Local Optima Networks with Escape Edges

Sebastien Verel[2], Fabio Daolio[3], Gabriela Ochoa[1], Marco Tomassini[3]

[1] School of Computer Science, University of Nottingham, Nottingham, UK.
[2] INRIA Lille - Nord Europe and University of Nice Sophia-Antipolis, France.
[3] Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland.

Abstract. This paper proposes an alternative definition of edges (*escape edges*) for the recently introduced network-based model of combinatorial landscapes: *Local Optima Networks (LON)*. The model compresses the information given by the whole search space into a smaller mathematical object that is the graph having as vertices the local optima and as edges the possible weighted transitions between them. The original definition of edges accounted for the notion of transitions between the basins of attraction of local optima. This definition, although informative, produced densely connected networks and required the exhaustive sampling of the basins of attraction. The alternative escape edges proposed here do not require a full computation of the basins. Instead, they account for the chances of escaping a local optima after a controlled mutation (e.g. 1 or 2 bit-flips) followed by hill-climbing. A statistical analysis comparing the two LON models for a set of NK landscapes, is presented and discussed. Moreover, a preliminary study is presented, which aims at validating the LON models as a tool for analyzing the dynamics of stochastic local search in combinatorial optimization.

1 Introduction

The performance of heuristic search algorithms crucially depends on the structural aspects of the spaces being searched. An improved understanding of this dependency, can facilitate the design and further successful application of these methods to solve hard computational search problems. Local optima networks (LON) have been recently introduced as a novel model of combinatorial landscapes [10, 11]. This model allows the use of complex network analysis techniques [7] in connection with the study of fitness landscapes and problem difficulty in combinatorial optimization. The model is based on the idea of compressing the information given by the whole problem configuration space into a smaller mathematical object, which is the graph having as vertices the optima configurations of the problem and as edges the possible transitions between these optima. This characterization of landscapes as networks has brought new insights into the global structure of the landscapes studied, particularly into the distribution of their local optima. Moreover, some network features have been found to correlate and suggest explanations for search difficulty on the studied domains.

The definition of the edges in the LON model critically impacts upon its descriptive power with regards to heuristic search. The initial definition of edges in [10, 11], *basin-transition* edges, accounted for the notion of transitions between the local optima basins' frontiers. This definition, although informative, produces highly connected networks and requires the exhaustive sampling of the basins of attraction. We explore in this article an alternative definition of edges, which we term *escape* edges, that does not require a full computation of the basins. Instead, the edges account for the chances (of a prospective heuristic search algorithm) of escaping a local optima after a controlled mutation (e.g. 1 or 2 bit-flips) followed by hill-climbing. This new definition produces less dense and easier to build LONs, which are more amenable to sampling and get us closer to a fitness landscape model that can be used to understand (and eventually exploit) the dynamics of local search on combinatorial problems.

The first goal of the present study is to compare and explore the relationships between the two LON models, based on (i) basin-transition edges and (ii) escape edges, respectively. Thereafter, we present a preliminary study that aims at validating the LON models in their descriptive power of the dynamics of stochastic local search algorithms. We conduct this validation by considering the behavior of two well-known stochastic local search heuristics, namely, Tabu Search [4] and Iterated Local Search [6]. The well known family of NK landscapes [5] is used in our study.

The article is structured as follows. Section 2, includes the relevant definitions and algorithms for extracting the LONs. Section 3, describes the experimental design, and reports a comparative analysis of the extracted networks of the two models. Section 4, presents our model validation study. Finally, section 5 discusses our main findings and suggest directions for future work.

2 Definitions and algorithms

A Fitness landscape [9] is a triplet (S, V, f) where S is a set of potential solutions i.e. a search space, $V : S \longrightarrow 2^S$, a neighborhood structure, is a function that assigns to every $s \in S$ a set of neighbors $V(s)$, and $f : S \longrightarrow R$ is a fitness function that can be pictured as the *height* of the corresponding solutions. In our study, the search space is composed of binary strings of length N, therefore its size is 2^N. The neighborhood is defined by the minimum possible move on a binary search space, that is the single bit-flip operation. Thus, for a bit string s of length N, the neighborhood size is $|V(s)| = N$.

