Local optima networks for continuous fitness landscapes

3 authors:

Jason Adair
University of Stirling
5 PUBLICATIONS   **2** CITATIONS

SEE PROFILE

Gabriela Ochoa
University of Stirling
165 PUBLICATIONS   **3,340** CITATIONS

SEE PROFILE

Katherine M. Malan
University of South Africa
41 PUBLICATIONS   **323** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

DAASE Project View project

Local Optima Networks for Continuous Fitness Landscapes

Jason Adair
University of Stirling
Stirling, United Kingdom
jason.adair@stir.ac.uk

Gabriela Ochoa
University of Stirling
Stirling, United Kingdom
gabriela.ochoa@stir.ac.uk

Katherine M. Malan
University of South Africa
Pretoria, South Africa
malankm@unisa.ac.za

ABSTRACT

Local Optima Networks (LONs) have been proposed as a coarse-grained model of discrete (combinatorial) fitness landscapes, where nodes are local optima and edges are search transitions based on an exploration search operator. This paper presents one of the first complex network analysis of continuous fitness landscapes. We use benchmark functions with well-known global structure, and an existing implementation of a Basin-Hopping algorithm to extract the networks. We also explore the impact of varying the Basin-Hopping perturbation step-size. Our results suggest that the landscape's connectivity pattern (global structure) strongly varies with the perturbation step-size, with extreme values of this parameter being detrimental to search and fragmenting the global structure. Our LON visualisations strikingly illustrate the landscape's global (funnel) structure, indicating that LONs serve as a tool for visualising high-dimensional functions.

CCS CONCEPTS

• **Computing methodologies** → **Search methodologies**;

KEYWORDS

Fitness Landscapes, Local Optima Networks, Continuous Optimization, Basin-Hopping, Global Structure, Funnels

1 INTRODUCTION

Complex optimisation problems are generally non-convex and multi-modal (i.e have multiple local optima and potentially multiple global optima). Moreover, in real-world scenarios the multiple local optima are likely to conform to some global structure, instead of being distributed uniformly at random in the search space.

Local Optima Networks (LONs) [21] are a coarse-grained model of fitness landscapes inspired by work on energy surfaces in theoretical chemistry [5]. The idea is to compress the search space

into a smaller mathematical object: a graph, where vertices are the local optima in the underlying landscape and edges are the possible search transitions among optima. A conspicuous limitation of heuristic optimisation is the danger of getting trapped at a local optimum. LONs capture the number, distribution and connectivity pattern of local optima in the underlying landscape. They are, therefore, a useful tool to study the global structure of fitness landscapes. Once a LON has been constructed, a variety of metrics and visualisation tools can be applied to enlighten our understanding of its structure [19]. These features can then be used for performance prediction and enhancing the selection and configuration of suitable optimisation methods to solve the problem at hand.

Despite their inspiration from the study of energy landscapes, the LON model was defined and developed to deal with discrete search spaces and combinatorial optimisation [21, 26], and has only been applied, to the authors' knowledge, to continuous spaces once before [2]. In the discrete case, local minima can be located precisely given sufficient computation time. In the continuous domain this is more difficult; concepts from differential topology, the presence of saddle points, and issues of precision and convergence come into play.

The main contributions of this paper are to:

(1) Adapt the LON model to represent continuous fitness landscapes.
(2) Propose a sampling methodology to extract and construct the networks.
(3) Analyse the funnel structure of well-known benchmark functions.
(4) Explore the effect of increasing the perturbation strength on the global structure of the studied landscapes.

We start with an overview of the notion of a *funnel*. Section 3 presents relevant definitions and algorithms to construct the LON models. Section 4 describes the benchmark instances, sampling procedure, and metrics computed, while 5 presents our results. Finally, Sections 6 and 7 summarises our findings and suggests directions of future work.

2 THE NOTION OF FUNNEL

Funnels can be loosely defined as groups of local optima which are close in configuration space within a group, but well-separated between groups. Funnels also constitutes a coarse-grained gradient towards a low cost optimum. The intuition is captured by Figure 1 where two funnels are depicted. A more formal definition is given in Section 3.3
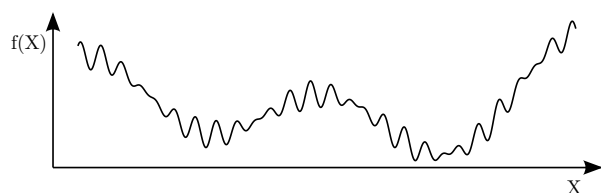
Figure 1: Depiction of two funnels.

2.1 Funnels in Energy Landscapes

In theoretical/computational chemistry, an energy landscape is a mapping of all possible conformations of a molecular entity (clusters, glasses and proteins) to their corresponding energy levels [27]. The structure and dynamics of these molecules are known to be related to the topographical features of their energy landscapes[6, 27].

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." [6]. The energy landscape of proteins is believed to be 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. Funnels have been widely studied on the so called *atomic clusters* (spatial arrangements of atoms), as they represent a convenient benchmark of controllable complexity for which excellent putative global minima have been tabulated [6].

2.2 Funnels in Continuous Fitness Landscapes

Energy landscapes in computational chemistry and fitness landscapes in optimisation are analogous. This is particularly true for continuous optimisation. Locatelli [12] found that search difficulty in continuous optimisation does not relate strictly to the number of local optima, but to how *chaotic* their positions are. He suggests that problems are structured in multiple levels; at each level, different 'objects' are observed, but all levels display a similar structure. This hints towards a fractal structure of fitness landscapes [25].

