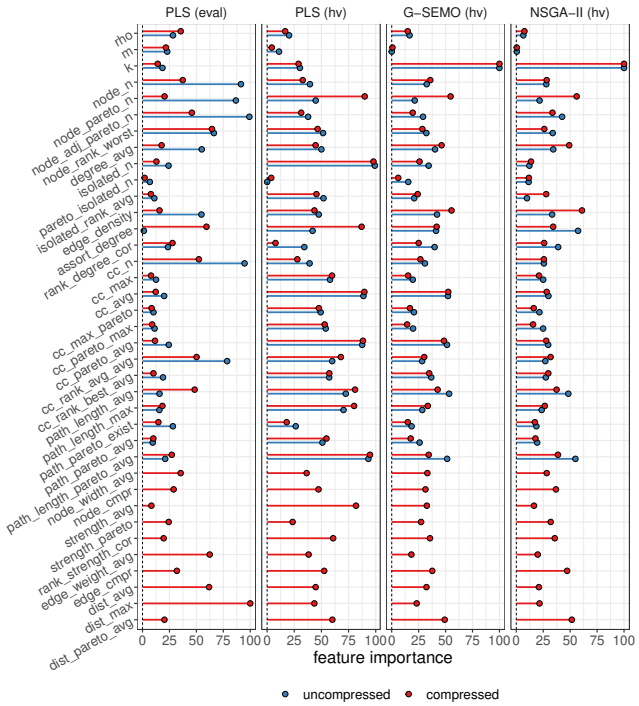

Figure 6: Relative importance of network metrics for predicting algorithm performance (number of evaluations for PLS, and hypervolume for all algorithms).

compressed network models. This suggests that PLS runs longer before getting stuck when there are more PLOS, and when those are more clustered. This results in PLS reaching better hypervolume values when the degree of connectedness of PLOS is high, and when the rank of connected components is low. Conversely, the higher the weight of compressed edges and the lower the distance between compressed nodes, the lower the number of PLS evaluations. Next, the higher the number of isolated nodes and the path length between nodes, the worse the hypervolume achieved by all algorithms. The average node strength in C-PLOS-net, that is, the sum of adjacent edge weights for each node on average, tends to increase the obtained hypervolume. At last, a strong reduction of edges from the uncompressed to the compressed model (low compression rate on the edges) positively affects the performance of all algorithms, whereas the compression rate on the nodes seems more impactful for PLS hypervolume.

5.3 Performance Prediction

We now study the combined effect of network metrics on algorithm performance by building a regression model for predicting the approximation quality of the different algorithms, and the runtime (number of evaluations) of PLS. We rely on well-established random forests [2, 18] with default parameters, using the network metrics as predictors and all considered ρmnk-landscapes for training. Fig. 6 gives the relative importance of predictors, as portrayed by random forests in terms of the mean decrease of prediction accuracy [2, 18]. The higher the value, the more important the predictor.

Once again, the PLS runtime seems mostly influenced by the number of nodes, the number of Pareto nodes (or nodes adjacent to them), and the connectedness of the uncompressed model. For the compressed model, the network diameter (maximum distance between nodes) is clearly the most important metric. Now looking at the hypervolume, the number of compressed nodes, of isolated nodes (for PLS), the connectedness (for PLS and G-SEMO), and the assortativity of nodes by degree are detected as important metrics. So are the path length (for PLS) and the distance (for G-SEMO and NSGA-II) between Pareto nodes, as well as the average compressed node strength (for PLS) and the density of compressed edges (for G-SEMO and NSGA-II). Here as well, results for Pareto resolution are omitted, but they are similar to those for hypervolume.

The above results suggest that there exists a strong connection between the descriptors of both network models and algorithm performance (runtime and approximation quality). We now assess the prediction accuracy of the regression models using 30 replicates of 10-fold cross-validation. In particular, we compare the variance explained by random forests using (1) benchmark parameters only (ρ, m, and k), (2) uncompressed network metrics from [20] (denoted as LDVLAT18), (3) new uncompressed network metrics, (4) compressed network metrics, and finally (5) the combination of all uncompressed and compressed network metrics. This is reported in Fig. 7, this time also including Pareto resolution (reso).