The $HillClimbing$ algorithm to determine the local optima and therefore define the basins of attraction, is given in Algorithm 1. It defines a mapping from the search space S to the set of locally optimal solutions S^*. Hill climbing algorithms differ in their so-called *pivot-rule*. In best-improvement local search, the entire neighborhood is explored and the best solution is returned, whereas in first-improvement, a neighbor is selected uniformly at random and is accepted if it improves on the current fitness value. We consider here a best-improvement local searcher (see Algorithm 1). For a comparison between first and best-improvement LON models, the reader is referred to [8]

2.1 Nodes

As discussed above, a best-improvement local search algorithm based on the 1-move operation is used to determine the local optima. A local optimum (LO), which is taken to be a maximum here, is a solution s^* such that $\forall s \in V(s)$, $f(s) \leq f(s^*)$.

Let us denote by $h(s)$, the stochastic operator that associates to each solution s, the solution obtained after applying the best-improvement hill-climbing algorithm (see

Algorithm 1 Best-improvement local search (hill-climbing).

Choose initial solution $s \in S$
repeat
  choose $s^{'} \in V(s)$, such that $f(s^{'}) = max_{x \in V(s)} f(x)$
  if $f(s) < f(s^{'})$ **then**
    $s \leftarrow s^{'}$
  end if
until s is a Local optimum

Algorithm 1) until convergence to a LO. The size of the landscape is finite, so we can denote by LO_1, LO_2, $LO_3 \ldots, LO_p$, the local optima. These LOs are the vertices of the *local optima network*.

2.2 Basin-transition edges

The basin of attraction of a local optimum $LO_i \in S$ is the set $b_i = \{s \in S \mid h(s) = LO_i\}$. The size of the basin of attraction of a local optimum i is the cardinality of b_i, denoted $\sharp b_i$. Notice that for non-neutral[4] fitness landscapes, as are standard NK landscapes, the basins of attraction as defined above, produce a partition of the configuration space S. Therefore, $S = \cup_{i \in S^*} b_i$ and $\forall i \in S \; \forall j \neq i$, $b_i \cap b_j = \emptyset$.

We can now define the weight of an edge that connects two feasible solutions in the fitness landscape. For each pair of solutions s and $s^{'}$, $p(s \rightarrow s^{'})$ is the probability to pass from s to $s^{'}$ with the given neighborhood structure. In the case of binary strings of size N, and the neighborhood defined by the single bit-flip operation, there are N neighbors for each solution, therefore, considering a uniform selection of random moves:
if $s^{'} \in V(s)$, $p(s \rightarrow s^{'}) = \frac{1}{N}$ and
if $s^{'} \notin V(s)$, $p(s \rightarrow s^{'}) = 0$.
The probability to go from solution $s \in S$ to a solution belonging to the basin b_j, is[5]:

$$p(s \rightarrow b_j) = \sum_{s^{'} \in b_j} p(s \rightarrow s^{'}) \quad .$$

Thus, the total probability of going from basin b_i to basin b_j, i.e. the weight w_{ij} of edge e_{ij}, is the average over all $s \in b_i$ of the transition probabilities to solutions $s^{'} \in b_j$:

$$p(b_i \rightarrow b_j) = \frac{1}{\sharp b_i} \sum_{s \in b_i} p(s \rightarrow b_j) \quad .$$

2.3 Escape edges

The escape edges are defined according to a distance function d (minimal number of moves between two solutions), and a positive integer $D > 0$.

[4] For a definition of basins that deals with neutrality, the reader is referred to [11].
[5] Notice that $p(s \rightarrow b_j) \leq 1$ and notice also that this definition, disregarding the fitness values, is purely topological and is not related to any particular search heuristic.

There exists an edge e_{ij} between LO_i and LO_j if it exists a solution s such that $d(s, LO_i) \leq D$ and $h(s) = LO_j$. The weight w_{ij} of this edge is then: $w_{ij} = \sharp\{s \in S \mid d(s, LO_i) \leq D$ and $h(s) = LO_j\}$, which can be normalized by the number of solutions within reach w.r.t. such a distance $\sharp\{s \in S \mid d(s, LO_i) \leq D\}$.

2.4 Local optima network

The weighted local optima network $G_w = (N, E)$ is the graph where the nodes $n_i \in N$ are the local optima, and there is an edge $e_{ij} \in E$, with weight $w_{ij} = p(b_i \rightarrow b_j)$, between two nodes n_i and n_j if $p(b_i \rightarrow b_j) > 0$.

According to both definitions of edge weights, $w_{ij} = p(b_i \rightarrow b_j)$ may be different than $w_{ji} = p(b_j \rightarrow b_i)$. Thus, two weights are needed in general, and we have an oriented transition graph.

