

Mapping the Global Structure of TSP Fitness Landscapes

Gabriela Ochoa · Nadarajen Veerapen

Received: date / Accepted: date

Abstract The global structure of combinatorial landscapes is not fully understood, yet it is known to impact the performance of heuristic search methods. We use a so-called local optima network model to characterise and visualise the global structure of travelling salesperson fitness landscapes of different classes, including random and structured real-world instances of realistic size. Our study brings rigour to the characterisation of so-called funnels, and proposes an intensive and effective sampling procedure for extracting the networks. We propose enhanced visualisation techniques, including 3D plots and the incorporation of colour, sizes and widths, to reflect relevant aspects of the search process. This brings an almost tangible new perspective to the landscape and funnel metaphors. Our results reveal a much richer global structure than the suggestion of a ‘big-valley’ structure. Most landscapes of the tested instances have multiple valleys or funnels; and the number, disposition and interaction of these funnels seem to relate to search difficulty on the studied landscapes. We also find that the structured TSP instances feature high levels of neutrality, an observation not previously reported in the literature. We then propose ways of analysing and visualising these neutral landscapes.

Keywords Fitness landscapes · Local optima networks · Funnels · Global structure · Neutrality · Travelling salesman problem · TSP · Lin-Kernighan heuristic · Iterated local search · Visualisation

1 Introduction

The global structure of realistic combinatorial fitness landscapes is still little understood, yet it clearly impacts the performance of heuristic search methods. A way of characterising a landscape global structure under a given neighbourhood operator is

G. Ochoa and N. Veerapen
Computing Science and Mathematics
University of Stirling, Scotland, UK
E-mail: {goc,nve}@cs.stir.ac.uk

by considering the distribution of its local optima. [Boese et al \(1994\)](#) conjectured that the search space of travelling salesman instances under 2-exchange moves has a ‘globally convex’ or ‘big-valley’ structure, in which local optima are clustered around one central global optimum. This globally convex structure has subsequently been observed in other combinatorial fitness landscapes such as the NK model ([Kauffman and Levin 1987](#)), graph bipartitioning ([Merz and Freisleben 1998](#)), and flowshop scheduling ([Reeves 1999](#)). Under this view, there are many local optima but they are easy to escape from, with the coarse level gradient leading to the global optimum. This hypothesis has become generally accepted and has inspired the design of local search heuristics referring to a similar principle with different names: adaptive multi-starts, large-step Markov chain, soft-restarts, chained local search and iterated local search. The notion of a big-valley is related to the notion of a ‘funnel’ structure from the study of energy landscapes in theoretical chemistry, as discussed further in [Section 2.2](#).

However, recent studies on TSP landscapes have revealed a more complex picture ([Hains et al 2011](#); [Ochoa and Veerapen 2016b](#)). The big-valley seems to decompose into several sub-valleys or multiple funnels. This helps to explain why certain iterated local search heuristics can quickly find high-quality solutions, but fail to consistently find the global optimum in cases where the global optimum is known. A similar multi-funnel structure has been observed on some continuous optimisation problems ([Locatelli 2005](#); [Lunacek and Whitley 2006](#); [Lunacek et al 2008](#)), where its impact on search difficulty has been established. In particular, landscapes with more than one funnel, where the global optimum is located in a deep, narrow funnel are significantly harder. The literature on characterising the multi-funnel structure of combinatorial landscapes is mostly lacking. This is partly due to the lack of adequate tools to study their complexity. We propose using local optima networks to analyse and visualise the global structure of combinatorial fitness landscapes. Local optima networks compress the whole search space into a graph, where nodes are local optima and edges are transitions among them with a given search operator ([Tomassini et al 2008](#); [Verel et al 2011](#)). The model emphasises the number, distribution and most importantly, the connectivity pattern of local optima. Modelling landscapes as networks introduces a new set of metrics to analyse fitness landscapes and, interestingly, the possibility of visualising them.

This article complements and extends our recent work on local optima network analysis of TSP fitness landscapes ([Ochoa and Veerapen 2016a,b](#)). Our previous work considered only a handful of instances of size up to 700 cities, and deliberately excluded degenerate instances (i.e. instances featuring neutrality). The sampling process was less intensive, and the characterisation of funnels incomplete. The main contributions of this article are:

1. An intensive sampling procedure, which allows the extraction of local optima networks for larger TSP instances (of up to 1500 or so cities) of different types, including both random and structured real-world instances (city instances and drilling problems).

2. An empirical characterisation and modelling of funnel floors, which requires both the identification of *sink* nodes, and the incorporation of higher-level *plateau-nodes* to model instances with neutrality.
3. A rigorous empirical characterisation of funnel basins, using the notion of *monotonic sequences* from theoretical chemistry (Berry and Kunz 1995).
4. Enhanced visualisation of the multi-funnel structure, including 3D plots, alternative graph layouts, and the incorporation of colour, nodes sizes and edge widths, to reflect relevant aspects of the search process.
5. A correlation study identifying connections between heuristic search performance and the global structure of the studied TSP instances.

The remainder of this article is organised as follows. The next section gives relevant background on TSP solvers and the notion of ‘funnel’. Section 3 defines the local optima network model considered. Section 4 describes the proposed empirical methodology, including the procedures and conceptual tools for extracting the network data, detecting the funnel floors and calculating the funnel basins. It also describes the TSP instances studied. Section 5 overviews our results including an analysis and visualisation of local optima networks, a more traditional fitness-distance analysis, a study of the impact of the sampling parameters and a correlation study between heuristic search performance and global landscape metrics. Finally, Section 6 summarises our main findings and suggests directions of future work.

2 Background and Related Work

2.1 TSP solvers

This section describes the TSP solvers used in our study both for comparison purposes and as part of the sampling procedure implemented.

2.1.1 Concorde

Concorde is currently the best-performing exact TSP solver (Applegate et al 2007, 2003a). It has been used to solve the largest non-trivial TSP instances (of up to 85900 cities) for which provably optimal solutions are known. Concorde is based on a complex Branch & Cut algorithm that uses a multitude of heuristic mechanisms to achieve good performance on a wide range of TSP instances. For example, it carries out a limited number of iterations of the Chained Lin-Kernighan heuristic (described below) during the initial stages of its computation to determine an initial upper bound to the objective value. Additionally, an exact mixed integer program solver is used to compute and refine the lower bound by solving a relaxed linear program of the problem.

2.1.2 The Chained Lin-Kernighan heuristic

The Lin-Kernighan (LK) algorithm (Lin and Kernighan 1973) is a powerful and well-known heuristic for finding approximate solutions to the TSP. For about two decades,

it was the best local search method, and nowadays it is a key component of many state-of-the-art TSP solvers. LK-search is based on the idea of k -exchanges: take the current tour and remove k different links from it, which are then reconnected in a new way to achieve a legal tour. Figure 1a illustrates a 2-exchange move. A tour is considered to be ' k -opt' if no k -exchange exists which decreases its length. LK-search applies 2, 3 and higher-order k -exchanges. The order of a change is not predetermined, rather k is increased until a stopping criterion is met. Thus many kinds of k -exchanges and all 3-exchanges are included. There are many ways to choose the stopping condition and the best implementations are rather involved. We use the implementation available in the Concorde software package (Applegate et al 2003a), which uses *do not look* bits and candidate lists.

The overall tour-finding strategy using LK-search was, previously, to repeatedly start the basic LK routine from different starting points keeping the best solution found. This practice ended in the 1990s with the seminal work of Martin et al (1992), who proposed the alternative of *kicking* (perturbing slightly) the LK tour and reapplying the algorithm. If a better tour is produced, the old LK tour is discarded and the new one kept. Otherwise, the search continues with the old tour and kicks it again. This simple yet powerful strategy is nowadays known as *iterated local search*. It was named *Chained Lin-Kernighan* (Chained-LK) by Applegate et al (2003b), who also provided an improved implementation to solve large TSP instances. The kick or escape operator in Chained-LK is a type of 4-exchange (depicted in Fig. 1b), named *double-bridge* by Martin et al (1992). It consists of two 2-exchanges, each of which is a 'bridge' as it takes a legal, connected tour into two disconnected parts. The combination of both bridges must then be chosen in order to produce a legal final tour.

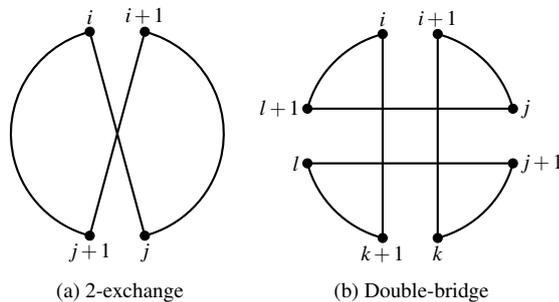


Fig. 1: Illustration of tours obtained after 2-exchange and double-bridge moves.

2.2 The notion of funnel

The intuition behind the concept of a 'funnel' is captured by Figure 2 where two funnels are depicted as two groups of local optima which are close in configuration space within a group, but well-separated between groups.

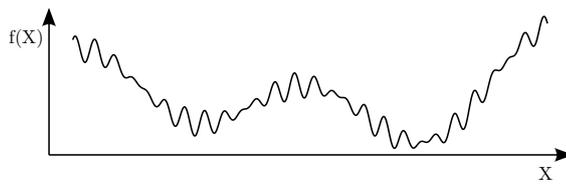


Fig. 2: Depiction of two funnels.

The term ‘funnel’ was introduced in the protein folding community to describe “a region of configuration space that can be described in terms of a set of downhill pathways that converge on a single low-energy structure or a set of closely-related low-energy structures” (Doye et al 1999). It has been suggested that the energy landscape of proteins is characterised by a single deep funnel, a feature that underpins their ability to fold to their native state. In contrast, some shorter polymer chains (polypeptides) that misfold are expected to have other funnels that can act as traps. Approaches to elucidate the global landscape structure have led to the concept of *disconnectivity graphs* (Becker and Karplus 1997; Doye et al 1999; Wales 2005), also known as *barrier trees* (Flamm et al 2002). In the context of energy landscapes, Berry and Kunz (1995) first introduced the term *monotonic sequence* to describe a sequence of local minima where the energy of minima is always decreasing. The set of monotonic sequences that lead to a particular minimum was termed ‘basin’; in this sense a ‘basin’ is analogous to a protein folding ‘funnel’. The collection of local optima associated to a funnel has also been termed ‘super-basin’ in the literature introducing disconnectivity graphs (Becker and Karplus 1997).

