

On the Fractal Nature of Local Optima Networks

Sarah L. Thomson¹ 0000-0001-6971-7817, Sébastien Verel² 0000-0003-1661-4093,
Gabriela Ochoa¹ 0000-0001-7649-5669, Nadarajen Veerapen¹ 0000-0003-3699-1080, and
Paul McMenemy¹ 0000-0002-5280-425X

¹ Computing Science and Mathematics, University of Stirling, Stirling, UK
{s.l.thomson,gabriela.ochoa,nadarajen.veerapen,paul.mcmenemy}@stir.ac.uk

² Université du Littoral Côte d'Opale, EA 4491 - LISIC, Calais, France
verel@uni-littoral.fr

Abstract. A *Local Optima Network* represents fitness landscape connectivity within the space of local optima as a mathematical graph. In certain other complex networks or graphs there have been recent observations made about inherent *self-similarity*. An object is said to be self-similar if it shows the same patterns when measured at different scales; another word used to convey self-similarity is *fractal*. The *fractal dimension* of an object captures how the detail observed changes with the scale at which it is measured, with a high fractal dimension being associated with complexity. We conduct a detailed study on the fractal nature of the local optima networks of a benchmark combinatorial optimisation problem (NK Landscapes). The results draw connections between fractal characteristics and performance by three prominent metaheuristics: Iterated Local Search, Simulated Annealing, and Tabu Search.

Keywords: Combinatorial Fitness Landscapes, Local Optima Networks, Fractal Analysis, NK Landscapes

1 Introduction

Weinberger and Stadler [1] noticed that certain fitness landscapes exhibit *self-similarity*. They saw that if they increased landscape diameter, they observed patterns of ruggedness that scaled in a way indicative of fractal geometry. In particular, the landscapes showed evidence of having a multilevel structure. The critical question is then whether we can exploit the fractal patterns.

The information used in the aforementioned study was obtained by conducting random walks on the fitness landscape at the solution level. Considering the solution level to be the base of the search space, we can then consider fractal patterns at higher levels of abstraction; to see the extent of fractal geometry in a model of the local optima space would be desirable.

A *local optima network* [2] models the local optima level of a fitness landscape. A network is formed by tracing the search connectivity between optima where the network nodes are local optima. An edge traced between two nodes means that the destination (optimum) can be reached from the source (optimum) by

carrying out a perturbation followed by hill-climbing. This measure of distance captures the notion of neighbourhood in the space of local optima.

Conducting a fractal analysis of a local optima network would give information about patterns at a raised level of abstraction in the fitness landscape. Of particular interest is the *fractal dimension* [3], hereafter denoted as FD. Here, a non-integer dimension can be assigned to a shape as a measure of complexity. This creates an index which quantifies how detail in an object is observed when the scale at which it is measured is changed. Figures 1a and 1b provide examples of this; both are shapes with FD somewhere between one and two, but they have markedly different complexities. We can see that the latter displays a significantly more complex pattern composition, containing much more detail at different scales than the former and filling much more of the overall space which it occupies. The shape in Fig. 1a contains a lot of empty space — a lack of ‘space-filling’ nature — and therefore low fractional dimension. Another perspective is to consider how different from the one-dimensional line the two are. Figure 1a is nearer to one dimension than two, at 1.1292; indeed, we can see that it is effectively a line with some ruggedness or detail. Figure 1b is almost two-dimensional, but is missing complicated segments.

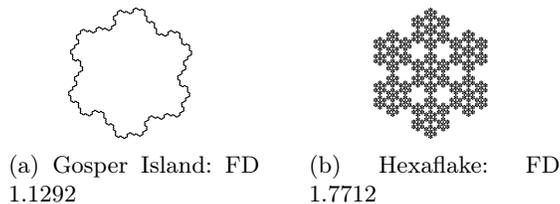


Fig. 1. Two fractals with different dimensions.

The study of complex networks as a field in its own right has created a wealth of measures to understand them, including algorithms for calculating the *FD* of a complex network. One of these, termed a ‘box-counting’ algorithm [4], tries to describe a network with as few *boxes* as possible, with each box containing nodes which are within m links of each other. The parameter m corresponds to the length of measure used in the equation to obtain fractal dimension. The number of boxes (as a proportion of the size of the network) is taken to be the extent of detail observed in the shape when using the scale m .

Because this method is agnostic of the semantics of the network, it can also be used to calculate the FD of a local optima network. However, distance between nodes is not the only important consideration. Node fitness, as well as link-distance, is of great significance in a local optima network. Accordingly, a modification of the box-count algorithm is required. A threshold should be incorporated, for the maximum fitness difference allowed between nodes which can be boxed together. In this way, nodes which satisfy both a distance and fitness criterion can be grouped, and the FD calculated from the end result.