We first observe that the prediction models using solely benchmark parameters as predictors are consistently outperformed by their counterparts using network metrics. This confirms that the

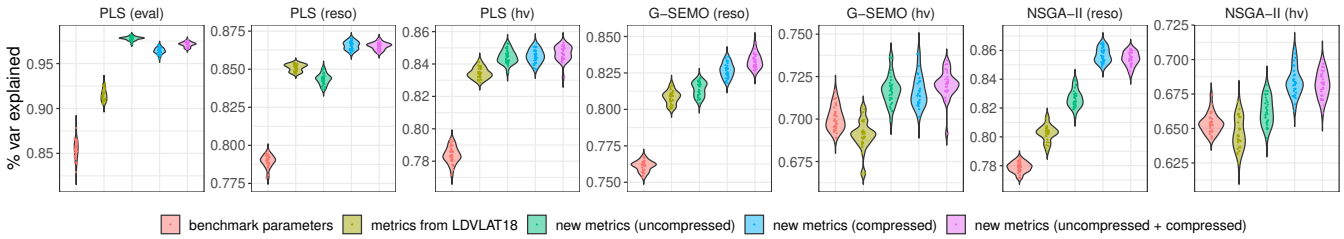

Figure 7: Prediction accuracy of regression models trained by algorithm performance (10-fold cross-validation, 30 replicates).

network models do indeed provide additional information that is not captured by the problem descriptors. Now comparing the network metrics from [20] to those considered in this paper, the latter typically outperform the former except for the Pareto resolution of PLS, where they surpass the uncompressed model's new metrics by a slight margin. In fact, the network models and metrics proposed in this paper lead to an accuracy higher than 0.8 for PLS and for the Pareto resolution of G-SEMO and NSGA-II (slightly less for hypervolume). This suggests that more than 80% of the variance of predicted values across all problems is explained by the prediction model, and thus by network metrics. Finally, the compressed model's metrics, whether considered by themselves or in addition to others, consistently lead to a better accuracy with two exceptions: (i) the PLS runtime, where the uncompressed model's new metrics are above, and (ii) the G-SEMO hypervolume, where there is no significant difference.

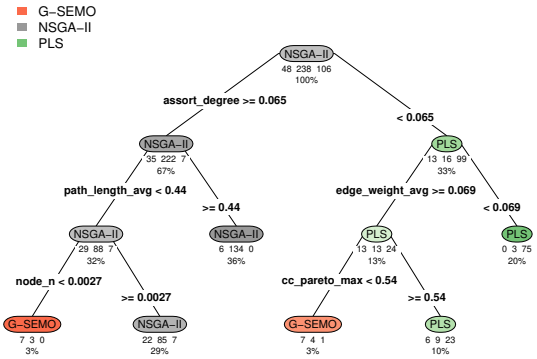

Figure 8: Hypervolume-based algorithm selection for compressed network metrics.

5.4 Algorithm Selection

We finish our study with a simple CART decision tree [3, 29] aiming to recommend which algorithm to select for a given landscape using network metrics as predictors. This is a classification task whose classes are the three considered algorithms, and the correct class for a given instance is the algorithm with the best average hypervolume. Out of the 520 initial instances, 128 were discarded because there was no best algorithm. The dataset thus consists of 392 instances. The decision tree is given in Fig. 8 for compressed network metrics. The numbers below each tree node indicate on how many instances G-SEMO, NSGA-II, and PLS is the best, respectively, followed by the proportion of instances covered by the node.

The cross-validated classification accuracy is 84%. This is substantially better than always selecting NSGA-II, which is better on 61% of instances from this dataset. As such, although the regression is not extremely accurate for hypervolume, the network metrics provide insightful enough information to make an informed decision on which approach to select from the portfolio of algorithms. The decision tree for the Pareto resolution is not reported due to space restriction, but the accuracy is on similar ranges (87%). We observe that the classifier recommends PLS over NSGA-II when the assortativity per degree is low, that is when nodes with similar degree are *not* often connected.