Energy landscapes in theoretical chemistry and fitness landscapes in optimisation are conceptually related, as has been already observed by Stadler (2002). This relationship is particularly close for continuous optimisation. Locatelli (2005) studied the sources of difficulty in continuous optimisation and finds that it is not strictly related to the number of local optima, but to how *chaotic* their positions are. Lunacek and Whitley (2006) propose a metric, *dispersion*, that quantifies the proximity of the best regions in the search space. A high dispersion metric indicates the presence of multiple funnels. In a follow up work, Lunacek et al (2008) studied abstract landscapes with two funnels and find that evolutionary algorithms mostly fail when the global optimum is located in a proportionally smaller funnel. Recent work on exploratory landscape analysis of continuous search spaces reveals that multimodality and global structure are among the most important high-level properties that help differentiate between problem classes (Bischi et al 2012; Kerschke et al 2015).

The literature is much more scarce for discrete search spaces. The notion of a big-valley discussed in the introduction is clearly related to the notion of a single funnel structure, and fitness distance scatter plots and correlation metrics are a standard tool in landscape analysis. Other approaches to understanding the global structure of combinatorial landscapes have been applied to small and simplified problems. Barrier trees have been applied to discrete optimisation problems where the notions of local optima, basins and saddle points are clearly defined, for instance in the context of

spin-glasses (Hordijk et al 2003). Flamm et al (2002) have extended these definitions so that barrier trees can be constructed for highly degenerate problems (i.e landscapes with neutrality). They present empirical results for binary strings of up to length 10. Hallam and Prugel-Bennett (2005) construct barrier trees for MAX-SAT problems with up to 40 variables using branch-and-bound to find only the best local optima in the space.

Daolio et al (2011) studied the community structure of local optima networks on two classes of instances of the quadratic assignment problem. The two problem classes give rise to different configuration spaces, with the so-called real-like instances revealing a modular structure. The approach is based on a full enumeration of local optima. Therefore, instances of size up to 10 were analysed. In a follow up work with a data-driven approach, the modularity of instances up to size 32 was studied (Iclanzan et al 2014). This work, however, did not relate the community structure to the notion of funnels. Herrmann et al (2016) recently established a connection between groupings (communities) in local optima networks and the notion of funnels. They extracted the networks of NK landscapes of length 20 and various levels of epistasis, and applied a community detection algorithm. Results confirm that landscapes consist of several clusters and the number of clusters increases with the epistasis level. A higher number of clusters, and a larger size of the cluster containing the global optimum were found to lead to a higher search difficulty.

3 The Local Optima Network Model

This section describes the local optima network model used in our study. We start by defining the notion of fitness landscapes, and follow by formalising the notions of nodes and edges of the network model.

A fitness landscape (Stadler 2002) is a triplet (S, N, f) where S is a set of potential solutions i.e. a search space; $N : S \rightarrow 2^S$, a neighbourhood structure, is a function that assigns to every $s \in S$ a set of neighbours $N(s)$, and $f : S \rightarrow R$ is an objective function (also called fitness function) that can be pictured as the *height* of the corresponding solutions.

The search space of a TSP instance of size m is the set of permutations of the m cities. The objective function f is given by the length of the tour, which is to be minimised. In order to model TSP fitness landscapes, we adapted the local optima network model with escape edges (Verel et al 2012). To construct these networks, we need to define their nodes and edges. The definitions are related to the search operators being modelled, specifically, the local search heuristic and *escape* operators. Our study considers those within the Chained Lin-Kernighan algorithm, namely, the Lin-Kernighan local search heuristic and the double-bridge escape move (described in Section 2.1.2).

Local optima. A local optimum, which in the TSP is a minimum, is a solution s^* such that $\forall s \in N(s^*), f(s^*) \leq f(s)$. Notice that the inequality is not strict, in order to allow the treatment of the neutral landscape case.

The neighbourhood N is imposed by LK-search, which considers variable values of k . The local optimality criterion is, therefore, rather stringent. Only a small proportion of all possible solutions are LK-optimal. The set of local optima, which corresponds to the set of nodes in the network model, is denoted by L and its cardinality by n .

Escape edges. Edges are directed and based on the double-bridge operator. There is an *escape* edge from local optimum x to local optimum y , if y can be obtained after applying a double-bridge kick to x followed by LK-Search. Edges are weighted with estimated transition probabilities between the connected nodes. These probabilities are estimated by the sampling process. Specifically, edge weights are integers indicating the number of times an edge was visited during the sampling process (described in Section 4.1). The set of escape edges is denoted by E .

Local Optima Network (LON). A local optima network is a graph $LON = (L, E)$ where nodes are the local optima L , and edges E are the escape edges. Edges are directed and weighted. Weights indicate transition probabilities.

4 Empirical Methodology

Our approach extracts and analyses local optima networks of TSP instances of realistic sizes and different types. The objective is to map and characterise the landscapes' global structure. Clearly, a full enumeration of the local optima for TSP instances of non-trivial size becomes unmanageable. Therefore, networks are constructed from a sample of high-quality local optima in the search space. This section starts by describing our sampling methodology and procedure for constructing the LONs. Thereafter, we describe the approach for characterising the funnel structures, which requires both detecting the funnel floors, and computing their basins.

4.1 Sampling the network data

To extract the network data, we considered the Chained-LK implementation of Concorde (see Algorithm 1). Instead of discarding the search history, we store, in L , every new improving LK local optima found during the search process. We also create and store, in E , a directed edge between the starting and ending optima after a double-bridge followed by LK. Edges only go from a node with higher cost to a node of equal or lower cost, in order to store the *monotonic sequences* and thus identify funnels (see Section 4.2). We keep count of the number of times an edge is visited, this number is stored as the edge's weight.

A thousand independent runs of Chained-LK are executed for each TSP instance. We consider two initialisation mechanisms, one producing better initial solutions than the other, in order to have a broader picture of the search space. Half of the runs use the Quick-Borůvka method, the default initialisation for Concorde's Chained-LK,

```

Data:  $I$ , TSP instance
Result:  $L$ , set of local optima,
          $E$ , set of escape edges
 $L \leftarrow \{\}; E \leftarrow \{\}$ 
for  $i \leftarrow 1$  to 1000 do
     $s_{start} \leftarrow \text{initialSolution}()$ 
     $s_{start} \leftarrow \text{LK}(s_{start})$ 
     $L \leftarrow L \cup \{s_{start}\}$ 
    while  $j < 10000$  do
         $s_{end} \leftarrow \text{applyKick}(s_{start})$ 
         $s_{end} \leftarrow \text{LK}(s_{end})$ 
         $j \leftarrow j + 1$ 
        if  $\text{Objective}(s_{end}) \leq \text{Objective}(s_{start})$  then
             $L \leftarrow L \cup \{s_{end}\}$ 
             $E \leftarrow E \cup \{(s_{start}, s_{end})\}$ 
             $s_{start} \leftarrow s_{end}$ 
             $j \leftarrow 0$ 
        end
    end
end

```

Algorithm 1: Local optima network sampling combining 1000 runs of Chained-LK.

which is based on the minimum-weight spanning tree algorithm of Borůvka (Applegate et al 2003b). The other half starts from a random solution. The default kicking procedure in Concorde’s Chained-LK is used: the edges involved in the double-bridge are selected using random walks along connected vertices. Each Chained-LK run continues until at least 10000 consecutive iterations are performed without finding an improving solution. The network is thereafter created by the combination of the unique nodes and edges produced by this sampling process.

4.2 Identifying funnel structures

The challenge is to devise an approach for identifying funnel structures, for different types of TSP instances, once the local optima networks have been extracted. Our approach adapts the notion of *monotonic sequences* from theoretical chemistry (Berry and Kunz 1995). We consider a monotonic sequence as a sequence of local optima where the evaluation of solutions is non-deteriorating. The collection of monotonic sequences leading to the same lowest minimum correspond to the so-called ‘monotonic sequence basins’ (Wales 2005). These structures have also been called ‘super-basins’ in the theoretical chemistry literature (Becker and Karplus 1997). We choose here to call them ‘funnel basins’ or simply ‘funnels’ borrowing from the protein folding literature. We can distinguish the *primary* funnel, as the one involving monotonic sequences that terminate at the global optimum. The primary funnel is separated from other neighbouring secondary funnels by transition states laying on a so-called ‘primary divide’ (Berry and Kunz 1995). Above such a divide, it is possible for a local optima to belong to more than one funnel through different monotonic sequences.

Our approach requires us to empirically locate the lowest cost minima which potentially lie at the bottom funnels, and thereafter compute each funnel basin. These procedures are described below.

4.2.1 Identifying the funnel floors

Hains et al (2011) considered funnel floors (or funnel bottoms) as those solutions empirically found after considerable search effort. Specifically, for each TSP instance, Chained-LK was run until at least 10 000 iterations without finding an improving tour, and this procedure was repeated 1000 times from different starting solutions. However, runs were considered separately and the intermediate local optima and transitions among them discarded, therefore, the number of funnel floors was overestimated.

We propose a refinement of this approach to more accurately estimate the number of funnels. We keep a similar computational effort, but store all the different local optima visited across the 1000 runs, and combine them into a single local optima network. This allows us to merge local optima transitions found across different runs. We name the solutions found at the end of each run (i.e. 10 000 consecutive iterations without an improvement) as *attractors* rather than as *funnel bottoms*. Our local optima network data shows two interesting features. First, solutions attracting the search process are not always single solutions, but are often part of connected local optima plateaus. Second, attractor nodes are not always at the bottom of funnels: a single run is trapped in an attractor, but when combining the 1000 runs into a single LON, many attractors show escape paths to solutions with lower evaluation. Therefore, in order to refine the process of detecting the conjectured funnel floors, we propose constructing a sub-network from the LON, where nodes are the attractors: the *Attractor Network (AN)*. Once this network is constructed for a given instance, we detect its sinks, i.e. nodes without outgoing edges. We consider these sinks as the solutions at the bottom of funnels, which produced a reduced number of funnels for the studied instances as compared to previous work (Hains et al 2011; Ochoa and Veerapen 2016b). We overview below the notation and procedures used to detect the number of funnel floors.

- *Attractor* nodes are empirically determined as those local optima at which Chained-LK stalls after a large search effort (10 000 consecutive iterations without finding an improving solution in our implementation). The set of attractor nodes is denoted by A , and its cardinality by a .
- A *sink* node, is an attractor node without outgoing edges to other attractor nodes of lower evaluation. Sink nodes are conjectured to be at the bottom of a funnel structure. The set of sink nodes is denoted by S and its cardinality by s .
- A *plateau* node, is a higher-level node compressing a group of local optima with the same objective function value belonging to a connected component according to the escape edges defined in Section 3. Plateau nodes can also be characterised as attractors or sinks.
- *Attractor-plateau* nodes are calculated by compressing connected attractor nodes at the same objective function level. The set of attractor-plateau nodes is denoted by A_p , and its cardinality by a_p .