Figure 1 depicts a representative example of the alternative LON models. The figures corresponds to a real NK landscape with N = 18, K = 2, which is the lowest ruggedness value explored in our study. The left plot illustrates the basin-transition edges, while the center and right plots the escape edges with $D = 1$ and $D = 2$, respectively. Notice that the basin-transition edges (left) produce a densely connected network, while the escape edges produce more sparse networks.

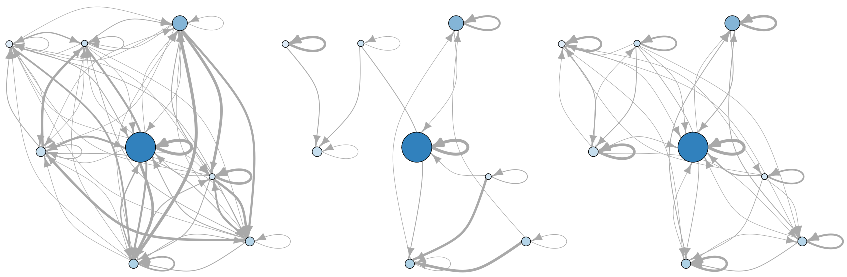

Fig. 1. Local optima network of an NK-landscape instance with N = 18, K = 2. Left: basin-transition edges. Center and Right: escape edges with $D = 1$ and $D = 2$, respectively. The size of the circles is proportional to the logarithm of the size of the corresponding basins of attraction; the darker the color, the better the local optimum fitness. The edges' width scales with the transition probability (weight) between local optima, according to the respective definitions. Notice that the basin-transition edges model (Left) is much more densely connected.

3 Comparative analysis of the LON models

In this section, we compare the LONs resulting from the different edges definitions discussed above. We chose to perform this analysis on the NK-model artificial landscapes, primarily to be able to compare directly with previous work [10, 11], but also because this problem provides a framework that is of general interest in studying the structure of complex combinatorial problems [5].

The NK family of correlated landscapes is in fact a problem-independent model for constructing multimodal landscapes that can gradually be tuned from smooth to rugged. In the model, N refers to the number of (binary) genes in the genotype, i.e. the string length, and K to the epistatic interaction, i.e. the number of genes that influence a particular gene. By increasing the value of K from 0 to $N-1$, the landscapes can be tuned from smooth to rugged. The K variables that form the context of the fitness contribution of a gene, can be chosen according to different models, the two most widely studied being the *random neighborhood* model and the *adjacent neighborhood* model. As no significant differences between the two were found, neither in terms of the landscape global properties [5] nor in terms of their local optima networks (preliminary studies), we conduct our full study on the more general random model.

In order to minimize the influence of the random creation of landscapes, we considered 30 different and independent problem instances for each combination of N and K parameter values. In all cases, the measures reported are the average of these 30 landscapes. In the present study, $N = 18$ and $K \in \{2, 4, 6, 8, 10, 12, 14, 16, 17\}$, which are the largest possible parameter combinations that allow the exhaustive extraction of local optima networks. LONs for the two definitions of edges: (i) basin-transition and (ii) escape edges with $D \in \{1, 2\}$, were extracted and analyzed[6].

Table 1. Local optima network features. Values are averages over 30 random instances, standard deviations are shown as subscripts. K = epistasis value of the corresponding NK-landscape ($N = 18$); N_v = number of vertices; D_{edge} = density of edges ($N_e/(N_v)^2 \times 100\%$); L_{opt} = average shortest path to reach the global optimum ($d_{ij} = 1/w_{ij}$).

K	N_v	D_{edge} (%)			L_{opt}		
	all	Basin-trans.	Esc.$D1$	Esc.$D2$	Basin-trans.	Esc.$D1$	Esc.$D2$
2	$43.0_{27.7}$	$74.182_{13.128}$	$8.298_{4.716}$	$22.750_{9.301}$	$21.2_{8.0}$	$16.8_{4.7}$	$33.5_{14.1}$
4	$220.6_{39.1}$	$54.061_{4.413}$	$1.463_{0.231}$	$7.066_{0.810}$	$41.7_{10.5}$	$19.2_{5.1}$	$53.7_{12.4}$
6	$748.4_{70.2}$	$26.343_{1.963}$	$0.469_{0.047}$	$3.466_{0.279}$	$80.0_{19.1}$	$22.2_{3.9}$	$66.7_{12.9}$
8	$1668.8_{73.5}$	$12.709_{0.512}$	$0.228_{0.009}$	$2.201_{0.066}$	$110.1_{13.8}$	$24.0_{4.9}$	$76.6_{9.1}$
10	$3147.6_{109.9}$	$6.269_{0.244}$	$0.132_{0.004}$	$1.531_{0.036}$	$152.8_{19.3}$	$27.3_{5.0}$	$90.7_{8.4}$
12	$5270.3_{103.9}$	$3.240_{0.079}$	$0.088_{0.001}$	$1.115_{0.015}$	$185.1_{23.8}$	$30.3_{6.7}$	$108.3_{12.3}$
14	$8099.6_{121.1}$	$1.774_{0.035}$	$0.064_{0.001}$	$0.838_{0.009}$	$200.2_{16.0}$	$38.9_{9.6}$	$124.7_{8.6}$
16	$11688.1_{101.3}$	$1.030_{0.013}$	$0.051_{0.000}$	$0.647_{0.004}$	$211.8_{15.0}$	$47.9_{11.4}$	$146.2_{11.2}$
17	$13801.0_{74.1}$	$0.801_{0.007}$	$0.047_{0.000}$	$0.574_{0.002}$	$214.3_{17.5}$	$55.7_{12.5}$	$155.9_{12.2}$