Lunacek and Whitley [14] propose a *dispersion* metric to quantify 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 [15], the authors studied abstract landscapes with two funnels and find that evolutionary algorithms mostly fail when the global optimum is located in a proportionally smaller funnel.

Work on exploratory landscape analysis of continuous landscapes suggests that multimodality and global structure are key high-level features help to differentiate between problem classes [3, 9]. The presence of funnels, captured with the dispersion metric, was also found to correlate with and explain the performance of Particle Swarm Optimisation algorithms [16].

2.3 Funnels and Local Optima Networks

Recent studies have used LONs to characterise the global structure of combinatorial fitness landscapes revealing in many cases a multi-funnel structure [7, 22, 23]. The presence of sub-optimal funnels

hinders the optimisation success, as the algorithms can become trapped, and thus fail to locate the funnel containing the global optimum.

3 ALGORITHMS AND DEFINITIONS

A full enumeration of the local optima for continuous spaces and benchmark instances of non-trivial size is unmanageable. Therefore, LONs are constructed from a sample of high-quality local minima in the search space. To construct the networks we need to define their nodes and edges. These definitions are closely related to the methodology for extracting the network data, which is based on a number of runs of a *Basin-Hopping* algorithm [11, 28]. The idea is to record and aggregate the local optima visited by several trajectories of the algorithm. Our local minima sample is, therefore, biased towards regions of high-fitness and is not algorithm-agnostic. We argue that this approach is relevant to analyse the global structure of the benchmark fitness landscapes studied. Moreover, we contrasted several perturbation step sizes to define the local optima connectivity. This section describes the Basin-Hopping algorithm and defines the LON models used.

3.1 Basin-Hopping

In computational chemistry, global optimisation algorithms are required to minimise the energy function of molecular conformation problems. A successful algorithm for these problem was termed *Basin-Hopping* by Wales and Doye [28], which is essentially the same as the previous *Monte Carlo-minimisation* algorithm of Li and Scheraga [11]. Interestingly, Basin-Hopping is analogous to the class of algorithms known as *Iterated Local Search* (ILS) in combinatorial optimisation [13].

Both Basin-Hopping and ILS are iterative methods with each cycle composed of the following steps: (i) random perturbation of the incumbent solution (ii) local minimisation, and (iii) acceptance criterion, which accepts or rejects the new solution based on the objective function value. The original Basin-Hopping algorithm depends on the Metropolis acceptance criterion of occasionally accepting uphill moves. A simple yet powerful variant is obtained if the Metropolis acceptance criterion is abandoned in favour of only accepting downhill (or non-deteriorating) steps [10]. Note that this corresponds to setting the Metropolis temperature parameter $T = 0$. This algorithm will follow a descending sequence of local minima until a funnel bottom is reached. Funnel bottoms can be empirically recognised (estimated) by the lack of improvement after a large number of move attempts. We used this *Monotonic Sequence Basin-Hopping* variant [10] (Algorithm 1) in our study in order to extract the LON models and study the landscapes funnel structure.

3.2 Model Description

A fitness landscape is defined as a triplet (X, N, f), where X is the set of configurations or candidate solutions; N is a notion of neighbourhood, distance or accessibility on X; and $f : X \to \mathbb{R}$ is the fitness function.

In the context of continuous optimisation, X is the set $\mathbb{R}^n$ of all possible real-valued solutions to the problem of n dimensions, $f : \mathbb{R}^n \to \mathbb{R}$, and we denote a candidate solution as the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. A common way of defining neighbourhood

---

Algorithm 1 Monotonic Sequence Basin-Hopping

Require: Search space X, Fitness function $f(X)$,
    Perturbation step size p, Stopping threshold R
1: Choose initial random solution $x_0 \in X$
2: $l \leftarrow \text{LocalMinimisation}(x_0)$
3: $r \leftarrow 0$
4: **repeat**
5:     $x' \leftarrow \text{Perturbation}(l, p)$
6:     $l' \leftarrow \text{LocalMinimisation}(x')$
7:     **if** $f(l') \leq f(l)$ **then**
8:         $l \leftarrow l'$
9:         $r \leftarrow 0$
10:    **else**
11:       $r \leftarrow r + 1$
12:    **end if**
13: **until** $r \geq R$
14: **return** l

---

of a solution $\mathbf{x}$ in continuous spaces is the set of points within the hypersphere with some small radius and centre $\mathbf{x}$ [24]. However, this definition of neighbourhood requires Euclidean distance calculations to test for neighbourhood. In this paper we define neighbourhood based on orthotopes (or hyperrectangles).

Formally, the neighbourhood set $N(\mathbf{x}_k)$ of an n-dimensional point $\mathbf{x}_k$ is an n-orthotope defined as follows:

$$\mathbf{x}_j \in N(\mathbf{x}_k) \iff |x_{ki} - x_{ji}| < s_i, \ \forall i \in [1, \ldots, n]$$

where $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ is a vector defining the size of the neighbourhood in all dimensions.

In order to study the multi-funnel structure of continuous fitness landscapes, we considered and adapted the Monotonic (MLON) and Compressed Monotonic (CMLON) models introduced in [23].

Monotonic LON. Is the directed graph $MLON = (L, E)$, where nodes are the local optima L, and edges E are the monotonic perturbation edges.