- A *sink-plateau* node, is an attractor-plateau node without outgoing edges to other attractor-plateau nodes. The set of sink-plateau nodes is denoted by S_p and its cardinality by s_p .

Our previous local optima network models of TSP landscapes (Ochoa and Veerapen 2016b) involved a less extensive sampling and did not consider plateau nodes. However, the study of a more varied set of instances revealed neutrality (i.e. different solutions with the same objective function value) on the landscapes of most structured instances, especially those of drilling problems. The random instances do not show neutrality, but many of the city instances and a large proportion of the drilling problems do reveal plateaus that can be extensive. It was, therefore, necessary to include plateau nodes, and to define two auxiliary network models (described below) in order to algorithmically detect the sink nodes, and thus funnel floors.

- *Attractor Network (AN)*, the sub-graph $AN = (A, E_a)$ of LON where nodes are the attractors A , and edges $E_a \subseteq E$ are the escape edges connecting them.
- *Attractor-plateau Network (AN_p)*, the graph $AN_p = (A_p, E_p)$ formed by contracting the nodes and edges of the Attractors network AN .

Figure 3 illustrates the attractor and attractor-plateau networks for the well-studied TSPLIB instance att532 (more details in Table 1), which consists of the 532 largest cities of the USA and considers pseudo-Euclidean distances between them. The procedure by Hains et al (2011), estimated four funnels for this instance. Our analysis, instead, suggests that it has only two funnels, whose sinks are represented in red and yellow in Figure 3. As the figure indicates, there are 48 attractor nodes (a), which are compressed into 7 plateau-attractor nodes (b). Not all the 7 plateau-attractor nodes are funnel floors, as most of them have outgoing edges to other nodes. Instead only two of them, the sinks coloured in yellow and red in plot (b), are conjectured to be at the bottom of funnels. The whole sampled local optima network for this instance, including nodes within 1% in evaluation from the global optimum cost, is visualised in Fig. 5 (a).

4.2.2 Identifying the funnel basins

Once the funnel sinks are detected, we can proceed to identify the funnel basins (see Algorithm 2). This is done by finding all the local optima in the network which are reachable from each funnel sink (sink-plateau for the instances with neutrality). Breadth-First-Search is used for this purpose. The set of unique solutions in the combined paths to a given funnel sink corresponds to the funnel basin. The cardinality of this set corresponds to the funnel size. Notice that the membership of a solution to a funnel might be overlapping, that is, a solution may belong to more than one funnel, in that there are paths from that solution to more than one funnel sink. The relative size of the primary funnel (or any other secondary funnel) is calculated as its size divided by the total number of local optima.

As described above, funnel sinks are nodes without outgoing edges in the attractor network, which is a sub-graph of the sampled local optima network (LON). When considering the whole LON, out-going edges might exist between sub-optimal sinks

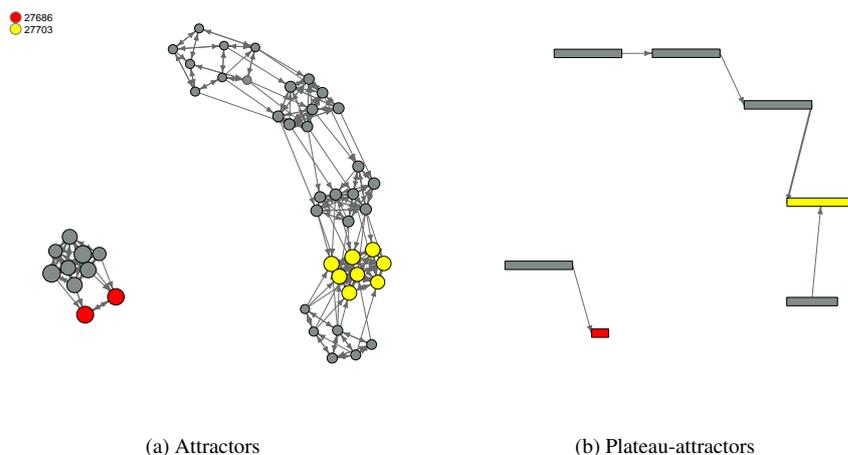


Fig. 3: Visualisation of the attractor network (a), and the same network (b) after compressing plateau nodes for a selected instance with 532 cities (**att532**). In the attractor network, each node is a local optimum with size proportional to its in-strength (i.e. the weighted in-degree). In the attractor-plateau network, nodes are rectangular with width proportional to the number of nodes in the plateau. The instance contains two funnel sink-plateaus, visualised in red (the global optima) and yellow (the secondary funnel sink), the legend indicates their cost values. Grey nodes are above (i.e higher cost) than their respective sinks.

and local optima with lower evaluation, which are not sinks themselves, but belong to a different basin. In consequence, before identifying the funnel basins for each sub-optimal sink, we remove all its outgoing edges. This is done by the function $disconnectSinks(LON, S)$ in Algorithm 2.

4.3 Instances

Table 1 summarises the TSP instances studied. We consider 20 instances of moderate size (in the range of 500 to 1500 or so cities) and different types. The first 10 instances are randomly generated using the DIMACS generator¹. Half of these are composed of uniformly distributed cities (prefixed by ‘E’), while in the other half, the cities are clustered (indicated by a ‘C’). The suffix number ‘.x’ in the instance name indicates, as per DIMACS convention, a seed of $x + 10\,000$. These synthetic instances are part of a larger set of 200 instances that we generated, with the number of cities being uniformly selected in the range [400, 1600]. These are used in Section 5.6 for a correlation study between landscape metric and heuristic search performance.

¹ dimacs.rutgers.edu/Challenges/TSP/download.html

Data: LON : sampled local optima network, S : funnel sinks
Result: $bsizes$: basin sizes vector, $basins$: funnel basins vector,
 $boverlap$: set of local optima in more than one basin

```

i ← 0
 $DLON$  ← disconnectSinks( $LON$ ,  $S$ )
for  $s \in S$  do
  |  $basins[i]$  ← breadthFirstSearch( $DLON$ ,  $s$ )
  |  $bsizes[i]$  ← length( $basins[i]$ )
  |  $i$  ←  $i + 1$ 
end
 $boverlap$  ← {}
for  $i, j \in range(1, length(S))$  do
  |  $boverlap$  ←  $boverlap \cup intersection(basins[i], basins[j])$ 
end

```

Algorithm 2: Identifying funnel basins. Function $disconnectSinks(LON, S)$ removes all outgoing edges from sub-optimal sinks in LON .

The bottom 10 instances in Table 1 are well-known instances from TSPLIB (Reinelt 1991). A popular way of constructing TSP instances is to choose a set of actual cities and define the cost of travel between any two cities as the distance between them. The first 5 TSPLIB instances are constructed in such a way. The last 5 arise from the task of drilling holes in printed circuit boards. The types of edge weights are as follows. EUC-2D refers to the Euclidean distance of points in a 2D plane rounded to the nearest integer. ATT refers to a pseudo-Euclidean distance where the sum of the squares is divided by 10 and the square root of this value is then rounded to an integer. GEO refers to the integer geographical distance computed from latitude and longitude coordinates on the surface of a sphere representing an idealised Earth.

The third column in Table 1 reports the success rate of the 1000 Chained-LK runs used for extracting the network data. By success rate, we mean the ratio of runs that found at least one global optimum. The last two columns give information on the solving difficulty of each instance. Specifically, we report the mean run time and the mean number of branch-and-bound nodes required by Concorde (interfaced with IBM ILOG CPLEX 12.6 as its mixed integer solver) to solve the instances to optimality on a 3.4 GHz Intel Core i7-3770 CPU across 10 runs. Although Concorde is an exact solver, the means are computed since it uses Chained-LK to generate initial solutions. Therefore, times include the Chained-LK time to compute initial solutions. This leads to different execution times and branch-and-bound trees. A number of branch-and-bound (B&B) nodes equal to 1 indicates that the lower and upper bounds found in the initial stages of the Concorde solving process match, and thus no actual tree was explored.

5 Results

The results and analysis are reported in the following six subsections. We start by reporting general statistics of the sampled local optima networks across the instance set. Subsection 5.2 concentrates on features characterising the landscapes' global struc-

Table 1: TSP instances with number of cities as suffix, edge type, Chained-LK success rate, and features resulting from running the Concorde solver: running time and number of branch-and-bound (B&B) nodes.

Instance	Edge Type	CLK Success	Concorde solver		
			Optimum	Run time (s)	B&B nodes
<i>Random uniform instances</i>					
E506.25	EUC-2D	0.957	16 313 719	12.0	6.2
E755.73	EUC-2D	0.128	20 158 565	28.0	8.0
E1010.37	EUC-2D	0.478	22 904 325	18.3	3.8
E1243.85	EUC-2D	0.030	25 227 141	118.4	38.0
E1521.33	EUC-2D	0.030	28 027 563	239.8	74.6
<i>Random clustered instances</i>					
C506.25	EUC-2D	0.329	6 816 950	8.0	5.6
C755.73	EUC-2D	1.000	9 867 050	3.2	1.0
C1010.37	EUC-2D	0.112	10 716 003	41.7	14.2
C1243.85	EUC-2D	0.136	12 943 477	110.3	33.6
C1521.33	EUC-2D	0.178	13 608 402	34.5	6.2
<i>Somewhat structured instances (city problems)</i>					
att532	ATT	0.437	27 686	10.3	6.2
gr666	GEO	0.183	294 358	7.3	3.4
pr1002	EUC-2D	0.673	259 045	3.3	1.0
rl1304	EUC-2D	0.286	252 948	19.4	1.0
nrw1379	EUC-2D	0.008	56 638	34.5	12.6
<i>Highly structured instances (drilling problems)</i>					
u574	EUC-2D	0.442	36 905	4.1	1.4
rat783	EUC-2D	0.959	8806	4.0	1.2
u1060	EUC-2D	0.214	224 094	35.2	15.4
d1291	EUC-2D	0.258	50 801	1098.5	37.4
fl1577	EUC-2D	0.012	22 249	292.3	9.8

ture. Subsection 5.3 visualises the local optima network of some selected instances. Thereafter, Subsection 5.4 provides a more classic fitness landscapes analysis. We finish with a study of the effects of sampling parameters on the landscape metrics (Subsection 5.5), and a correlation study between landscape metrics and heuristic search performance (Subsection 5.6).

5.1 General network metrics

Table 2 reports basic statistics of the sampled local optima networks, including the number of unique global optima go , the number of different local optima n , the number of unique objective function evaluations $evals$, and the relationship between these two values $n/evals$ as an indication of the amount of neutrality in the landscape. The table also reports the number of connected components c , the proportion of nodes in the network from where there is a path to a global optimum p_{go} ; and the average \bar{l}_{go} and maximum l_{go} path length from any node in the network to a global optimum. The last two columns report the average \bar{d} and maximum d^{max} in-strength (i.e. incoming

weighted degree, where the weight of an edge is the number of times it was traversed during sampling) of nodes in the networks.