3.1 Network features and connectivity

Number of nodes and edges. The 2^{nd} column of Table 1, reports the number of nodes (local optima), N_v, which is the same for all the studied landscapes and models. The

[6] Some of the the tools for fitness landscape analysis and the local search heuristics, were used from the "ParadisEO" library [2]; data treatment and network analysis are done in "R" with the "igraph" package [3].

number of nodes increase exponentially with increasing values of K. The networks, however, have a different number of edges, as can be appreciated in the 3^{rd}, 4^{th}, and 5^{th} columns of Table 1, which report the number of edges normalized by the square of the number of nodes (density of edges). Clearly, the density is higher for the basin-transition edges, followed by the escape edges with $D = 2$, and the smaller density corresponds to $D = 1$. The trend is, however, that density decreases steadily with increasing values of K, which supports the correlation between the two models.

With the basin-transition edges, LONs are densely connected, especially when K is low: 74% and 54% of all possible edges are present, on average, for $K \in \{2, 4\}$. The escape edges produce sparsely connected graphs. Indeed, the $D = 1$-escape edge model, produces networks that are not completely connected, with the number of connected components ranging between 1.67 and 8.37 in average. The global optimum, though, always happens to belong to the largest connected component, which, in our analysis, comprises an average proportion of solutions raising, with increasing values of K, from 0.9392764 to 0.9999879.

The networks with escape edges and $D = 2$, are always connected. The density decreases with the epistasis degree. For high Ks, the density values are close to those of the basin-transition networks. Figure 2 (Left) illustrates what is happening in terms of the average degree of the outgoing links. First, notice that the difference with the basin-transition networks is maximal when K is between 4 and 12. Whereas for $D = 1$ the outgoing degree only increases from 1.7 to 5.5 across the range of K values, for $D = 2$ the growth is faster and reaches 78.2, not far from the 109.5 score of the LON with basin-transition edges. The size of the basins could provide an explanation for this: at high values of K basins are so small that a 2-bit mutation from the local optimum is almost enough to recover the complete topology.

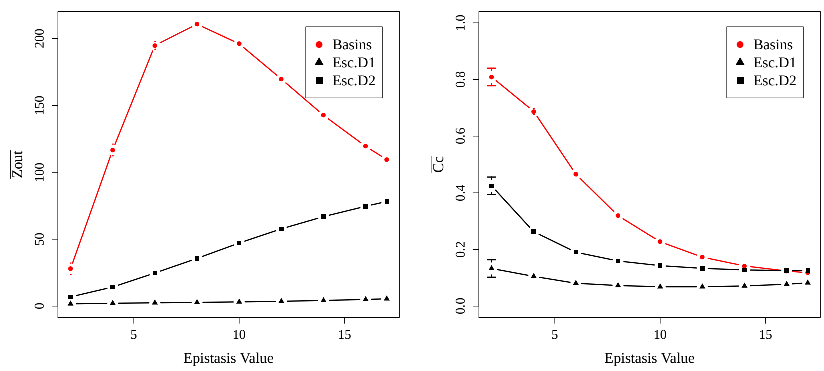

Fig. 2. Average out-degree (Left) and average clustering coefficient (Right) vs epistasis value.