Local optima. We assume a search space $\mathbb{R}^n$ with a fitness function $f(\mathbf{x})$ and a neighbourhood function $N(\mathbf{x})$. A local optimum, which in the studied benchmark functions is a minimum, is a solution $\mathbf{l}$ such that $\forall \mathbf{x} \in N(\mathbf{l}), f(\mathbf{l}) \leq f(\mathbf{x})$. Notice that the inequality is not strict, in order to allow the treatment of neutrality (local optima of equal fitness), which is generally present in complex real-world problems. The set of local optima, which corresponds to the set of nodes in the network model, is denoted by L.

Monotonic perturbation edges. Edges are directed and based on the perturbation operator (adding a stepsize). There is an edge from local optimum $\mathbf{l}_1$ to local optimum $\mathbf{l}_2$, if $\mathbf{l}_2$ can be obtained after applying a random perturbation to $\mathbf{l}_1$ followed by local minimisation, and $f(\mathbf{l}_2) \leq f(\mathbf{l}_1)$. These edges are called *monotonic* as they record only non-deteriorating transitions between local optima. Edges are weighted with estimated frequencies of transition. The edge weights are estimated after the sampling process. The weight is the number of times a transition between two local optima basins occurred with a given perturbation. The set of edges is denoted by E.

3.3 Compressed Monotonic LON Model

This is a coarser LON model compressing connected local optima at the same fitness level (according to a given accuracy) into single nodes. The instances studied in this paper only show a small amount of neutrality. We consider this model, however, for the sake of generality, and because it allows a crisper characterisation of the funnel bottoms or sinks.

Compressed Monotonic LON. Is the directed graph $CMLON = (CL, CE)$, where nodes are compressed local optima CL, as defined below, and the edges CE are aggregated from the monotonic edge set E by summing up the edge weights.

Compressed local optima. A compressed local optimum is a set of connected nodes in the MLON with the same fitness value. Two nodes in the MLON are connected if there is a perturbation edge between them. The set of connected MLON optima with the same fitness, denoted by CL, corresponds to the set of nodes in the Compressed Monotonic LON model.

Monotonic Sequence. A monotonic sequence is a path of connected nodes $MS = \{\mathbf{cl}_1, \mathbf{cl}_2, \ldots, \mathbf{cl}_s\}$ where $\mathbf{cl}_i \in CL$. By definition of the edges, $f(\mathbf{cl}_i) \leq f(\mathbf{cl}_{i-1})$. There is a natural end to every monotonic sequence, $\mathbf{cl}_s$, when no improving transitions can be found. This corresponds to the funnel 'bottom'. In the directed CMLON network, $\mathbf{cl}_s$ will be a node without outgoing edges (or 'sink')[1].

Funnel. We characterise funnels in the CMLON as the aggregation of all monotonic sequences ending at the same point (funnel bottom or sink). Funnels can be seen as basins of attraction at the level of local optima.

4 EMPIRICAL METHODOLOGY

4.1 Benchmark Functions

Our experiments used three well-known benchmark functions: Ackley's, Rastrigin, and Birastrigin (also known as the Double Rastrigin or Lunacek's function). These are detailed below, and visualised for two variables in Figure 3.

Ackley's Function. This function was initially defined in two dimensions, but the extended form [1] has been used for our experimentation in order to investigate multiple dimensions. It is expressed as

$$f(\mathbf{x}) = f(x_1, ..., x_n) = -20 \exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) -$$
$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right) + 20 + e$$

This function has one global minimum at $[0, 0, \ldots, 0]$ with a fitness of 0.0, and typically evaluated within the range $[-32.768, 32.768]$ for all x_i.

Ackley's function centres around a single funnel with a high number of local minima. While it appears to be relatively simple,

---

[1]In directed graphs, a node without outgoing edges is called a *sink*.

it is an important benchmark as it is reminiscent of the free energy landscape of proteins, and therefore has potential real-world application [4].

Rastrigin Function. Like Ackley's function, Rastrigin has a single funnel landscape, but with considerably fewer local minima. The difficulty of this search space is in the increased distance between those local minima, and flattened search space. These properties mean that local searches tend to become trapped, whereas global searches must possess an adequate step-size. It is expressed as

$$f(\mathbf{x}) = f(x_1, ..., x_n) = 10n + \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i) \right)$$

This function has one global minimum at $[0, 0, \ldots, 0]$ with a fitness of 0.0, and is typically evaluated within the range $[-5.12, 5.12]$ for all x_i.

Birastrigin Function. The Birastrigin (or Lunacek's function [15]) is the resulting hybrid of the double-sphere and Rastrigin functions; effectively creating a double-funnel problem with local optima. This function is intended to create a more difficult benchmark for global optimisation methods that closely resembles those found in real world problems – specifically those in computational chemistry, such as Lennard-Jones clusters [5]. This double-funnel landscape is particularly challenging for population based methods. It is expressed as

$$f(\mathbf{x}) = f(x_1, ..., x_n) = 10 \sum_{i}^{n} (1 - \cos 2\pi(x_i - \mu_1))$$
$$+ \min \left(\sum_{i}^{n} (x_i - \mu_1)^2, \quad d \cdot n + s \cdot \sum_{i}^{n} (x_i - \mu_2)^2 \right)$$

and is typically evaluated within the same domain as the Rastrigin function: $[-5.12, 5.12]$ for all x_i. In this study, the parameter settings used were $d = 1$, $s = 1 - \left(1/(2\sqrt{n + 20} - 8.2) \right)$, $\mu_1 = 2.5$, and $\mu_2 = \sqrt{|(\mu_1^2 - d)/s|}$.