Results show that, as expected, the number of global optima and the average and maximum path length to a global optimum generally increase with the number of cities for each instance class. There are, however, striking differences among the classes. The randomly generated instances always show a single global optimum and each local optima has a different objective function value (as indicated by columns *evals* and *n/evals*). This is not the case for the structured instances, where several global optima are generally the norm. Indeed global optima seem to be located in large plateaus (of several thousands of nodes) for some of the drilling instances. The high amount of neutrality is also revealed by the smaller number of objective function levels as compared to the number of optima, with the ratio being as large as several hundreds or thousands. This is probably because several pairwise distances between cities are the same, and some of these distances are rather small. Therefore, several city orderings will have the same evaluation. The structured instances also show longer path lengths to a global optimum, which can be in part explained by high amount of neutrality.

The average in-strength, \bar{d} , is usually quite low since most nodes are visited only once. Higher values indicate either that multiple runs find the same nodes or that there is cycling between nodes of equivalent evaluation. This cycling is also why d^{max} is quite high for some instances.

A surprising consequence of sampling and modelling the local optima networks for the very neutral instances such as u1060 and fl1577, is that the number of connected components (c in Table 2) equals the number of runs in the sampling process (1000). This means that each run traverses a different set of solutions, in other words, there are no shared local optima among the runs, even though many runs reach the same objective function values. We suspect that these instances contain such large plateaus, that runs (starting from different initial points) explore completely different parts of them. For example, the number of different global optima found by our sampling process on instance u1060 is 163 569. The metric *n/evals* in Table 2, gives an estimate of the average size of plateaus for each instance, which can be as large as several hundred.

5.2 Funnel metrics

Table 3 reports metrics describing the global structure of the random instances. We report the proportions of: solutions in the primary funnel (i.e the funnel containing the global optimum) f_{go} , solutions in the largest funnel (which can be the primary funnel) f_{lg} , and solutions belonging to more than one funnel f_{ov} . The table also reports the number of attractors a and funnel sinks s . The proportion of the in-strength (i.e. the weighted in-degree) of global optimum sinks to the in-strength of all sinks, d_{go} , is given and the Chained-LK success rate from Table 1 is reproduced in the last column for comparison purposes.

Results suggest that the number of funnels generally increases, and the size of the primary funnel decreases, with increases to the number of cities on both instance

Table 2: Local optima network metrics. Number of unique global optima go , number of unique local optima n , number of unique objective function evaluations $evals$, relationship between number of nodes and unique evaluations $n/evals$, number of connected components c , proportion of nodes from where there is a path to a global optimum p_{go} , average path length to a global optimum \bar{l}_{go} , maximum path length to a global optimum l_{go} , average in-strength of nodes \bar{d} and maximum in-strength d^{max} .

	go	n	$evals$	$n/evals$	c	p_{go}	\bar{l}_{go}	l_{go}	\bar{d}	d^{max}
E506.25	1	14 730	14 304	1.03	1	1.00	10.08	30	1.37	957
E755.73	1	24 774	23 569	1.05	1	0.42	16.87	47	1.14	255
E1010.37	1	39 685	36 923	1.07	1	1.00	23.69	74	1.12	478
E1243.85	1	50 779	46 366	1.10	46	0.12	31.46	79	1.04	105
E1521.33	1	64 146	58 309	1.10	190	0.11	37.57	103	1.00	30
C506.25	1	19 099	16 842	1.13	1	1.00	12.99	36	1.24	622
C755.73	1	32 040	28 937	1.11	1	1.00	18.68	64	1.12	1000
C1010.37	1	48 319	40 724	1.19	1	0.85	30.00	91	1.07	250
C1243.85	1	59 894	52 929	1.13	3	0.57	34.75	105	1.04	349
C1521.33	1	81 957	70 488	1.16	1	0.48	44.79	142	1.03	257
att532	2	23 851	827	28.84	1	1.00	14.95	48	9.01	17 803
gr666	2	33 892	7598	4.46	1	0.78	21.04	62	1.86	5227
pr1002	1	54 821	7542	7.27	13	0.98	30.82	95	1.08	673
rl1304	1	50 670	11 328	4.47	2	0.79	29.19	92	1.10	286
nrw1379	96	226 763	1330	170.50	292	0.07	71.26	238	2.28	180
u574	4	28 115	1230	22.86	1	0.79	16.86	51	9.77	30 388
rat783	1024	112 911	264	427.69	1	1.00	37.08	281	7.76	1748
u1060	163 569	1 396 071	5579	250.24	1000	0.16	586.04	3431	1.02	4
d1291	512	943 430	3398	277.64	215	0.11	111.53	2143	1.49	837
fl1577	22 628	4 266 040	3445	1238.33	1000	0.01	1878.45	9695	1.01	4

classes. The random uniform instances (prefixed with ‘E’), show a larger number of funnels and a smaller primary funnel, when compared to the clustered instances (prefixed with ‘C’). These metrics change faster with the instance size on the uniform instances, specially the number of funnels. The proportion of solutions in more than one funnel is also larger for the uniform instances, which is expected given the larger number of funnels. By virtue of its definition, the relative in-strength of globally optimal sinks is closely related to success rate, but it is not exactly the same, as can be observed from the table.

Table 4 reports global metrics for the structured instances. Since these instances contain neutrality, the attractor-plateaus networks (AN_p) are constructed in order to identify the plateau-sink nodes, and thus the number of funnels. Results suggest that the number of funnels is less correlated to the instance size as it is the case for the random instances. Since some of the instances here exhibit several primary funnels, note that f_{go} represents the relative size of largest of those funnels. Other instance features seem to influence the landscape global structure. The number of funnels on the structured instances is somewhere in between that of the uniform and the

Table 3: Random instances global structure metrics. Relative sizes of the primary funnel f_{go} , the largest funnel f_{lg} , proportion of solutions in more than one funnel f_{ov} , number of attractor local optima a , number of funnel sinks s , and relative in-strength of global optimum sinks d_{go} . Chained-LK success rates are reproduced in the last column for comparison purposes.

	f_{go}	f_{lg}	f_{ov}	a	s	d_{go}	CLK
E506.25	0.99	0.99	0.50	4	2	0.75	0.96
E755.73	0.37	0.59	0.44	17	10	0.21	0.13
E1010.37	0.92	0.92	0.86	33	12	0.50	0.49
E1243.85	0.05	0.22	0.46	214	148	0.03	0.03
E1521.33	0.05	0.05	0.17	372	327	0.03	0.03
C506.25	1.00	1.00	0.00	5	1	1.00	0.33
C755.73	1.00	1.00	0.00	1	1	1.00	1.00
C1010.37	0.81	0.81	0.72	39	7	0.32	0.11
C1243.85	0.43	0.43	0.37	29	9	0.16	0.14
C1521.33	0.48	0.51	0.07	28	8	0.47	0.18

Table 4: Structured instances global structure metrics. Relative sizes of the primary funnel f_{go} , the largest funnel f_{lg} , proportion of solutions in more than one funnel f_{ov} , number of attractor local optima a , number of plateau attractors a_p , number of sink-plateau nodes s_p , and relative in-strength of global optimum sinks d_{go} . Chained-LK success rates are included. For the very neutral instances (nrw1379, u1060, d1291, fl1577) a range instead of a single value is reported.

	f_{go}	f_{lg}	f_{ov}	a	a_p	s_p	d_{go}	CLK
att532	0.89	0.89	0.37	52	8	2	0.55	0.44
gr666	0.74	0.74	0.66	55	29	13	0.26	0.18
pr1002	0.64	0.64	0.46	60	53	50	0.51	0.67
rl1304	0.74	0.74	0.59	38	33	16	0.34	0.29
nrw1379	<0.01–0.02	0.02–0.04	0.11–0.95	54 325	86– 473	61– 442	0.01	0.01
u574	0.79	0.79	0.29	26	6	2	0.53	0.44
rat783	0.82	0.82	0.10	10 104	5	4	0.86	0.96
u1060	<0.01	<0.01	0 –0.94	757 533	90–1000	90–1000	0.21	0.21
d1291	0.03–0.08	0.03–0.08	0.50–0.86	476 107	116– 294	98– 275	0.22	0.26
fl1577	<0.01	<0.01	0 –0.94	1 838 082	55–1000	55–1000	0.01	0.01

clustered random instances (see Table 3). The uniform random instances show the largest number of funnels in the studied set.

For the instances with high levels of neutrality, i.e. nrw1379 and the drilling problems except u574 and rat783, it was not possible to assess whether the attractor local optima at a given objective function level fully connected into a single plateau or groups of plateaus sharing the same objective function value. The plateaus are potentially very large, with the sampling process not guaranteeing their full exploration. Alternatively there may be multiple plateaus for a single objective function level. We therefore used two different methods to estimate bounds for the different metrics. The first method approximates the plateaus by assuming that they are indeed con-

nected. With this approach, instances such as d1291 and fl1577, reveal a relatively small number of funnels despite their large number of different local optima. The second method does not assume that there is a single plateau for each objective function level, instead it considers all connected nodes sharing the same evaluation as different plateaus. This leads to 1000 such plateaus, the same as the number of runs, for instances u1060 and fl1577. Further discussion of this issue, through the analysis of distances between solutions, is presented in Section 5.4. Note that for these two instances, f_{go} and f_{lg} have very small ranges and are thus summarised by a single value.

A thorough study of these highly neutral instances may require additional sampling efforts and additional analysis to characterise the extent of the plateaus. We leave that for future work. Nevertheless, the global structure in terms of the number of funnels does not seem to differ significantly from that of the city instances or clustered random instances, if we assume that plateaus are indeed connected. We argue that search difficulty on very neutral landscapes relates not only to the multi-funnel structure, but also to the size and location of the plateaus. A large plateau at the global optimum level may reflect an easy to search instance, while a large sub-optimal plateau may act as a trap that is difficult to escape from.

5.3 Network visualisation

One of the advantages of modelling a system as a network, is the possibility of visualising it. This is one of the strengths of the proposed approach, as it allows a more accessible way of grasping the complexities of landscapes global structure.