Clustering coefficient. LONs with basins-transitions edges present a somewhat symmetric structure: when two nodes i and j are connected, both edges e_{ij} and e_{ji} are present (even though their weights are in general different, $w_{ij} \neq w_{ji}$). Moreover, those connections often form triangular closures whose frequency is given by the global clustering coefficient. As Figure 2 (Right) shows, this measure of transitivity is lower with the escape-transition edges, but the difference could be due to the different number of edges of those LONs. Values for $D = 1$ are remarkably low, even if the calculation disregarded the direction of edges. Overall though, the decreasing trend w.r.t. the landscape ruggedness, remains common. In other words, even with escape-transition edges, the clustering coefficient can be retained as a measure related to problem complexity: it decreases with the non-linearity of the (NK-) problem.

Shortest paths. Due to the differences in topology, in the escape edges networks not all the paths are possible: few nodes might be disconnected, or they might not be reachable due to direction constraints (these are more "asymmetric" networks, as can be seen in Fig. 1). Thus, while evaluating shortest paths, only paths connecting reachable couples of nodes are averaged. Moreover, there are different ranges of weights, so the values displayed in Figure 3 have been normalized. An unexpected behavior can be observed for $D = 2$: the average path length peaks at $K = 4$ and stays always high. Maybe the increasing connectivity of nodes (see Fig. 2) counteracts their increase in numbers. However, some paths are more important than others, for example those who lead to the global optimum (see Fig. 3 (right)). With respect to these paths, all the LON models show the same trend: the paths increase in length as ruggedness increases.

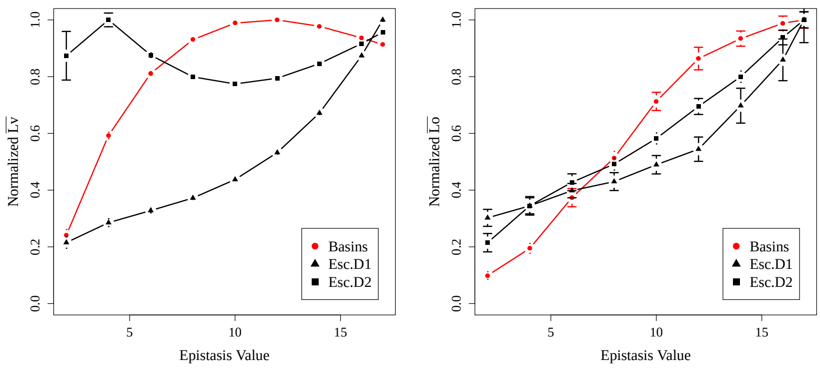

Fig. 3. Shortest paths over the LON vs epistasis value. Left: average geodesic distance between optima. Right: average shortest path to the global optimum. Each curve has been divided by its respective maximum value.

3.2 Characterization of weights

Disparity. Figure 2 (Left) gave the average connectivity of vertices, counting outgoing links. One might then ask whether or not there are preferential directions when leaving a particular basin, i.e. if for a given node i, the weights w_{ij} (with $j \neq i$) are equivalent. For this purpose, we measure the disparity Y_2 [1], which gauges the heterogeneity of the contributions of the edges of node i to the total weight. For a large enough degree

z_i, when there is not a dominant weight, then $Y_2 \approx 1/z_i$. The connectivity of LONs with escape edges with $D = 1$ is weak, and it is difficult to draw conclusions based on disparity only, as Figure 4 (left) illustrates. In the example illustrated, where $K = 4$, (i.e. relatively low epistasis, and so not a random structure), Y_2 approaches $1/z$ for escape-$D = 1$, whereas it has distinctively higher values for both the escape-$D = 2$, and the basin-transition edges.

However, the common trend is that disparity decreases with increasing epistasis: as the landscapes become more rugged, the transition probabilities to leave a particular basins appear to become more uniform, which could relate to the search difficulty. This is clear from Fig 4 (right), where Y_2 approaches $1/z$ on average as K grows.

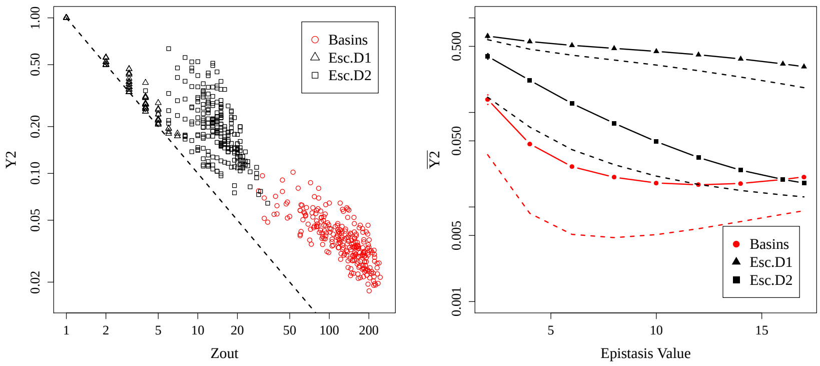

Fig. 4. Disparity of edges' weights. Left: scatter plot of disparity against outgoing degree (lin-log scale) for an instance with epistasis $K = 4$. Right: average vertex disparity for outgoing edges vs epistasis value. Dotted lines represent the inverse of the outgoing degree.