6 CONCLUSION

We revisited the concept of Pareto local optimal solutions network, an extension of the well-established single-objective local optima network for multi-objective landscapes. First, we introduced a compressed model that reduces the number of nodes while maintaining a satisfying level of information, especially regarding the network's connected components. This allowed us to enhance the network visualization under a rank layout, which is particularly relevant for 3+ objectives. However, for landscapes with many local optima, we still had to prune some network nodes in order to visualize it. In the future, we plan to consider rank *bins* (instead of separate ranks) in order to further compress nodes with close ranks, following a logarithmic scale in order to put more emphasis on good solutions. Second, we proposed a new set of network metrics providing insights from an optimization perspective. Some of them were found to strongly correlate with search performance. Interestingly, the compressed model, in addition to providing a clearer visualization, achieves a prediction accuracy as high as the original model, and even surpasses it for some prediction tasks. We consider to extend the notion of funnels, known to impact the performance of single-objective local search [30], by investigating possible definitions for *rank-based sinks*. Finally, we confronted the network metrics for algorithm selection, showing that even a simple decision tree recommends the right algorithm in nearly 85% of the cases. In order to address larger landscapes , we plan to adapt sampling techniques from single-objective local optima networks [31]. We believe the proposed metrics could be smoothly adapted to sampled networks, although it remains open whether they will achieve a satisfactory prediction accuracy. This would allow us to validate our findings on other, larger and more practical multi-objective problems.

REFERENCES

[1] Hernán Aguirre and Kiyoshi Tanaka. 2007. Working principles, behavior, and performance of MOEAs on MNK-landscapes. *European Journal of Operational Research* 181, 3 (2007), 1670–1690.

[2] Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (2001), 5–32.

[3] Leo Breiman, Jerome Friedman, Charles J. Stone, and R. A. Olshen. 1984. *Classification and regression trees.* Taylor & Francis, Andover, UK.

[4] Carlos A. Coello Coello, Gary B. Lamont, and David A. Van Veldhuizen. 2007. *Evolutionary Algorithms for Solving Multi-Objective Problems* (second ed.). Springer.

[5] Fabio Daolio, Arnaud Liefooghe, Sébastien Verel, Hernán E. Aguirre, and Kiyoshi Tanaka. 2017. Problem Features versus Algorithm Performance on Rugged Multiobjective Combinatorial Fitness Landscapes. *Evolutionary Computation* 25, 4 (2017).

[6] Kalyanmoy Deb. 2001. *Multi-Objective Optimization using Evolutionary Algorithms.* John Wiley & Sons.

[7] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (2002), 182–197.

[8] Jonathan E. Fieldsend and Khulood Alyahya. 2019. Visualising the Landscape of Multi-Objective Problems Using Local Optima Networks. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO 2019.* ACM, Prague, Czech Republic, 1421–1429.

[9] Jonathan E. Fieldsend, Tinkle Chugh, Richard Allmendinger, and Kaisa Miettinen. 2022. A visualizable test problem generator for many-objective optimization. *IEEE Transactions on Evolutionary Computation* 26, 1 (2022), 1–11.

[10] Carlos M. Fonseca and Peter J. Fleming. 1996. On the performance assessment and comparison of stochastic multiobjective optimizers. In *Parallel Problem Solving from Nature, PPSN IV.* Springer, Berlin, Heidelberg, 584–593.

[11] Emden R. Gansner, Yehuda Koren, and Stephen North. 2005. Graph Drawing by Stress Majorization. In *Graph Drawing.* Springer, Berlin, Heidelberg, 239–250.

[12] David E. Goldberg. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, Boston, MA, USA.

[13] Stuart A. Kauffman. 1993. *The Origins of Order.* Oxford University Press.

[14] Pascal Kerschke and Christian Grimme. 2017. An Expedition to Multimodal Multi-objective Optimization Landscapes. In *Evolutionary Multi-Criterion Optimization, EMO 2017 (Lecture Notes in Computer Science, Vol. 10173).* Springer, 329–343.

[15] Pascal Kerschke and Mike Preuss. 2019. Exploratory landscape analysis. In *Genetic and Evolutionary Computation Conference Companion, GECCO 2019.* ACM, Prague, Czech Republic, 1137–1155.

[16] Joshua D. Knowles and David Corne. 2007. Quantifying the Effects of Objective Space Dimension in Evolutionary Multiobjective Optimization. In *Evolutionary Multi-Criterion Optimization Conference, EMO 2007 (Lecture Notes in Computer Science, Vol. 4403).* Springer, Matsushima, Japan, 757–771.