4.2 Sampling Method

The sampling procedure consists of aggregating the local minima and transition edges obtained by 100 runs of basin-hopping (Algorithm 1). The stopping condition was set to 1000 iterations without an improvement. We used the basin-hopping implementation provided in Scipy[8], with the following configuration and parameter values:

Initialisation. The initial solution is a random n-dimensional coordinate vector $\mathbf{x}_0$ with independent identically distributed components generated from a uniform distribution within the given problem bounds.

Local minimisation. The local minimisation step used the limited-memory variant of the Broyden-Fletcher-Goldfarb-Shanno algorithm [20] (L-BFGS-B). This version of BFGS uses less memory by storing only a few vectors instead of the estimation of the entire inverse Hessian matrix, ensuring the scalability of this sampling technique to much larger search-spaces. As this variant is also

Table 1: Base perturbation step sizes (β).

Dimensions (n)	Ackley	Rastrigin	Birastrigin
	Functions		
3	0.4546	0.4746	0.5156
5	0.4646	0.4749	0.4946

bounded, the search can be constrained to the specified domain. The local search was terminated when the change in value of the solution between iterations was less than 10^{-7}, or 15,000 iterations occurred.

At the end of each local optimisation step, the solution found is stored as a local optimum to form the nodes of the LON. In continuous spaces, there is the issue of deciding when two solutions found through different local search runs are close enough in solution space to be regarded as the same local optimum. In this study, we set the local optimum position threshold to 10^{-5} in all dimensions.

Perturbation. The perturbation strength is an essential parameter in defining the search-landscape as it has been shown to have dramatic effects on the shape of said landscape [18]. If it is inadequate, the search will not be able to escape the basin in which the solution originates, but if it is too strong, it will move through the search-space, randomly sampling different regions without learning from the landscape. This is especially problematic when the search has progressed to a region with relatively high fitness, where large perturbations are likely to be deteriorating.

We determined the base perturbation step (β) for each function in 3D and 5D spaces by randomly sampling the space, performing a local search, and perturbating that solution based on a β value. This β was then varied until approximately 50% of the steps resulted in escaping the original local optimum. This method for selecting perturbation strength via a sampling technique was adopted from that used by Leary [10]. The resulting step sizes are detailed in Table 1. For each function and dimension combination, four perturbation step-sizes were used: $p \in \{0.5\beta, \beta, 2\beta, n\beta\}$ (where n is the dimension). The perturbation was performed by applying a displacement to each coordinate of the vector sampled from a uniform distribution within the perturbation step-size, $x' = x + Uniform(-p, p)$.

4.3 Metrics

For each instance, problem dimension, and perturbation step size, we extracted the LON models and computed the measurements described in Table 2. Metrics are reported as aggregations over 100 runs.

Table 2: Metrics.

	Performance Metrics
success	Proportion of runs that reached the global optimum.
deviation	Mean deviation from the global optimum value.
	Network Metrics
noptima	Number of optima (including local and global).
neutral	Proportion of neutral to total number of optima (reciprocal).
nfunnels	Number of funnels (including local and global).
strength	Normalised incoming strength of the global funnel.

5 RESULTS

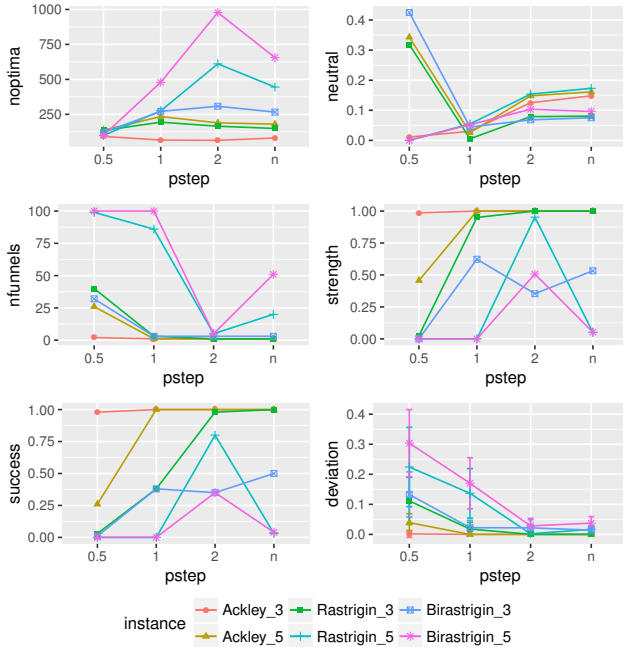## 5.1 Structural and Performance Metrics

Figure 2: Structural and performance metrics (as defined in Table 2) for four perturbation step sizes, which are multiples of the base (β as defined in Table 1).

As seen in Figure 2, the perturbation strength (*pstep*) has substantive effects on the structural metrics derived from LONs. The total number of optima found over the course of the basin-hopping experiments generally increased as more suitable *pstep* values were used, before decreasing as it became excessive. In regards to Ackley's function, 0.5β best explored the search space, increasing to β in 5 dimensions, with a similar effect seen with the Rastrigin peaking at β in 3 dimensions and 2β in 5. This increase in perturbation strength was not required for the Birastrigin function, as 2β was favoured in both investigated dimensions, there was, however, a much more stark increase in discovered optima in the higher dimensional space.

The number of discovered optima correlates well with the number of funnels detected, and the strength of their basins. With 0.5β, we can see that a low number of optima were discovered, but a large number of funnels, indicating that the basin-hopping algorithm rarely escaped the nearest local optima to its origin point. This is most pronounced in the Rastrigin and Birastrigin 5D functions showing a stark decrease in the number of funnels between *pstep* $= \beta$ and *pstep* $= 2\beta$, where we were able to uncover the underlying single and double funnel landscapes. This suggests that, the right perturbation step size better exploits the underlying global structure.