Strength. In a general weighted network, the degree of a vertex naturally extends into its strength, which measures its weighted connectivity. In LON model, basins size, as well as connectivity, generally correlate with fitness value [8]. Thus we ask if the incoming strength of a given node, i.e. the sum of the transition probabilities for all the incoming connections, correlates with the fitness of its LO. Figure 5 gives a clear affirmative answer for all the definition of edges.

Correlation of weights among edges' definitions. The LON models resulting from the alternative definition of edges show different structures but common trends w.r.t. the features that are related to problem difficulty. We conclude this subsection by directly comparing the transition probabilities that result from the alternative edge definitions. For this purpose, Figure 6 shows the Spearman's rank correlation between the corresponding rows of the weighted adjacency matrix, for different LON models of the same instance. The statistic is always positive, but, given the sparser nature of the escape-$D = 1$ networks, the result is weaker in this case. With $D = 2$, though, the correlation with the basin-transition definition is consistently good, which agrees with the previous

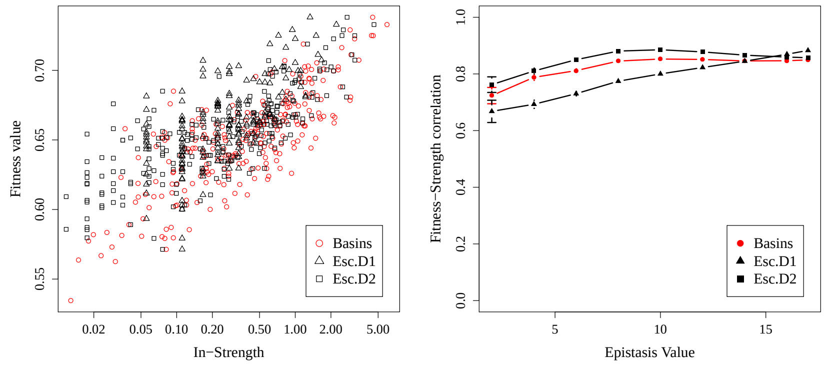

Fig. 5. Correlation between the strength of a node and its fitness value. Left: scatter plot of fitness against weighted connectivity (lin-log scale) for an instance with epistasis $K = 4$. Right: average correlation Spearman's coefficient between in-strength and fitness value vs epistasis value.

findings on this topology. Indeed, as the landscape ruggedness increases, the correlation between $D = 1$ and $D = 2$ becomes smaller.

4 Model validation: local search dynamics

In this section, we analyze the connection between the LON model and dynamics of local search (LS) in an attempt to validate its descriptive power. Does a LS follow the edges of the LON? Which edge definition is the most accurate to predict the dynamics of a local search heuristic? This is a preliminary study on one particular NK-landscape instance with $N = 18$ and $K = 4$, a larger analysis will be the subject of future work.

4.1 Experiment setup

We choose two simple but efficient stochastic LS heuristics, namely Iterated Local Search (ILS) [6], and Tabu Search (TS) [4]. Both are given $2.6 * 10^4$ function evaluations ($\approx 10\%$ of the search space), or stop when reaching the global optimum.

The search trajectory is traced and filtered according to the basins of attraction: each solution belongs to one basin, so it is labelled by the corresponding local optimum. We then considered that a LS stays in the same basin, if the solutions s_t and s_{t+1} both belong to the same basin and the fitness increases: $f(s_t) \leq f(s_{t+1})$. Otherwise, the LS jumps from one basin to another one. In that manner, when accumulated from a number of independent runs of the LS, 10^4 in our experiments, it is possible to compare the empirical transition frequencies between basins with the corresponding edge weights of a given LON. Such a high number of independent runs is necessary because often times single trajectories go through few LOs and only once.