[17] Marco Laumanns, Lothar Thiele, and Eckart Zitzler. 2004. Running time analysis of evolutionary algorithms on a simplified multiobjective knapsack problem. *Natural Computing* 3, 1 (2004), 37–51.

[18] Andy Liaw and Matthew Wiener. 2002. Classification and Regression by randomForest. *R News* 2, 3 (2002), 18–22.

[19] Arnaud Liefooghe, Fabio Daolio, Sébastien Verel, Bilel Derbel, Hernán Aguirre, and Kiyoshi Tanaka. 2020. Landscape-Aware Performance Prediction for Evolutionary Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation* 24, 6 (2020), 1063–1077.

[20] Arnaud Liefooghe, Bilel Derbel, Sébastien Verel, Manuel López-Ibáñez, Hernán Aguirre, and Kiyoshi Tanaka. 2018. On Pareto Local Optimal Solutions Networks. In *Parallel Problem Solving from Nature, PPSN XV.* Springer International Publishing, Cham, 232–244.

[21] Katherine M. Malan. 2021. A Survey of Advances in Landscape Analysis for Optimisation. *Algorithms* 14, 2 (2021), 40.

[22] Katherine M. Malan and Gabriela Ochoa. 2021. Recent advances in landscape analysis for optimisation and learning. In *Genetic and Evolutionary Computation Conference Companion, GECCO 2021.* ACM, Lille, France, 899–917.

[23] Gabriela Ochoa, Marco Tomassini, Sébastien Verel, and Christian Darabos. 2008. A study of NK landscapes' basins and local optima networks. In *Genetic and Evolutionary Computation Conference, GECCO 2008.* ACM, New York, NY, 555–562.

[24] Gabriela Ochoa, Nadarajen Veerapen, Fabio Daolio, and Marco Tomassini. 2017. Understanding Phase Transitions with Local Optima Networks: Number Partitioning as a Case Study. In *Evolutionary Computation in Combinatorial Optimization, EvoCOP 2017 (Lecture Notes in Computer Science, Vol. 10197).* 233–248.

[25] Luís Paquete, Tommaso Schiavinotto, and Thomas Stützle. 2007. On local optima in multiobjective combinatorial optimization problems. *Annals of Operations Research* 156, 1 (2007), 83–97.

[26] Hendrik Richter and Andries Engelbrecht (Eds.). 2014. *Recent Advances in the Theory and Application of Fitness Landscapes.* Springer.

[27] Lennart Schäpermeier, Christian Grimme, and Pascal Kerschke. 2020. One PLOT to Show Them All: Visualization of Efficient Sets in Multi-objective Landscapes. In *Parallel Problem Solving from Nature, PPSN XVI.* Springer International Publishing, Cham, 154–167.

[28] Shoichiro Tanaka, Keiki Takadama, and Hiroyuki Sato. 2022. Impacts of Single-objective Landscapes on Multi-objective Optimization. In *IEEE Congress on Evolutionary Computation, CEC 2022.* IEEE, Padua, Italy, 1–8.

[29] Terry Therneau and Beth Atkinson. 2022. *rpart: Recursive partitioning and regression trees.* R package version 4.1.16.

[30] Sarah L. Thomson, Fabio Daolio, and Gabriela Ochoa. 2017. Comparing Communities of Optima with Funnels in Combinatorial Fitness Landscapes. In *Genetic and Evolutionary Computation Conference, GECCO 2017* (Berlin, Germany). ACM, 377–384.

[31] Sarah L. Thomson, Gabriela Ochoa, Sébastien Verel, and Nadarajen Veerapen. 2020. Inferring Future Landscapes: Sampling the Local Optima Level. *Evolutionary Computation* 28, 4 (2020), 621–641.

[32] Tea Tusar and Bogdan Filipic. 2015. Visualization of Pareto Front Approximations in Evolutionary Multiobjective Optimization: A Critical Review and the Prosection Method. *IEEE Transactions on Evolutionary Computation* 19, 2 (2015), 225–245.

[33] Sébastien Verel, Arnaud Liefooghe, Laetitia Jourdan, and Clarisse Dhaenens. 2013. On the structure of multiobjective combinatorial search space: MNK-landscapes with correlated objectives. *European Journal of Operational Research* 227, 2 (2013), 331–342.

[34] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. 2003. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation* 7, 2 (2003), 117–132.