This is further reinforced in the success rate, that is, the percentage of runs finding the global optimum. With 0.5β, only the simpler

Ackley's function showed any signs of success, with higher dimensionality requiring β, much in line with the number of observed optima and funnels. Most notably, the success rates on the 5D Rastrigin and Birastrigin functions share a very a similar behaviour to the strength, number of optima, and number of funnels. We can also see a similar detrimental behaviour to increased βs, suggesting that higher dimensional functions may be more sensitive to excessive perturbation sizes than those with lower dimensions.

The mean distance of the final solutions found from the global optimum (deviation) clearly displays a decrease in accordance with the strength of the perturbation parameter. When considering this with the success rates, we can suggest that larger step sizes in basin hopping may decrease the likelihood of finding the global optimum, but increases the chances of achieving, on average, a good solution.

5.2 Visualisation

Figure 3 provides a 3D representation of the three benchmark instances selected. For practical purposes, we can only visualise functions with two variables in this way as there is no clear way of visually conveying fitness landscapes for functions with many variables.

One of the advantages of modelling fitness landscapes with LONs is the possibility of visualising these higher dimensional functions. To illustrate this point, Figure 4 visualises the compressed monotonic LONs for the three benchmark functions with five variables, $n = 5$. In these plots, the nodes are local minima, and edges are perturbation transitions with step size $p = 2\beta$ (the values of β, for each function and dimension can be found in Table 1). The size of nodes is proportional to their incoming weighted degree (also called strength), which indicates how much a node 'attracts' the search process. Node colours indicate funnel membership, with pink highlighting the funnel containing the global optimum, and light blue demonstrating the sub-optimal funnels. The red node identifies the global optimum, while the dark blue nodes represent the bottom of sub-optimal funnels (which we also call sinks).

By contrasting the images in Figures 3 and 4, we can appreciate the overall resemblance of the respective functions. In both cases, Ackley and Rastrigin have a single funnel structure, while Birastrigin has a double funnel structure. The smoother shape of the Rastrigin funnels can also appreciated.

In order to illustrate the effect of the perturbation step-size in the LONs connectivity pattern, Figure 5 shows 2D LON visualisations for the three benchmark functions with 5 variables $n = 5$ and three increasing perturbation step-sizes, *pstep*, as indicated in the sub-figures captions. It is worth noting that the sampling process aggregates local optima and transition edges from 100 Basin-Hopping independent runs, so the maximum possible number of estimated funnels is 100. The success rate achieved by these 100 independent runs is also shown in the sub-figures caption.

In the images, the size of nodes is proportional to their incoming strength, and the colours indicate funnel membership. The pink nodes belong to the funnel containing the global optimum, while light blue nodes belong to sub-optimal funnels. The red node (when visible) is the global optimum, while dark blue nodes indicate the sub-optima funnel bottoms or sinks.
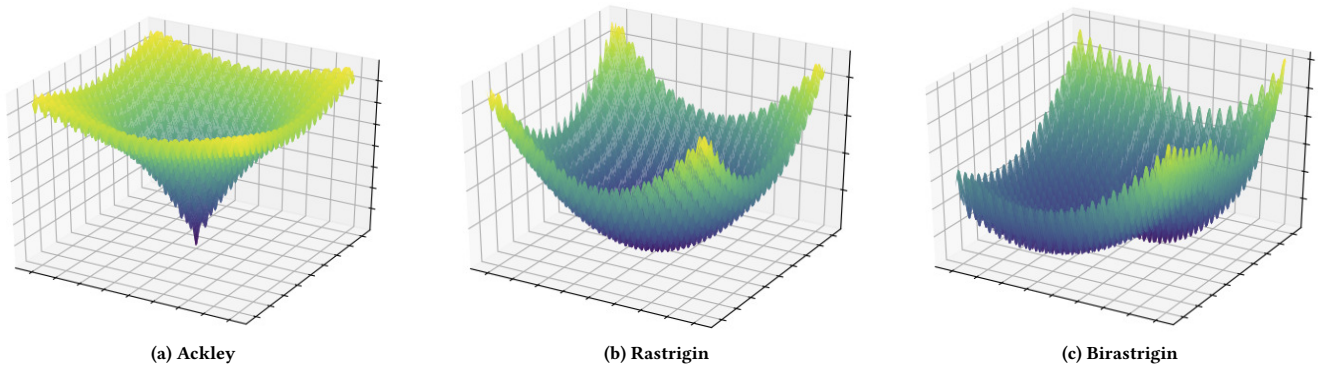
(a) Ackley

(b) Rastrigin

(c) Birastrigin

Figure 3: 3D visualisations for the three benchmark functions with two variables $n = 2$