This study is intended as an introductory investigation into the use of fractal measures in the space of local optima in a fitness landscape. We compute the FD and associated metrics for a set of NK Landscape instances. The obtained results suggest links between the fractal geometry in the networks and the empirical difficulty for search algorithms.

2 Background

2.1 The Study of Fitness Landscapes

A fitness landscape [5] is a triplet (S, N, f) where S is the set of all possible solutions, $N : S \rightarrow 2^S$, a neighbourhood structure, is a function that assigns to every $s \in S$ a set of neighbours $N(s)$, and f is a fitness (objective value) function such that $f : S \rightarrow \mathbb{R}$, where the fitness value is a real number that can be viewed as the *height* of a given solution in the landscape.

2.2 The Local Optima Network

A *local optima network* is a representation of the fitness landscape at the level of local optima. We now formally define the constituent parts of a local optima network, before proceeding to describe the object as a whole.

Nodes. The set of nodes, LO , is comprised of local optima, i.e. a solution lo_i satisfies the condition that it has superior fitness to all other solutions in its neighbourhood: $\forall n \in N(lo_i) : f(lo_i) \geq f(n)$, where $N(lo_i)$ is the neighbourhood and n is a single neighbour.

Edges. The set of edges, E , consists of directed and weighted links. An edge is traced if the probability of ‘escape’ — using perturbation and then hill-climbing — from the source node to the destination is greater than zero, and is weighted with the probability. Formally, local optima lo_i and lo_j form the source and destination of an edge iff $w_{ij} > 0$.

Local optima network (LON). The weighted local optima network $LON = (LO, E)$ is a graph where the nodes $lo_i \in LO$ are the local optima, and there exists an edge $e_{ij} \in E$, with weight w_{ij} , between two nodes lo_i and lo_j if $w_{ij} > 0$. Note that w_{ij} may be different than w_{ji} . Thus, two weights are needed in general, and so a local optima network is an oriented transition graph.

2.3 The Fractal Dimension

This study aims to draw links between the fractal detail in local optima networks and search success on the underlying problem instances. Specifically, the *fractal dimension* [3] (FD) is used to characterise the dimensional complexity of the networks. To understand FD we start with the more familiar notion of

typical geometric dimensions; the one-dimensional line, for example, or the two-dimensional square. A line is one-dimensional, as if we try to measure it with a scale half the length of the line (we zoom in by a factor of two), we measure or obtain exactly two copies of the original shape. In general, to calculate dimension for a given object, we need the scaling factor (two in this example), and the extent of detail observed using that scale (here, the two copies). The result is found by solving for x the equation

$$scale^x = detail \tag{1}$$

which in this example is $2^x=2$. The value for x is one, meaning it is a one-dimensional object.

It follows that a shape containing fractal geometry does not have an integer dimension. Instead it has a fractal dimension, which lies somewhere on the real number line. An object with dimension just above an integer (for example 2.12) is only slightly more complex or detailed than the dimension below. Conversely, a dimension just below an integer (e.g. 2.89) indicates a shape with that dimension but with complicated patterns removed. In essence, fractal dimension is a complexity index and captures how detail in an object relates to a scaling factor. Another way to approach this notion is to consider how well a pattern fills the geometric space where it resides. For example, recall Fig. 1b, where the patterns fill much of the possible space. This shape has a high FD; the way it fills space is much more complicated than a shape with a smaller integer dimension. Rearranging Equation 1, we obtain the fractal dimension:

$$FD = \frac{\log(detail)}{\log(scale)} \tag{2}$$

2.4 Fractals and Fitness Landscapes

Weinberger and Stadler [1] noticed self-similar behaviour in certain fitness landscapes. They used the well-known *autocorrelation* metric [6] in their analysis. The way autocorrelation scaled with landscape diameter was indicative, in some cases, of fractal geometry.

Several years later, Locatelli conducted a detailed study [7] on the phenomenon of patterns re-appearing at different levels of abstraction in fitness landscapes. They termed this the ‘multilevel’ structure of optimisation problems, and noted that it could be exploited.

Zelinka *et al.* [8] also demonstrated the potential of using fractal analysis for learning more about the nature of fitness landscapes, focusing on low-dimensional, continuous spaces.

Until now, there has been a lack of study regarding fractal patterns within a local optima network. One consideration is how precisely to define the dimensionality: a complex network is quite different to the typical two-dimensional pictures used in fractal analysis. Methods have been proposed for calculating dimension in the specific case of a network. In this study, a ‘box-counting’ algorithm is used and extended to cater for the local optima network case.