Software for analysing and visualising networks is currently available in various languages and environments. Here we use the R statistical language together with the *igraph* package (Csardi and Nepusz 2006). Layout algorithms are at the core of network visualisation, they assign vertices to positions in a metric space. *Force-directed* methods model the pairwise attraction and repulsion of vertices, and are known to reflect the *community structure* or modularity of a network (Noack 2009). We, therefore, use them in order to visually characterise the landscapes' multi-funnel structure. Funnels can be visually identified as modules in the network. As our model indicates, nodes are LK-search local optima and edges represent escape transitions according to double-bridge moves. We decorated them to reflect features relevant to search dynamic. The size of nodes is proportional to their incoming strength (weighted incoming degree), therefore, it reflects the extent to which nodes attract the search dynamics. The colour of nodes reflects their funnel membership. We used the *heat colours* palette, a sequential colour scheme skewed to the reds and yellows. Red identifies the global optima, and the yellow colour gradient reflects an increase in cost. The edges' widths are proportional to their weight, which indicates the frequency of transitions. That is, the most frequently visited edges are thicker. We present both 2D and 3D images. In Ochoa and Veerapen (2016a), we proposed a 3D visualisation where the x and y coordinates are, as usual, determined by a graph layout algorithm; the innovation is to use the objective function as the z coordinate. This provides a clearer representation of the funnel sink and basin concepts, bringing an almost tangible as-

pect to the landscape and funnel metaphors. The global optimum can be identified in the 3D plots as the node with the lowest z coordinate.

In order to have manageable images, we plotted the networks corresponding to the subset of local optima within 0.1% or 0.05% in evaluation from the global optimum. We also removed self-loops for improved visibility. Figures 4 and 5 illustrate local optima networks for selected random and structured instances, respectively.

The plots in Figures 4a and 4b illustrate how the number of funnels rapidly increases with the number of cities for the uniform random instances. Instance E1243.85 features 76 and 19 funnels for solutions within 0.1 and 0.05% in evaluation from the global optimum, while instance E755.73, have 4 and 2, respectively. The size of the funnel containing the global optimum is much smaller for E1243.85 as compared with the smaller uniform instance E755.73 and the clustered instance of the same size C1243.85. Indeed, instances E755.73 (Fig. 4a) and C1243.85 (Fig. 4b) have a similar Chained-LK success rate despite their difference in size, which indicates the importance of global structure in search difficulty. Both instances feature a large secondary funnel coloured in orange, which acts as a strong attractor of the search process as indicated by the size of its funnel sink.

Figure 5 illustrates the local optima networks for solutions within 0.1% in objective function value from the optimal solution for 2 city instances att532 and pr1002, and one drilling problem u574. Instances att532 (Fig. 5a) and u574 (Fig. 5c) have a similar CLK success rate and also a similar global structure for solutions within 0.1% in evaluation from the global optimum cost. They both feature two funnels visualised in red (the primary funnel) and yellow (the secondary funnel). In att532 the two funnels are overlapping, whereas in u574 they are separated (for solutions within 0.1% in evaluation, but they may overlap at higher solution costs). An important difference between the random (Fig. 4) and the structured instances (Fig. 5) is that the latter show neutrality. This is reflected by the number of global optima (a single one for all the studied random instances), and 2 and 4 for att532 and u574, respectively. The neutrality can also be appreciated at higher solution costs on the 3D plots for att532 and u574, where several solutions are located at the same objective function level.

The global structure of city instance pr1002 (Fig. 5b) is strikingly different. It clearly has a large primary funnel sink visualised in red, which reflects the high CLK success rate (0.67) of this instance. There are however 3 secondary funnels whose sink solutions have an evaluation 80 units higher than the optimum (visualised in orange and yellow). It is important to note that these 3 nodes do not form a plateau despite having the same objective function values, they are not connected with double-bridge moves according to our sample. These 3 funnels have several connections to other solutions in the primary funnel, although they do not connect directly to the primary funnel sink. Therefore, they are separate funnels according to our empirical definition. The visual depiction however, seems to reflect that they belong to the primary funnel, which suggests that there might be hierarchies of funnel structures.

For some of the studied instances, each sampling run visits a different set of solutions. This leads to as many connected components as there are runs, as is the case for instances u1060 and fl1577. Furthermore, each component consists of a long chain of nodes with some small loops when plateaus are explored. The previously used network visualisation provides relatively little information for such cases. Figure 6

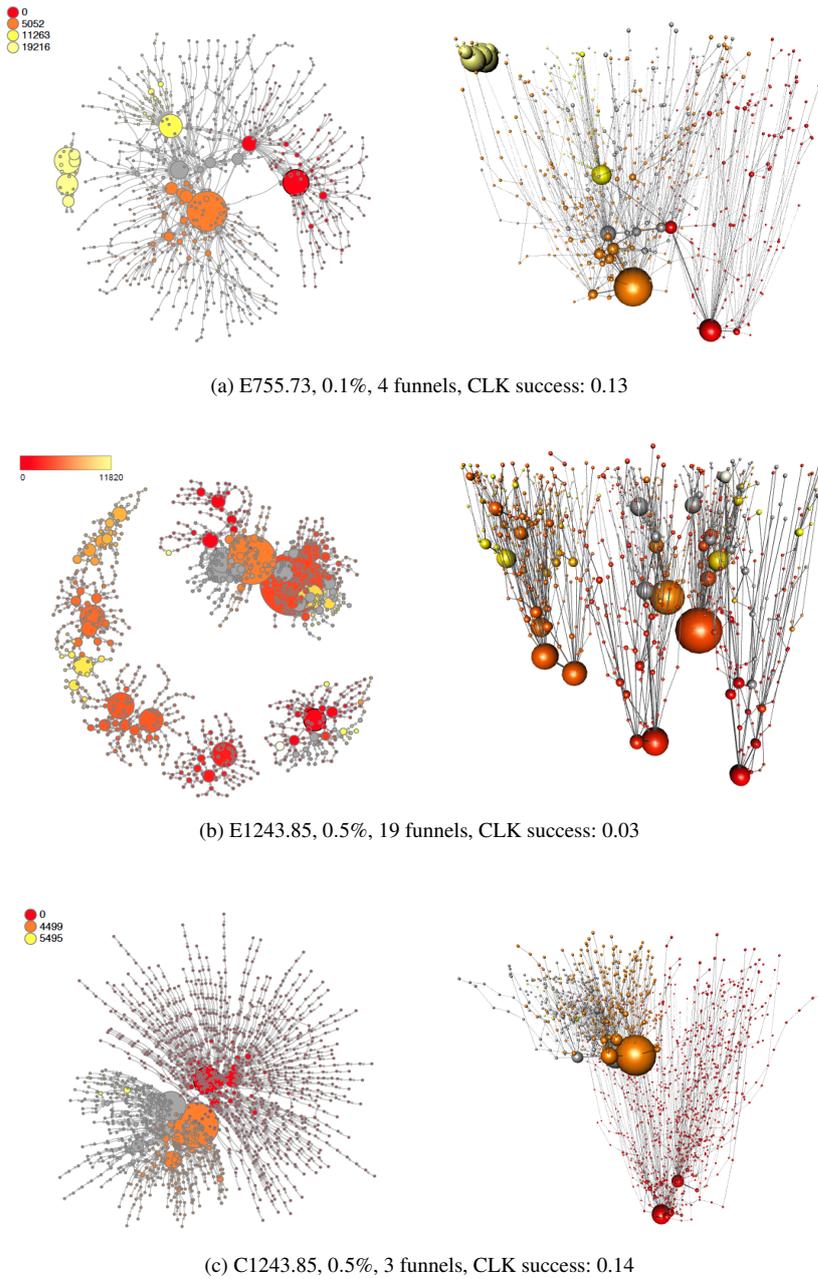
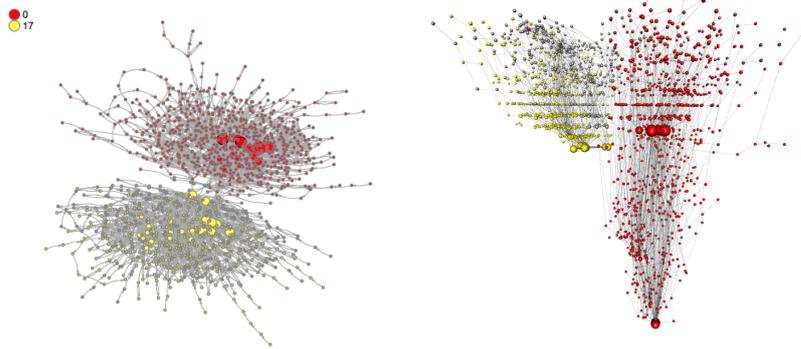
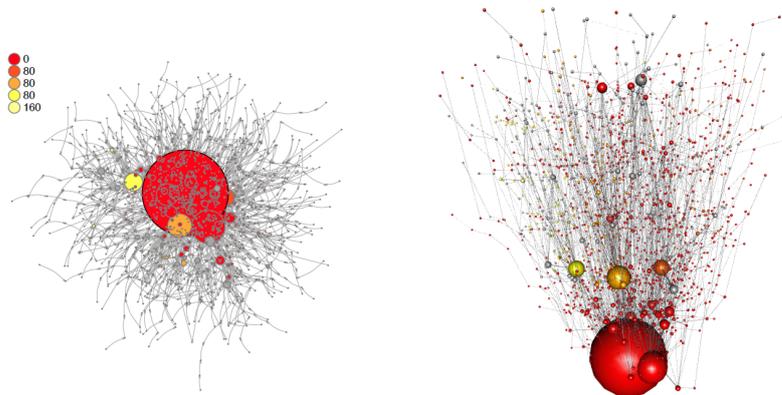


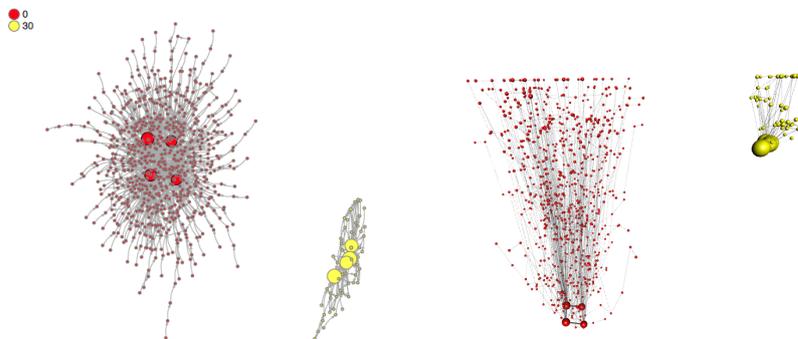
Fig. 4: Local optima networks for selected random instances. Images are shown in 2D and a 3D projection. Funnel structures are visualised in colours, with red indicating the funnel containing the global optimum, and the yellow gradient an increase in cost. Grey nodes are those belonging to more than one funnel. The legend indicates the sinks cost difference from the global optimum. Node sizes are proportional to their incoming strength.