(a) Ackley, $pstep = 2\beta$

(b) Rastrigin, $pstep = 2\beta$
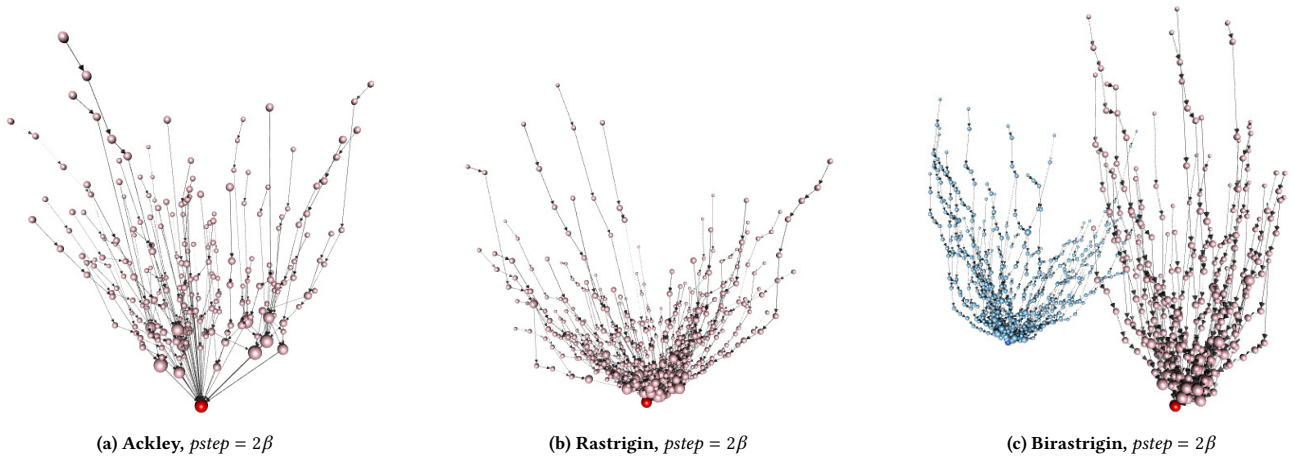
(c) Birastrigin, $pstep = 2\beta$

Figure 4: 3D LON visualisations for the three benchmark functions with five variables $n = 5$ and perturbation step-size $p = 2\beta$. The z coordinate represents fitness. The size of nodes is proportional to their incoming strength. Red highlights the global optimal funnels, while blue suboptimal funnels. Funnel bottoms (sinks) are indicated in more intense colours.

With the lowest perturbation step in the Ackley function (plot (a)) only a portion of the runs (*success* = 0.26) reached the global optimum. The successful trajectories are indicated by the pink nodes ending at the global optimum (red node). The remaining runs end up trapped in sub-optimal sinks, with trajectory lengths of one or two hops. For the largest perturbation step (plots (b) and (c)) a single funnel can now be appreciated, where all the search trajectories converge to the global optimum.

For the Rastrigin function and a low perturbation step (plot (d)), the global optimum cannot be reached; the 100 runs all end up trapped in a sub-optimal funnel after a few search transitions. For a perturbation step of 2β (plot (e)) a single funnel is observed, with most trajectories converging to the global optimum (which is hidden behind the pink nodes in this 2D projection). Increasing the perturbation step to 5β (plot (f)) causes the landscape to fragment again into multiple funnels, with several trajectories trapped in sub-optimal sinks.

The Birastrigin function shows a similar trend to the Rastrigin function, with the differences being observed at the intermediate perturbation step-size $pstep = 2\beta$ (plot (h)), where two funnels appear, with the sub-optimal funnel in blue attracting a larger proportion (65%) of the search trajectories.

6 DISCUSSION

This proof-of-concept paper shows that it is possible to use LONs to model and visualise the global structure of continuous search spaces. There are, however, a number of limitations and challenges regarding the sampling methodology that need to be addressed before the approach can be applied more widely.

One challenge relates to the sampling of initial solutions to be used as the starting points for local minimisation. The approach used in this paper is to sample each component from a uniform random distribution. A problem with this approach is that, as the number of dimensions increase, the position of the full solution vector is biased towards the centre of the search space. In addition,
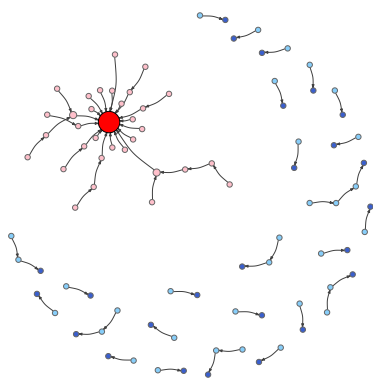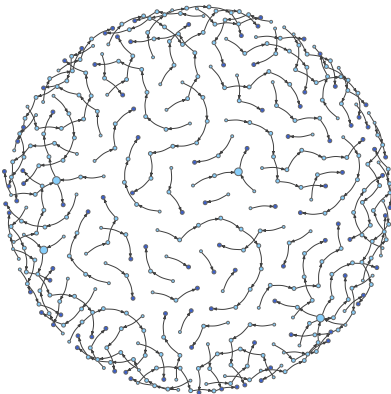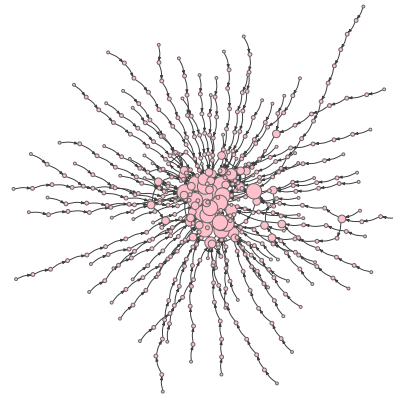
(a) **Ackley**, *pstep* = 0.5*β*, *success* = 0.26
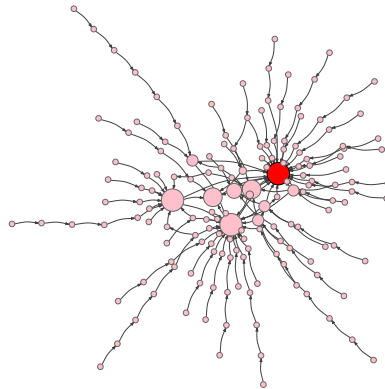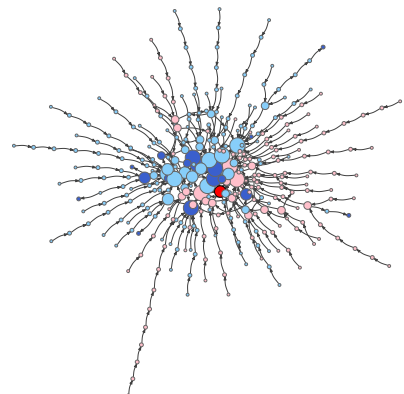
(b) **Ackley**, *pstep* = 2*β*, *success* = 1.0
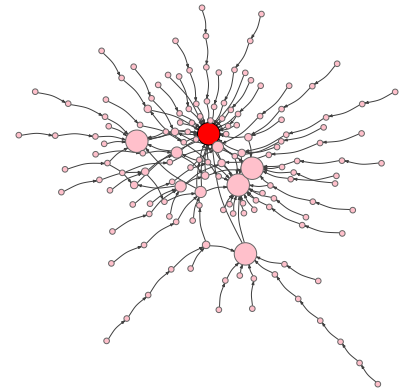
(c) **Ackley**, *pstep* = 5*β*, *success* = 1.0