When a LS performs a transition between two basins that are not connected in the LON model, we add to the latter a virtual edge with weight equal to 0.0. Of course, we

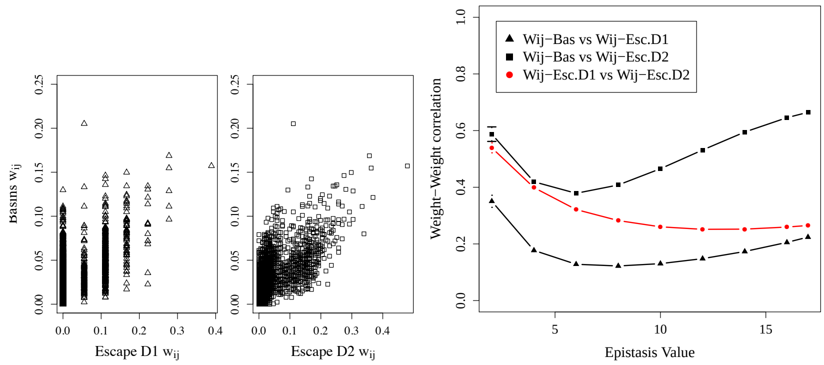

Fig. 6. Correlation between the weights resulting from the the new and the old definition of edges. Left: scatter plot for an instance with epistasis $K = 4$. Right: average Spearman's rank correlation coefficient vs epistasis value. No self-loops.

do not consider the LON edges that are not sampled, as it is not possible to compute the transition frequency between nodes that have not been visited. Finally, the LS under study are based on hill-climbing, thus the LON self-loops are also discarded.

4.2 Results

Figure 7 shows the scatter plots for the correlations between the edges weights of the different LON models and the empirical transition frequencies between basins.

Iterated Local Search. Our implementation is based on a steepest-ascent (best-improvement) hill-climbing, which, at each iteration, moves to the best neighboring solution, and stops on a local optimum. The ILS perturbation is the k bit-flip mutation, which flips k bits at random, $k \in \{1, 2\}$ in the present study. A mutation is accepted as a new solution if its fitness value is strictly better than the current one. Intuitively, the ILS follows the same edges of the escape definition: from the LO, k bits are flipped; the the main difference lies in the acceptance criterion used by the ILS. Indeed (see Fig. 7), correlations are all significant. As expected, for the 1-bit-flip perturbation, Spearman coefficient is higher for escape edges with $D = 1$, and lower for $D = 2$-escape edges (0.49 for $D = 1$, 0.45 for $D = 2$, and 0.48 for basin edges). For the 2-bit-flips perturbation, the highest correlation is for the escape edges with $D = 2$, and the smallest for escape edges with $D = 1$ (0.41 for $D = 1$, 0.75 for $D = 2$, and 0.72 for basin edges). Notice how figures for basin edges and $D = 2$-escape edges are very similar.

Tabu Search. Our implementation uses the 1-bit-flip mutation, and, in order to obtain a better diversification, a set of moves are forbidden during the search; those tabu-moves are stored in a memory that lasts $\lambda = N = 18$ iterations. A tabu move is nonetheless accepted when it finds a new best-so-far solution. Referring to Figure 7, the correlation values are 0.27 for the $D = 1$-escape edges, 0.53 for $D = 2$, and 0.79 for the basin edges. All values are significant and positive, particularly so for the more interesting LON definitions: $D = 2$-escape edges and basin edges.

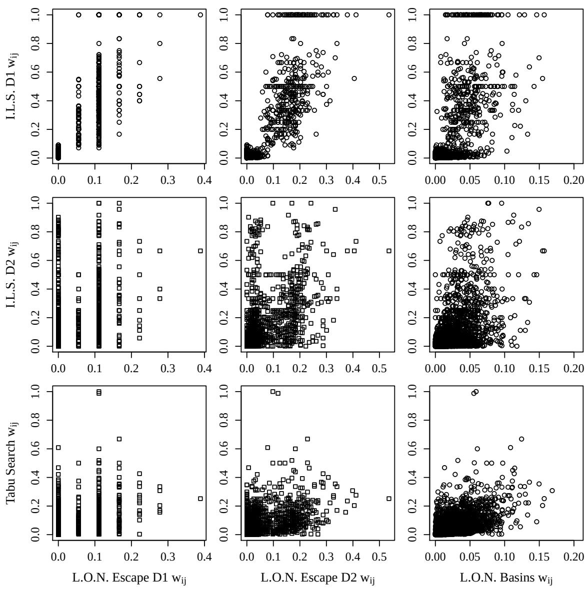

Fig. 7. Correlation (Spearman coefficient) between the edge weight and the empirical transition frequency of a local search. 10^4 independent runs on an NK-instance with $N = 18$, $K = 4$. From top to bottom, Iterated Local Search with 1- and 2-bit-flips, and Tabu Search have been tested. From left to right, the empirical frequencies are plotted against the corresponding edge weights according to the Escape $D \in \{1, 2\}$ and Basins definition, respectively. Only transitions between different basins are considered (no self-loops).