(a) att532, 0.1%, 2 funnels, CLK success: 0.44



(b) pr1002, 0.1%, 5 funnels, CLK success: 0.67



(c) u574, 0.1%, 2 funnels, CLK success: 0.44

Fig. 5: Local optima networks for selected structured instances. Images are shown in 2D and a 3D projection. Funnel structures are visualised in colours, with red indicating the funnel containing the global optimum, and the yellow gradient an increase in cost. Grey nodes are those belonging to more than one funnel. The legend indicates the sinks cost difference from the global optimum. Node sizes are proportional to their incoming strength.

presents an alternative visualisation for instances where different runs do not share any common solutions. Blue dots represent single solutions. Red lines represent consecutive solutions in a plateau, i.e. the solutions have the same cost, and the length of the line is proportional to the number of solutions in the plateau. Each run is displayed as one column of dots and lines since there is no overlap between the solutions found in different runs.

Figure 6 specifically shows the top 25 largest components, i.e. runs, containing the 25 largest plateaus. A marked difference can be observed in the distribution of solutions and plateaus between the two instances. For u1060, there is usually a single plateau at the end of the run that is very close to the global optimum. The latter is not actually reached for the top 25 runs with largest plateaus. For fl1577, runs usually encounter a number of different plateaus, manage to escape from them, and finally get stuck in a non-optimal plateau. The globally optimal value is reached in only 3 out of 25 runs. The fact that u1060 has fewer plateaus to escape than fl1577 provides an explanation for its higher success rate in finding a global optimum out of 1000 runs. The visualisation provides a straight-forward view of the distribution of plateaus.

A question that remains is whether all the plateaus with equal evaluation found across the different runs are actually a single very large plateau. We attempt to answer this by analysing the distance between solutions in Section 5.4.

5.4 Distance analysis

In addition to visualising and analysing the networks, it is also useful to examine the landscapes using the pairwise distances between nodes. We consider the bond distance in particular, which is defined as the difference in the number of common edges, or bonds, between two solutions. It corresponds to the number of edges within a solution minus the number of edges that are in common between the two solutions considered.

Bond distance heatmaps can be useful to quickly assess how close pairs of solutions are to each other. Figure 7 displays the pairwise bond distance between sink nodes or nodes within sink-plateaus for the six instances visualised in Figures 4 and 5. The objective function of each node is displayed on the vertical axis. The plot is mirrored along the diagonal. For E1243.85 and pr1002, only the ten best local optima are considered and, in both cases, these nodes are sink nodes and are not part of a plateau. For the other four instances, there are ten or fewer nodes to consider. The colour gradient allows us to clearly distinguish the two plateaus within instances att532 and u574. The bond distance within these plateaus is lower than 15 units and generally only in single digits, showing that the solutions are very similar, as expected.

When considering highly neutral instances, like u1060 and fl1577, where each run produced a new connected component, an important question is whether the different plateaus found at the same objective function value are actually part of some very large plateau or if they are an artefact of the stochasticity of the sampling method.

Figure 8 attempts to provide some insight by visualising the distribution of pairwise bond distances between sink-plateau nodes sharing the same evaluation. In practice, because each sink-plateau contains several hundred nodes, we choose the node

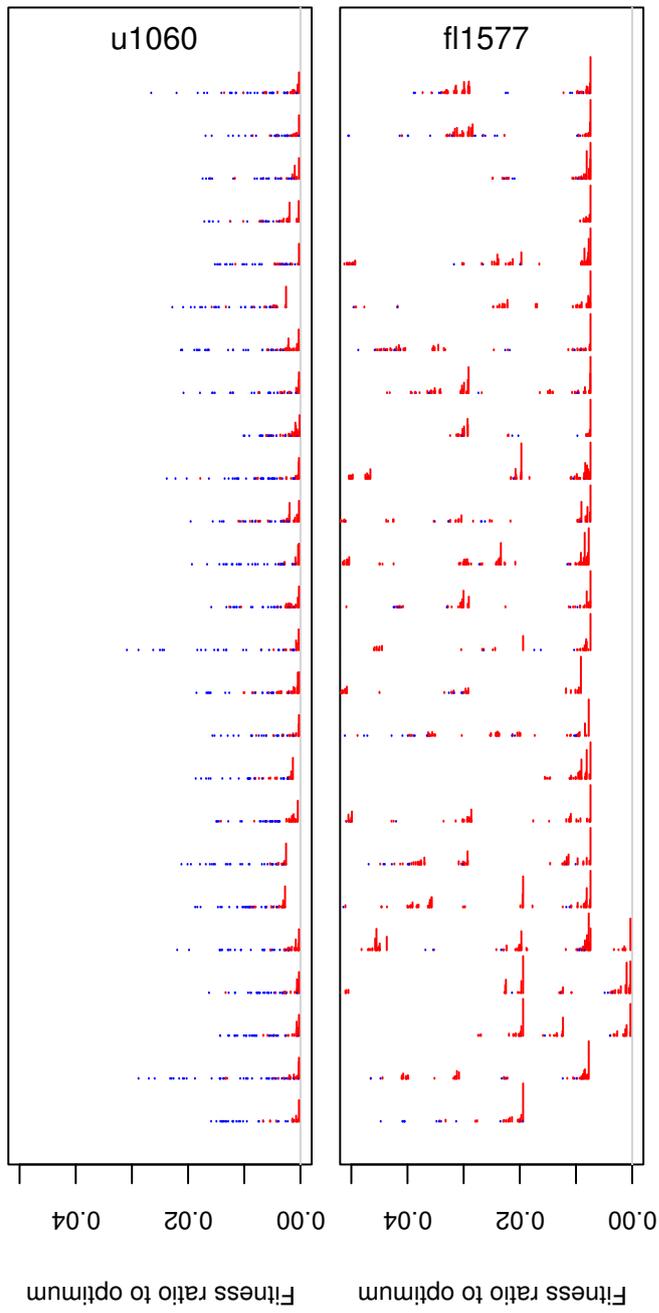


Fig. 6: Local optima networks for u1060 and fl1577. For both, the 25 components containing the 25 largest plateaus are plotted. Each component, which corresponds to a run, is displayed as a column of blue dots for single solutions and red lines for plateaus. The length of the red lines is proportional to the square root of the number of unique solutions found in a given plateau in a given component. The y-axis corresponds to the ratio of objective function values to the that of the global optima. The gray line indicates ratio 0, the global optima objective value.

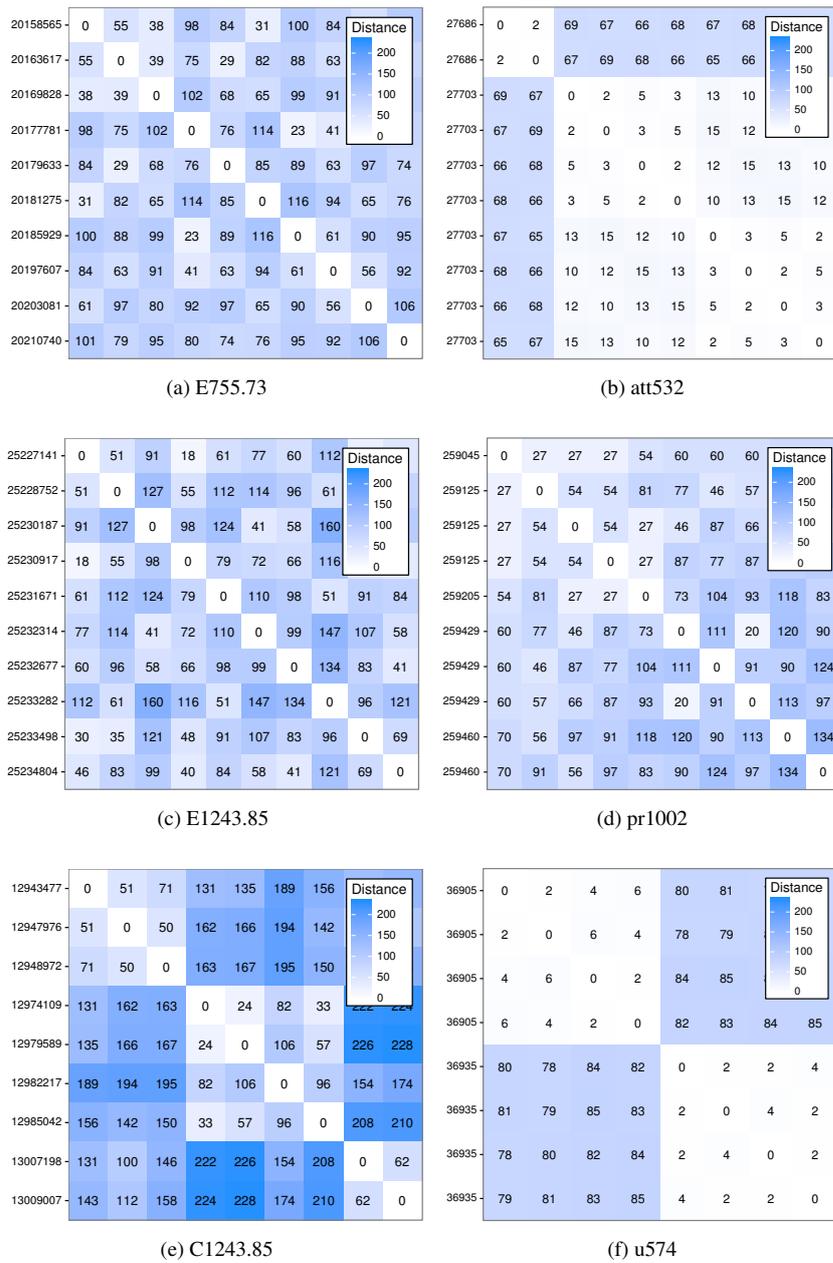


Fig. 7: Heatmaps for pairwise bond distance between sink nodes or nodes within sink-plateaus. Labels on the y-axis indicate objective function but omitted on the x-axis. Only the 10 best nodes are considered for E1243.85 and pr1002. For both, the nodes considered are sink nodes and are not part of a plateau. The two sink-plateaus in att532 and u574 are clearly identified by the nearly-white shading between neighbouring nodes (low bond distance).

with the highest density as the representative of a given sink-plateau. Ties are broken randomly. Each representative node is then compared to the other representative nodes that share the same evaluation. The plot has bond distance as x-axis, frequency of occurrence as y-axis and the objective value of each node pair is colour-coded as shown in the legend.

For the two instances, the distributions appear multimodal with one major peak. This was confirmed using Hartigan’s Dip test (Hartigan and Hartigan 1985; Maechler 2016) which rejected the null hypothesis of unimodality (p-value $< 2.2 \times 10^{-16}$). The smallest bond distance found for u1060 is 17, while it is 42 for fl1577. The largest bond distances are 193 and 199 for u1060 and fl1577, respectively. The distances are fairly large, meaning that it is not easy to move between plateaus sharing the same evaluation. Still, for u1060, the smaller distances may indicate that these is a roundabout way to bridge the disconnected plateaus. For fl1577, the larger distances indicate it is unlikely that the presence of multiple sink-plateaus is purely the result of the sampling process. However, we cannot conclusively confirm that the sink-plateaus found are actually separate either. There might be some pathways between sink-plateaus with the same evaluation if plateaus were explored for more than 10 000 iterations.