(d) **Rastrigin**, *pstep* = 1*β*, *success* = 0.0

(e) **Rastrigin**, *pstep* = 2*β*, *success* = 0.8

(f) **Rastrigin**, *pstep* = 5*β*, *success* = 0.03
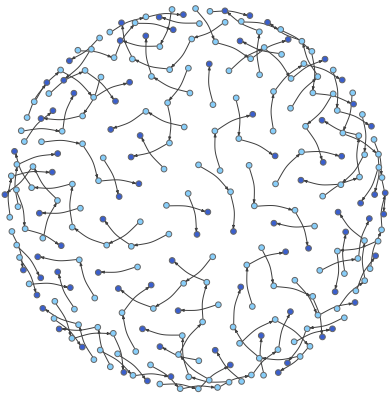
(g) **Birastrigin**, *pstep* = 1*β*, *success* = 0.0
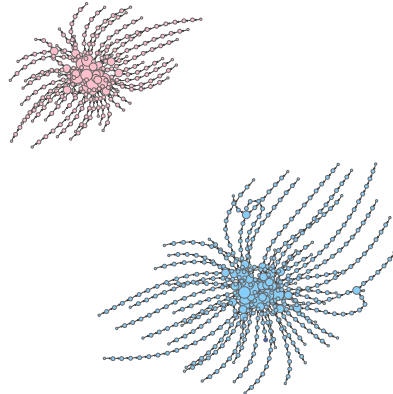
(h) **Birastrigin**, *pstep* = 2*β*, *success* = 0.35
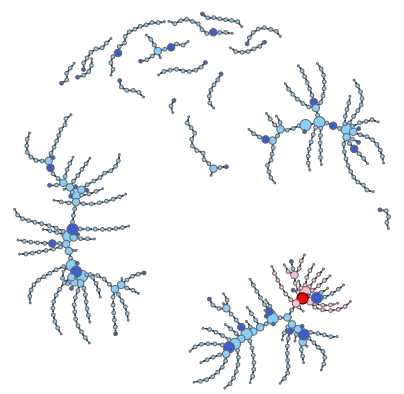
(i) **Birastrigin**, *pstep* = 5*β*, *success* = 0.04

Figure 5: Local optima networks for the three benchmark functions with five variables $n = 5$ and three perturbation step-sizes, p, as indicated in the captions. The size of nodes is proportional to their incoming strength. Red highlights the global optimal funnels, while blue sub-optimal funnels. Funnel bottoms (sinks) are indicated in more intense colours.

each initial solution is sampled independently and so there is no guarantee that the different starting positions provide a good coverage of the search space. Latin hypercube sampling [17] is a random approach that provides a better distribution in multidimensional spaces than uniform random sampling. Future work could investigate the effect of the sampling approach of initial solutions on the LON models.

The choice of method for performing local optimisation is an important consideration in the sampling for LON construction. In this study we used a variant of the BFGS algorithm, which is a quasi-Newton method that uses an approximation of the gradient of the objective function to direct the search. An alternative local search approach may need to be used for black-box optimisation scenarios with expensive function evaluations. There are, however, scenarios where the gradient information is available.

The approach used in this study is computationally expensive. It may not be necessary to perform as many runs of the Basin-Hopping algorithm or to perform as many iterations in the local optimisation step. Further work could investigate cheaper ways to sample continuous search spaces in order to construct LONs that are less detailed, but still informative.

7 CONCLUSION

This paper detailed a methodology of extracting local optima networks from continuous functions. We demonstrated that the perturbation strength of the Basin-Hopping algorithm is an essential parameter to tune in order to extract the underlying nature of the function – landscapes appear fractured at values without appropriate strengths, coming together to form the single and double-funnel structures known to exist in the problem domain when the perturbation strength is better chosen. We show that LONs can be used as a tool for visualising fitness landscapes, displaying that we can preserve the fundamental topology of the high-dimensional functions. Future work will address the limitations discussed, aiming at analysing larger dimensions and real-world problems.