2.5 Fractals and Complex Networks

We use a state-of-the-art ‘box-counting’ methodology to define the FD of a complex network, proposed in Song *et al.* [4]. The process iteratively boxes together nodes which are within m links of each other, with the aim of describing the network in as few boxes as possible. The parameter m is an integral part of computing the FD: it is one of the two values needed for the relation between detail and scale, corresponding to the *scale* parameter introduced in Section 3. For *detail*, we calculate the number of boxes (when the algorithm has converged) as a proportion of total network size, P .

In the specific case of a *local optima* network, link-distance is not the only consideration when boxing together nodes. Crucial information about landscape structure is encoded in the fitness values of the local optima. The process of boxing identifies how much *detail* is observed measuring the object at a certain scale. Omitting the fitness values would ignore fitness detail (for example, a large fitness difference between two optima near each other in the space). Therefore, an extension of the algorithm is required. Two nodes should be boxed together only if they satisfy both a distance *and* a fitness condition. More specifically, two nodes, lo_i and lo_j , can be boxed together in the algorithm if the distance between them is less than m , where m is the maximum links allowed between nodes boxed together, and $|f(lo_i) - f(lo_j)| < \epsilon$, where ϵ is the maximum allowed fitness disparity. In subsequent text, we denote the link distance between two nodes in the network as $d(lo_i, lo_j)$.

Pseudo-code for iterative box-counting of a local optima network is shown in Algorithm 1. The notation $mass(v)$ is used to represent the ‘mass’ of the network which can be covered using the vertex v as a reference point. Upon convergence of the algorithm, we need to examine the relation between detail and scaling to obtain the FD. We have the number of boxes needed, b , which we can take as a proportion of the network as P . For the scaling factor, we use the link-distance parameter m . This corresponds to the level of abstraction being applied to measure the shape. We insert P and m into Equation 2 to derive the FD of the local optima network.

3 Experimental Setting

3.1 Test Problem

We consider instances from a benchmark combinatorial optimisation domain in this work: the NK Landscape model. The instances are from the work by Ochoa *et al.* [2]. The problems are deliberately small in size, such that a full enumeration of the local optima is possible. This is particularly necessary due to the introductory nature of this study into whether fractal analysis of local optima networks is helpful.

NK Landscapes are a family of synthetic fitness functions. They give rise to fitness landscapes which can be tuned from completely smooth to completely random. There are two parameters: N and K . Solutions are binary-encoded and

of length N . The parameter K dictates how many of the binary variables are dependent on each other — epistasis. Each bit has a numeric value assigned from a uniform distribution of floating-point numbers. The fitness of a given solution,

Algorithm 1 Box-counting a Local Optima Network

```

Initialisation:
 $V, CV, NCV =$  nodes in network, center nodes, non-center nodes
 $CV = []$ ,  $NCV = V$ 
 $Cov, NCov =$  covered nodes, non covered nodes
 $NCov = V, Cov = []$ 
Stage 1:
repeat
  for  $v$  in  $NCV$  do
     $mass(v) =$  count( $v'$  in  $NCov$  where  $d(v, v') < mb$  and  $diff(f_v, f_{v'}) < \epsilon$ )
  end for
   $next.center = v$  where  $mass(v) == max(mass(v \in NCV))$ 
  for  $v$  in  $NCov$  do
     $distance = d(next.center, v)$ 
     $e = diff(f(next.center), f(v))$ 
    if  $e < \epsilon$  and  $distance < mB$  then:
       $NCov = NCov - v$ 
    end if
  end for
until  $\forall v \in V, v \in Cov$  or  $v \in CV$ 
Stage 2:
for  $c$  in  $CV$  do
   $id(c) =$  generate(distinct id)
  for  $v$  in  $V$  do
     $lowest =$  lowest combined fitness-distance and link-distance seen so far
     $closest =$  centre node which is closest
    for  $c$  in  $CV$  do
       $distance = d(v, c)$ ,  $e = diff(f_v, f_c)$ 
      if  $distance + e < lowest$  and  $distance < mB$  then:
         $closest = c$ 
         $lowest = distance + e$ 
      end if
    end for
     $closest.centers[v] = lowest$ 
  end for
end for
 $NCov = NCov$  according to  $closest.centers$  values (ascending)
for  $v$  in  $NCV$  do
   $v' =$  neighbour of  $v$  with lower value in  $closest.centers$  array
   $id(v) = id(v')$ 
  remove  $v$  from  $NCV$ 
end for

```

S , is the average of the fitness contributions of the N bits:

$$F(S) = \sum^i f(s_i) \quad (3)$$

when calculating the contribution of a bit, s_i , the values of K other bits are considered:

$$f(s(i)) = f(s_i, s_1^i, \dots, s_k^i) \quad (4)$$

Here we use $N=18$, and $K=\{2,4,6,8\}$, with 30 instances for each K , for a total of 120. NK Landscapes are often used as a test-bed for new fitness landscape techniques because ruggedness can be introduced in a controlled way, by increasing the value for K .