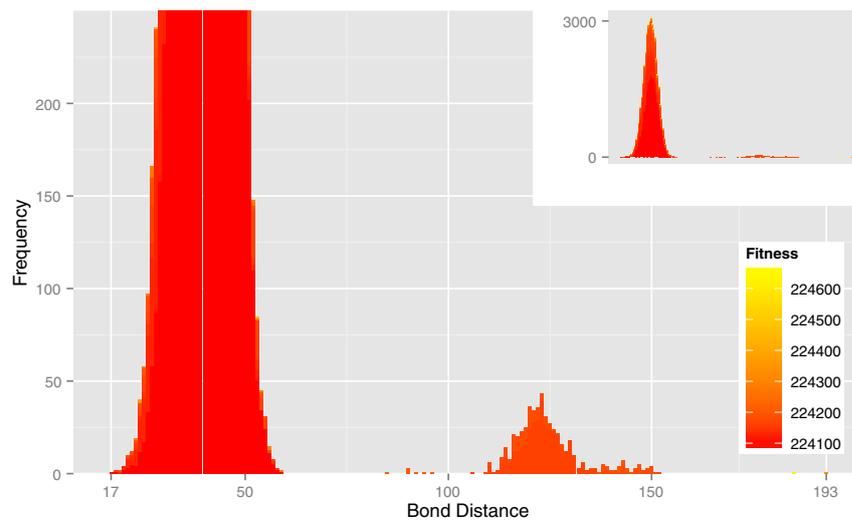
5.5 Effects of sampling methodology

In this subsection we investigate the influence of the stopping criterion and of the number of runs on the networks and their metrics. In a sense, these parameters affect the “depth” and “width” of the sampled networks.

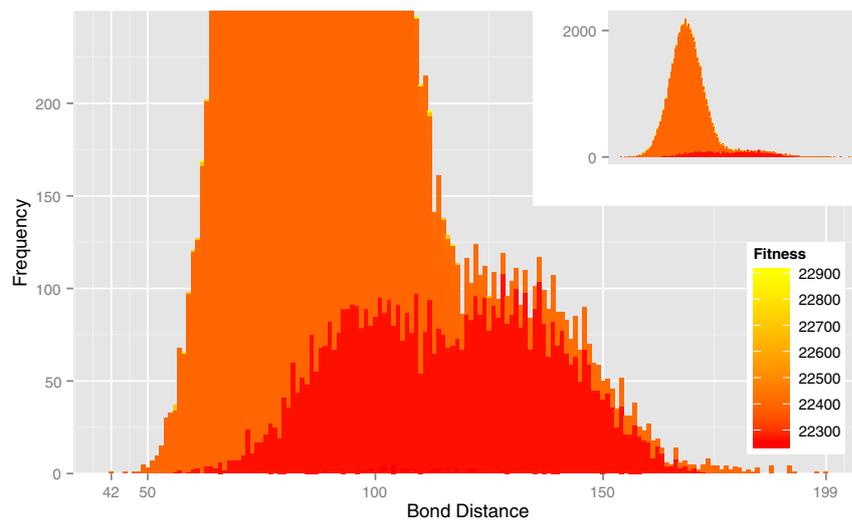
We first consider the impact of using fewer than 10 000 consecutive iterations without improvement as the stopping criterion, and whether this threshold is sufficient for the different metrics to converge. For 18 of the 20 instances, we generate networks for 1000 to 10 000 iterations without improvement, with steps of 1000. The instances left out are those with high neutrality, u1060 and fl1577, for which 1000 runs are not enough to find common solutions between runs. For these instances, a new sampling methodology that considers the extensive neutrality needs to be devised, but this is left as future work.

For each (metric, instance) couple, values outside the $[0, 1]$ range are normalised and, for each stalling iteration increment, the average is computed across the instances. Results are presented in Figure 9 where each line joins the average values across instances for each metric. The success rate is highlighted by the red line. Some of the curves are markedly non-monotonic. This is an artefact of using independent subsets of runs for each stalling threshold. In addition, the normalisation exacerbates the differences between those subsets.

As can be observed, on average the different metrics have approximately converged when 10 000 non-improving iterations are used as stopping criterion. For many metrics the majority of the change in values occurs before the 5000 iterations mark. For a few (a , a_p and s_p), this happens later, around the 9000 iterations mark. The two weighted degree metrics (\bar{d} and d^{max}) naturally do not converge since more connections between already discovered nodes appear, especially in sink plateaus.



(a) u1060



(b) f1577

Fig. 8: Frequency of occurrence of bond distances when considering one node per sink-plateau of u1060 and f1577. For each objective function level of sink-plateaus, bond distances are computed between all pairs of representative nodes. Columns are stacked and coloured according to the objective value of the plateau. The full plot is inset on the top right and the main plot is a zoomed-in version showing the bottom region.

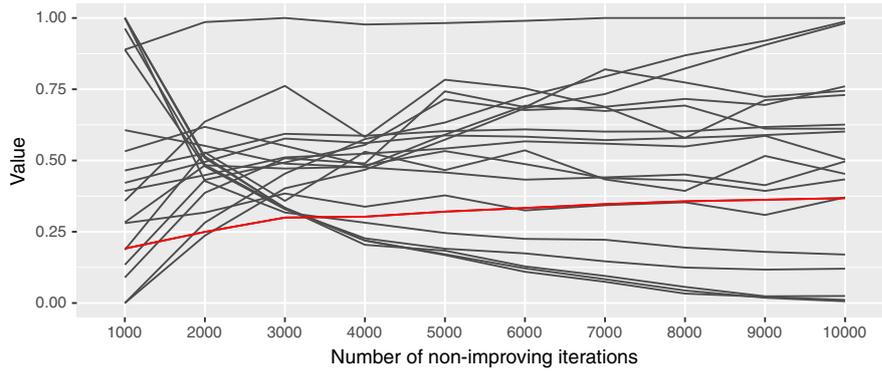


Fig. 9: Influence of stopping criterion on metrics. Each line represents the average value of one metric across instances. The red line is the success rate.

Overall, these findings suggest that the stopping criterion is appropriate and the sampled landscapes would not differ greatly with a higher threshold.

We repeated a similar experiment for the number of runs with values from 100 to 1000 with steps of 100. The stopping criterion is fixed to 1000. Normalisation was performed where necessary as described earlier. We discarded a few data points for which no global optimum was found since several of our metrics take the global optima into account. Results are presented in Figure 10.

In this scenario, the metrics fall within two broad categories: ones that are relatively constant and those whose value increases. The latter are metrics that depend on the size of the sample: for instance, the number of local optima and unique evaluations, the in-strength, and the maximum length of a path to a global optimum. The number of clusters and of attractor and sink plateaus also increases on average across the instances. However, when looking at instances individually, they have converged by 1000 runs for smaller instances such as att532. Let us note that the number of global optima remains more or less constant, except for those instances that have a lot of them.

Metrics that maintain relatively constant values are those describing properties of the sampled network with respect to the whole: the proportion of nodes for which there is a path to a global optimum, the average length of such a path, the relative sizes of funnels and the relative in-strength of global optimum sinks. On the whole, these metrics might seem preferable since they do not depend on the number of runs. Nevertheless, metrics where convergence with respect to the number of runs is unlikely may still be useful and some arbitrary number of runs should be chosen, probably related to the available computing budget.

5.6 Correlation study

In order to obtain more general insights into how different landscape features affect search difficulty, we conduct a correlation study of the different metrics studied in

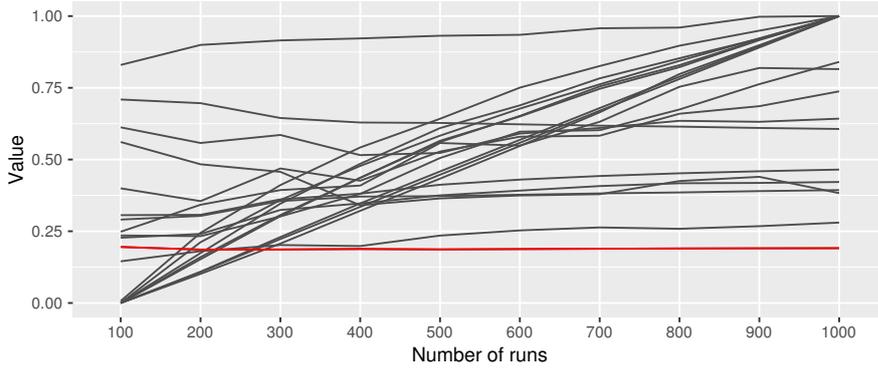


Fig. 10: Influence of number of runs on metrics. Each line represents the average value of one metric across instances. The red line is the success rate.

Table 5: Strongest correlations between success rate and features. T indicates the type of feature: instance (I) or network structure (S) related.

Uniform Instances			Clustered Instances		
T	Feature	Corr.	T	Feature	Corr.
S	Rel. in-strength of sub-opt. sinks (d_{ngo})	-0.95	S	Rel. in-strength of sub-opt. sinks (d_{ngo})	-0.84
I	Sum of lowest edge values	-0.85	S	Path length to global opt. (l_{go}^-)	-0.71
S	Avg. path length to global opt. (l_{go}^-)	-0.85	S	Unique opt. per unique evals ($n/evals$)	-0.65
I	Number of cities	-0.84	S	Number of unique local optima (n)	-0.65
I	Edge value mode frequency	-0.84	S	Num. of unique evaluations ($evals$)	-0.65
I	Nearest-neighbour dist. mean	0.89	I	Nearest-neighbour dist. median	0.58
I	MST distance median	0.89	S	Rel. size of largest funnel (f_{lg})	0.60
I	MST distance mean	0.89	I	MST distance mean	0.60
I	MST distance sum	0.89	S	Rel. size of global opt. funnel (f_{go})	0.66
S	Rel. in-strength of opt. sinks (d_{go})	0.95	S	Rel. in-strength of opt. sinks (d_{go})	0.84

this paper with respect to the success rate across 1000 runs. To assess the relative quality of local optima network and global structure metrics as estimators of success, we also include 64 features based on TSP instance characteristics. These include features that describe the edge cost distribution, cluster characteristics and minimum spanning trees (MSTs). They are computed using the `tspmeta` R package (Mersmann et al 2013).

Pearson correlations are calculated across two separate sets of instances: 100 uniform and 100 clustered instances. For the sake of clarity, only the top 5 positively and negatively correlated features are presented in Table 5.