The result of this preliminary study is encouraging because it shows that the LON could capture the coarse-grained dynamics of a LS. In particular, the escape edge definition with $D = 2$ seems to be informative enough to correlate with the trajectories of a simple ILS, or a simple TS. Of course, further studies have to confirm this result for a larger class of LS, a broader number of instances, and on other problems, as to delineate the validity domain of the LON model.

5 Concluding remarks

The local optima networks (LON) model is a mesoscopic representation of a problem search space, which deals with the local-optima basins of attractions as the meso-state level of description. In this contribution, an alternative definition of edges (*escape edges*), has been proposed. Our statistical analysis, on a set of NK-landscapes, shows that the escape edges are as informative as the original (basin- transition) edges. We reach this conclusion because, for both LON models, the analyzed network features (such as the clustering coefficient, disparity, correlation between in-strength and fitness of local optima, and path length to global optimum) are always consistent with the non-linearity of the problem (the K parameter), which tunes the landscape ruggedness. Indeed, the edges' weights are positively correlated between the different definitions.

We also present a preliminary analysis that aims at validating the model. We show that the dynamics of simple stochastic local search heuristics such as Iterated Local Search, or Tabu Search, tend to follow the edges of LONs according to the rate defined by the weights. The model validation in the present study is a first step. Starting from the work by Reidys and Stadler [9] on combinatorial fitness landscapes, we could conduct a spectral analysis of the LON model. From the adjacency matrix of the LON graphs, it should be possible to build a Markov Chain having the previously discussed transition probabilities. That would allow us to compare the stationary distribution of different LON models of the same search space. In this case, an empirical assessment could be performed in a more informed way, because we could estimate the number of local search runs that are necessary to have a good sampling, as in a Monte Carlo method. The model validation could, then, go together with a prediction of the LS dynamics.

Overall, the present study opens up relevant perspectives. The escape edges definition will allow us to design a sampling methodology of the LONs. The enumeration of the basins of attraction is impossible on realistic search spaces. Therefore, the original definition of edges is restricted to study small search spaces. With the new definition, and through sampling of large networks such as breadth-first search, forest-fire or snowball sampling, we will be able to study the properties of LONs for real-world combinatorial search spaces. Our hope is that, by combining the LON model of the search dynamics and the ability to build LONs for large search spaces, we will be able to perform off-line parameter tuning of evolutionary algorithms and local search heuristics. Moreover, if the LON features are collected along the search process, on-line control of the parameters could also be achieved.

References

1. Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A.: Characterization and modeling of weighted networks. Physica A 346, 34–43 (2005)
2. Cahon, S., Melab, N., Talbi, E.G.: Paradiseo: A framework for the reusable design of parallel and distributed metaheuristics. Journal of Heuristics 10, 357–380 (2004)
3. Csardi, G., Nepusz, T.: The igraph software package for complex network research. InterJournal Complex Systems, 1695 (2006)
4. Glover, F., Laguna, M.: Tabu Search. Kluwer Academic Publishers, Norwell, MA, USA (1997)
5. Kauffman, S.A.: The Origins of Order. Oxford University Press, New York (1993)
6. Lourenço, H.R., Martin, O., Stützle, T.: Iterated local search. In: Handbook of Metaheuristics, International Series in Operations Research & Management Science, vol. 57, pp. 321–353. Kluwer Academic Publishers (2002)
7. Newman, M.E.J.: The structure and function of complex networks. SIAM Review 45, 167–256 (2003)
8. Ochoa, G., Verel, S., Tomassini, M.: First-improvement vs. best-improvement local optima networks of nk landscapes. In: Parallel Problem Solving from Nature - PPSN XI. Lecture Notes in Computer Science, vol. 6238, pp. 104–113. Springer (2010)
9. Reidys, C., Stadler, P.: Combinatorial landscapes. SIAM review 44(1), 3–54 (2002)
10. Tomassini, M., Verel, S., Ochoa, G.: Complex-network analysis of combinatorial spaces: The NK landscape case. Phys. Rev. E 78(6), 066114 (2008)
11. Verel, S., Ochoa, G., Tomassini, M.: Local optima networks of NK landscapes with neutrality. IEEE Transactions on Evolutionary Computation (to appear)