3.2 Metaheuristics

For the fractal analysis to be useful, a view of the relationship between FD and search difficulty in the underlying problem instances should be sought. However, the notion of difficulty is subjective to the algorithm used. We deploy common trajectory-based metaheuristics: iterated local search (ILS), simulated annealing (SA) and tabu search (TS). For the local search element of all three algorithms, a bit-flip is deployed. For the perturbation in the case of the ILS, this mechanism is doubled. The SA parameters are those suggested in Thomson *et al.* [9] in a study on NK Landscapes. The start and end temperatures are 1.4 and 0.0, respectively; α is set at 0.8; and the maximum iterations at the same temperature is 262. The length of the tail in the tabu search is set at n , i.e. 18, the length of the solutions.

All three metaheuristics were implemented with *Paradiseo* [10], an open-source package in C++, and were run 1000 times each per problem instance.

3.3 Fractal Analysis

To calculate the FD of our networks, we employ Algorithm 1: a box-counting algorithm specialised to the local optima network case. An implementation for box-counting a network in C was obtained from the work in Song *et al.* [4].

As we stated in Section 2.5, two important parameters exist in our consideration of FD in the specific case of a local optima network. The parameter m , used in the general box-count algorithm, controls edge-distance between boxed nodes; we also have the parameter ϵ , which is the maximum fitness difference allowed between boxed nodes. This should operate on normalised fitness ranges, such that a single ϵ value can be applied to all networks. A fitness value can be normalised as $f = \frac{f-E(f)}{sd(f)}$, where $E(f)$ is the expected fitness value and $sd(f)$ is the standard deviation. With this, the mean becomes zero while the standard deviation is one.

4 Results

We compute the FD of the local optima networks for each of the 120 problem instances considered. This is done using the modified box-counting algorithm outlined in Section 3.3, and represented in Algorithm 1. For each network, we compute 20 FDs, differentiated by setting the fitness-mandate parameter ϵ , which is in the range $[0.0, 1.0]$ in step sizes of 0.05.

The three metaheuristics described in Section 3.2 are applied to each problem instance. Then, with the fractal information, the feature information and the performance data from the metaheuristics we can proceed to examine the relationships between these measures. The ultimate aim of the experiments in this study was to assess whether fractal analysis of a local optima network could be helpful.

4.1 Fractals and Epistasis

Firstly, we examine the distribution of FD in the local optima networks. The instances can be split by their *epistasis* level, i.e. how many variables are interdependent in a solution. Figure 2 shows box-plots for each of the four levels used.

We can see from the Fig. 2 that the FDs of the local optima networks increase with the epistatic parameter K , which is known to increase ruggedness and randomness in NK Landscapes. We can see this by, for example, comparing the interquartile range in category $K2$ with category $K4$. This would suggest that more ordered, predictable problem instances give rise to lower dimension objects. The interquartile range for the $K=2$ instances spans ~ 1.008 to ~ 1.379 . The median of the $K2$ LON dimension is 1.260. These are objects between one and two-dimensional; somewhere between a line and a 2-d shape. The dimension is closer to one, implying a structure akin to a 2d shape but with lots of parts removed from it, bringing it closer to resembling a line. In our context, the shapes we are dealing with are local optima networks, which represent search connectivity between local optima. A dimension such as this might imply these networks form somewhat linear sequences, with some deviation.

The instances with $K=4$ generally have higher fractional dimension, with the interquartile range spanning ~ 1.363 to ~ 1.878 . Again this implies that the networks are located between one and two dimensions. This time they are closer to two, with a median of 1.623. The fractal in Fig. 3 has dimension ~ 1.585 . We can view it as a 2-d triangle which has complicated portions cut out of it. Of course, the structure of a local optima network with this dimension is unlikely to look like this; however, the way the detail in local optima network scales is similar to this fractal. Having a dimension of around 1.585 implies that when we reduce our scale to measure the shape in Fig. 3 to one fourth (a scaling factor of four), we obtain *nine* copies of our original shape. Using Equation 1, i.e. solving $4^x = 9$ for x , does not result in an integer dimension, but instead ~ 1.585 , a fractional dimension.