We found both instance and network structure features that strongly correlate with Chained-LK success rate. The features showing the strongest correlations differ between both sets of instances. On clustered instances, a higher number of structural features show a strong correlation while the opposite is observed on uniform instances. However, on the two instance sets, the average path length to a global

optimum, \bar{l}_{go} , is strongly correlated to success. Also, d_{go} , the normalised incoming strength of global-optimum sinks and its complement, d_{ngo} , the incoming strength of non-globally optimum sinks show strong correlations. This is because sub-optimal sinks act as traps to the search process, from where Chained-LK search cannot escape with its perturbation operator. For clustered instances, several of the top correlations are related to neutrality (n , $evals$, $n/evals$).

When considering all network structure metrics discussed previously in this paper, almost all of them produced correlation values below -0.4 or above 0.4 . The features with low correlation values for clustered instances are the number of connected components (c), the proportion of solutions in more than one funnel (f_{ov}) and the mean and maximum in-strength (\bar{d} and d^{max}). In contrast, for uniform instances, the d^{max} shows a strong correlation (0.74) and it is only f_{ov} that exhibits a low correlation value with success. Overall, this suggests that the metrics studied reflect the search dynamics of Chained-LK.

6 Conclusions

Revealing what makes a combinatorial problem hard remains an open challenge. In this quest, understanding the global structure of the underlying landscapes is essential. There are few attempts in the literature to analyse, let alone visualise, the global structure of combinatorial landscapes. This is due in part to the lack of adequate tools to study their complexity. Local optima networks help to fill this gap. In our sample of TSP problem instances, we found evidence of multiple funnels, instead of a single big-valley as previously believed, in TSP landscapes of moderate size (500 to over 1500 cities). Good local optima decompose into multiple valleys of different depths, each channelling the search process to a separate low cost solution or group of solutions. We also found evidence of high amounts of neutrality or extensive plateaus at the local optima network level in several of the structured instances (i.e. city and drilling problems).

Our data-driven approach models realistic landscapes as networks, empirically characterises the notion of funnels according to notions from theoretical chemistry, and proposes novel 2D and 3D network visualisations. This brings new quantitative and visual insights into landscapes' global structure. The 3D plots provide a concrete and intuitive illustration of the fitness landscape and funnel metaphors. The depiction of edges incoming strength as node sizes, visually reveals the strong attractors of the search process. We also propose alternative visualisation tools for analysing very neutral, and conduct a distance analysis. Finally, we conduct a correlation study between Chained-LK success rate and both landscape and TSP instance features.

We found significant differences among the studied instance classes. Randomly generated instances have a single global optimum and reveal null or very small neutrality. On the other hand, structured instances from the TSPLIB generally have more than one different global optima, with some instances featuring a large global optimum plateau. This is probably because there are many pairwise city distances with the same value. Within the same instance class, the number of funnels generally increases, while the size of the funnel containing the global optima generally de-

creases with the size of the instance. However, the instance class strongly influences the global structure. The random uniform instances revealed a much larger number of funnels as compared with the other instances studied, which explains why these instances are generally harder to solve with heuristic search methods.

Our correlation study reveals that Chained-LK success is strongly correlated to several landscape structural features. On clustered instances a higher number of structural features show high correlations, as compared to the uniform instances where a large number of instance features show strong correlations. For both instance types, however, the features showing the strongest correlation are the incoming weighed degree of the global optimal sink and its complement, the incoming weighed degree of the sub-optimal sinks. This confirms the importance of landscapes global structure in explaining search difficulty: sub-optimal sinks act as traps to the search process, from where Chained-LK cannot escape with its perturbation operator.

The impact of neutrality on search difficulty is harder to assess, and this will motivate future work with additional sampling mechanisms to explore the extent of local optima plateaus. Preliminary experiments adding a stronger perturbation to Chained-LK proved to help in *smoothing* the funnel structure, that is, reducing the number of funnels and making the global optima more reachable (Ochoa and Veerapen 2016a). Future work will expand this study, explore the role of crossover operators in landscapes with multiple funnels, and apply the methodology to other combinatorial optimisation problems.

Acknowledgements This work was supported by the Leverhulme Trust ('The Cartography of Computational Search Spaces', award number RPG-2015-395). All data generated during this research will be made available from the Stirling Online Repository for Research Data, upon acceptance. Results were obtained using the EPSRC funded ARCHIE-WeSt High Performance Computer (www.archie-west.ac.uk, EP-SRC grant EP/K000586/1).

References

- Applegate D, Bixby R, Chvátal V, Cook W (2003a) Concorde TSP solver. URL <http://www.math.uwaterloo.ca/tsp/concorde.html>
- Applegate D, Cook W, Rohe A (2003b) Chained Lin-Kernighan for Large Traveling Salesman Problems. *INFORMS Journal on Computing* 15:82–92
- Applegate DL, Bixby RE, Chvátal V, Cook WJ (2007) *The Traveling Salesman Problem: A Computational Study*. Princeton University Press
- Becker OM, Karplus M (1997) The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *The Journal of chemical physics* 106(4):1495
- Berry RS, Kunz RE (1995) Topography and dynamics of multidimensional interatomic potential surfaces. *Phys Rev Lett* 74:3951–3954
- Bischl B, Mersmann O, Trautmann H, Preuß M (2012) Algorithm selection based on exploratory landscape analysis and cost-sensitive learning. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2012, GECCO '12*, ACM, New York, NY, USA, pp 313–320
- Boese KD, Kahng AB, Muddu S (1994) A new adaptive multi-start technique for combinatorial global optimizations. *Operations Research Letters* 16:101–113
- Csardi G, Nepusz T (2006) The igraph software package for complex network research. *InterJournal Complex Systems*:1695
- Daolio F, Tomassini M, Verel S, Ochoa G (2011) Communities of minima in local optima networks of combinatorial spaces. *Physica A: Statistical Mechanics and its Applications* 390(9):1684–1694

- Doye JPK, Miller MA, Wales DJ (1999) The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. *Journal of Chemical Physics* 110(14):6896–6906
- Flamm C, Hofacker IL, Stadler PF, Wolfinger MT (2002) Barrier Trees of Degenerate Landscapes. *Phys Chem* 216:155–173
- Hains DR, Whitley LD, Howe AE (2011) Revisiting the big valley search space structure in the TSP. *Journal of the Operational Research Society* 62(2):305–312
- Hallam J, Prugel-Bennett A (2005) Large barrier trees for studying search. *IEEE Transactions on Evolutionary Computation* 9(4):385–397
- Hartigan JA, Hartigan PM (1985) The Dip Test of Unimodality. *The Annals of Statistics* 13(1):70–84
- Herrmann S, Ochoa G, Rothlauf F (2016) Communities of Local Optima as Funnels in Fitness Landscapes. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2016, GECCO '16*, ACM, New York, NY, USA, pp 325–331
- Hordijk W, Fontanari JF, Stadler PF (2003) Shapes of tree representations of spin-glass landscapes. *Journal of Physics A-Mathematical and Theoretical* 36(13):3671
- Iclanzan D, Daolio F, Tomassini M (2014) Data-driven local optima network characterization of QAPLIB instances. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2014, GECCO '14*, ACM, New York, NY, USA, pp 453–460
- Kauffman S, Levin S (1987) Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* 128:11–45
- Kerschke P, Preuss M, Wessing S, Trautmann H (2015) Detecting funnel structures by means of exploratory landscape analysis. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2015, GECCO '15*, ACM, New York, NY, USA, pp 265–272
- Lin S, Kernighan BW (1973) An Effective Heuristic Algorithm for the Traveling-Salesman Problem. *Operations Research* 21:498–516
- Locatelli M (2005) On the Multilevel Structure of Global optimization problems. *Computational Optimization and Applications* pp 5–22
- Lunacek M, Whitley D (2006) The dispersion metric and the CMA evolution strategy. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2006, GECCO '06*, ACM, New York, NY, USA, pp 477–484
- Lunacek M, Whitley D, Sutton AM (2008) The impact of global structure on search. In: *Parallel Problem Solving from Nature - PPSN X, 10th International Conference Dortmund, Germany, September 13-17, 2008*, Proceedings, pp 498–507
- Maechler M (2016) diptest: Hartigan's Dip Test Statistic for Unimodality - Corrected. URL <https://CRAN.R-project.org/package=dipstest>, r package version 0.75-7
- Martin O, Otto SW, Felten EW (1992) Large-step markov chains for the TSP incorporating local search heuristics. *Operations Research Letters* 11:219–224
- Mersmann O, Bischl B, Trautmann H, Wagner M, Bossek J, Neumann F (2013) A novel feature-based approach to characterize algorithm performance for the traveling salesperson problem. *Annals of Mathematics and Artificial Intelligence* 69(2):151–182
- Merz P, Freisleben B (1998) Memetic algorithms and the fitness landscape of the graph bi-partitioning problem. In: *Proceedings of Parallel Problem Solving from Nature, PPSN V*, Springer-Verlag, Lecture Notes in Computer Science, vol 1498, pp 765–774
- Noack A (2009) Modularity clustering is force-directed layout. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 79(2)
- Ochoa G, Veerapen N (2016a) Additional Dimensions to the Study of Funnels in Combinatorial Landscapes. In: *Proceedings of the Genetic and Evolutionary Computation Conference 2016, GECCO '16*, ACM, New York, NY, USA, GECCO '16, pp 373–380
- Ochoa G, Veerapen N (2016b) Deconstructing the Big Valley Search Space Hypothesis. In: *Evolutionary Computation in Combinatorial Optimization, EvoCOP 2016*, Springer International Publishing, LNCS, vol 9595, pp 58–73
- Reeves CR (1999) Landscapes, operators and heuristic search. *Annals of Operations Research* 86:473–490
- Reinelt G (1991) TSPLIB—A Traveling Salesman Problem Library. *ORSA Journal on Computing* 3(4):376–384, URL <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>
- Stadler PF (2002) Fitness landscapes. *Appl Math and Comput* 117:187–207
- Tomassini M, Vérel S, Ochoa G (2008) Complex-network analysis of combinatorial spaces: The NK landscape case. *Phys Rev E* 78(6):066,114

-
- Verel S, Ochoa G, Tomassini M (2011) Local optima networks of NK landscapes with neutrality. *IEEE Transactions on Evolutionary Computation* 15(6):783–797
- Verel S, Daolio F, Ochoa G, Tomassini M (2012) Local optima networks with escape edges. In: *Proceedings of the International Conference on Artificial Evolution, EA-2011*, Springer, Lecture Notes in Computer Science, vol 7401, pp 49–60
- Wales DJ (2005) Energy landscapes and properties of biomolecules. *Physical Biology* 2(4):S86–S93