REFERENCES

[1] David H. Ackley. 1987. *A Connectionist Machine for Genetic Hillclimbing.* Kluwer Academic Publishers, Norwell, MA, USA.
[2] Andrew J. Ballard, Ritankar Das, Stefano Martiniani, Dhagash Mehta, Levent Sagun, Jacob D. Stevenson, and David J. Wales. 2017. Energy landscapes for machine learning. *Physical Chemistry Chemical Physics* 19, 20 (2017), 12585–12603. https://doi.org/10.1039/c7cp01108c
[3] Bernd Bischl, Olaf Mersmann, Heike Trautmann, and Mike Preuß. 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, 313–320.
[4] Johannes M. Dieterich and Bernd Hartke. 2012. Empirical review of standard benchmark functions using evolutionary global optimization. *CoRR* abs/1207.4318 (2012). arXiv:1207.4318 http://arxiv.org/abs/1207.4318
[5] J. P. K. Doye. 2002. The network topology of a potential energy landscape: a static scale-free network. *Phys. Rev. Lett.* 88 (2002), 238701.
[6] J P K Doye, M A Miller, and D J Wales. 1999. The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. *Journal of Chemical Physics* 110, 14 (1999), 6896–6906.
[7] Sebastian Herrmann, Gabriela Ochoa, and Franz Rothlauf. 2016. Communities of Local Optima as Funnels in Fitness Landscapes. In *Genetic and Evolutionary Computation Conference, GECCO 2016.* ACM, 325–331.

[8] Eric Jones, Travis Oliphant, Pearu Peterson, et al. 2001–. SciPy: Open source scientific tools for Python. (2001–). http://www.scipy.org/ [Online; accessed <today>].
[9] Pascal Kerschke, Mike Preuss, Simon Wessing, and Heike Trautmann. 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, 265–272.
[10] Robert H. Leary. 2000. Global Optimization on Funneling Landscapes. *Journal of Global Optimization* 18, 4 (01 Dec 2000), 367–383. https://doi.org/10.1023/A:1026500301312
[11] Z Li and H A Scheraga. 1987. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences* 84, 19 (1987), 6611–6615. https://doi.org/10.1073/pnas.84.19.6611
[12] Marco Locatelli. 2005. On the multilevel structure of global optimization problems. *Computational Optimization and Applications* 30, 1 (2005), 5–22.
[13] Helena R Lourenço, Olivier C Martin, and T. Stützle. 2003. Iterated Local Search. *Handbook of Metaheuristics* (2003), 320–353.
[14] Monte Lunacek and Darrell Whitley. 2006. The dispersion metric and the CMA evolution strategy. In *Genetic and Evolutionary Computation Conference, GECCO 2006.* 477–484.
[15] Monte Lunacek, Darrell Whitley, and Andrew M. Sutton. 2008. The Impact of Global Structure on Search. In *Parallel Problem Solving from Nature - PPSN X (LNCS)*, Vol. 5199. Springer, 498–507. https://doi.org/10.1007/978-3-540-87700-4_50
[16] K. M. Malan and A. P. Engelbrecht. 2013. Ruggedness, funnels and gradients in fitness landscapes and the effect on PSO performance. In *IEEE Congress on Evolutionary Computation, CEC 2013.* IEEE, 963–970.
[17] M. D. Mckay, R. J. Beckman, and W. J. Conover. 2000. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics* 42, 1 (2000), 55–61.
[18] Paul McMenemy, Nadarajen Veerapen, and Gabriela Ochoa. 2018. How perturbation strength shapes the global structure of TSP fitness landscapes. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10782 LNCS (2018), 34–49. https://doi.org/10.1007/978-3-319-77449-7_3
[19] M. E. J. Newman. 2010. *Networks: An Introduction.* Oxford University Press, Oxford, UK.
[20] George Nocedal and Stephen J. Wright. 1999. Quasi-Newton Methods. In *Numerical Optimization.* Springer, New York, Chapter 8.
[21] Gabriela Ochoa, Marco Tomassini, Sebástien Vérel, and Christian Darabos. 2008. A study of NK landscapes' basins and local optima networks. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation.* ACM, 555–562.
[22] Gabriela Ochoa and Nadarajen Veerapen. 2017. Mapping the global structure of TSP fitness landscapes. *Journal of Heuristics* (2017), 1–30.
[23] Gabriela Ochoa, Nadarajen Veerapen, Fabio Daolio, and Marco Tomassini. 2017. Understanding Phase Transitions with Local Optima Networks: Number Partitioning as a Case Study. In *European Conference on Evolutionary Computation in Combinatorial Optimization.* Springer, 233–248.
[24] El-Ghazali Talbi. 2009. *Metaheuristics: From Design to Implementation.* John Wiley & Sons, Inc., Hoboken, New Jersey.
[25] Sarah L. Thomson, Sébastien Vérel, Gabriela Ochoa, Nadarajen Veerapen, and Paul McMenemy. 2018. On the Fractal Nature of Local Optima Networks. In *Evolutionary Computation in Combinatorial Optimisation, EvoCOP 2018 (Lecture Notes in Computer Science)*, Vol. 10782. Springer, 18–33.
[26] S. Verel, G. Ochoa, and M. Tomassini. 2011. Local Optima Networks of NK Landscapes with Neutrality. *IEEE Transactions on Evolutionary Computation* 15, 6 (2011), 783–797.
[27] David J Wales. 2005. Energy landscapes and properties of biomolecules. *Physical Biology* 2, 4 (2005), S86–S93.
[28] David J. Wales and Jonathan P. K. Doye. 1997. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A* 101, 28 (1997), 5111–5116. https://doi.org/10.1021/jp970984n