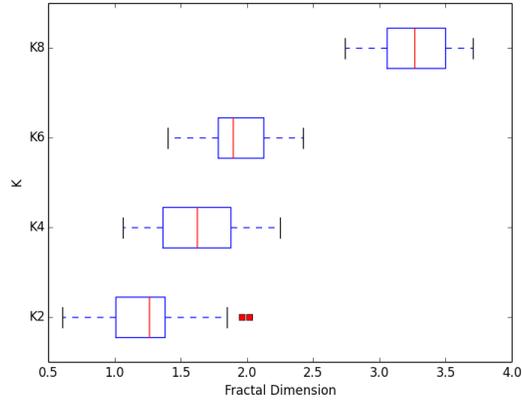


Fig. 2. Boxplots showing the FD of the local optima networks, grouped by epistasis level; 30 instances of each group are considered for each of $K=\{2,4,6,8\}$, all with $N=18$.

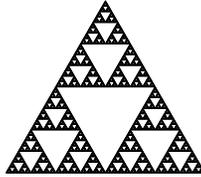


Fig. 3. A shape with dimension between one and two. Sierpinski triangle — FD ~ 1.585 .

Consider again Fig. 2, this time for the $K=6$ instances. Here the interquartile range spans ~ 1.779 to ~ 2.126 , with a median value of ~ 1.893 . This tells us that the networks of local optima are either highly complex in being just below two-dimensional (see Fig. 1b for a fractal like this), or three-dimensional with a lot of the shape absent, leaning towards a 2d shape.

Notice that the increase in dimension is stark between the $K6$ and $K8$ groups. While the majority of dimensions in the lower-epistasis LONs were somewhere between one and two, these highly rugged fitness landscapes appear to give rise to local optima networks with dimension mostly between three and four. Of course, this situation is virtually impossible to visualise or conceptualise.

4.2 Fractal Dimension and Search Performance

We now consider how the FD of local optima networks is connected to empirical search difficulty. To gain insight into this, we must select appropriate measures of hardness.

For search algorithms, two things matter: efficiency and effectiveness. To quantify these, we should have a measure of success (or lack thereof) and a measure of speed. Here we use a proportional measure to represent success: the number of runs which reached the global optimum divided by the total runs. To measure efficiency, we consider the number of function evaluations used in successful runs.

Correlation Study. A logical place to start when contrasting problem features with performance is by correlation study. We compute the Spearman correlation coefficient and corresponding p-value for the observed fractal features and the performance measures.

In the interest of space, Fig. 4 shows the correlation matrix for a selected sample of features of the 120 considered problems. We show the correlation coefficient between variables (upper triangle of plot) density plots (middle diagonal), and scatter-plots (lower triangle). In the case of the density and scatter-plots, difference in colour indicates a split into epistasis level, as outlined in the caption. The size of the text is proportional to the strength of the absolute value of the correlation. An indication of p-value level is given by the asterisk, as described in the caption.

We include all measures of metaheuristic performance in the variable set. There are two metrics for each algorithm, giving six in total; one each for efficiency, and one for effectiveness. Indication of the measure is given after the abbreviation for the algorithm name; for example, *ILS.s* is the success rate of the ILS, while *ILS.t* is the run-time.

The remaining variables included are fitness landscape features. We are particularly interested in those relating to fractal geometry in the local optima networks. In our experiments, we computed correlations between performance and dimension at four different values for ϵ , namely $\{0.25, 0.50, 0.75, 1.00\}$, operating on normalised fitness ranges. The lower the value of this parameter, the more strict the fractal analysis algorithm is when boxing together nodes: recall that they are considered neighbours iff $d(lo_i, lo_j) < d$ and $|f(lo_i) - f(lo_j)| < \epsilon$. If ϵ is nearer to 1.0, the condition for grouped nodes is more lenient.

In the Figure, we show FD correlations with this parameter set at 0.5 (*FD1*) and 1.0 (*FD2*), but 0.25 and 0.75 showed similar trends.

In addition to the fractal dimension relationships, we show those of the number of local optima (*num.optima*) and of the number of landscape funnels (*num.funnels*). A *funnel* is a basin of attraction at the level of local optima and has been linked to search complications in NK Landscape problems [9].

Surveying the correlation matrix, the drastic effect of the parameter ϵ — the local optima network-specific detail parameter — on the resultant FD can be seen as quite apparent when comparing the two FD columns. We can check against the intersections with the performance metric rows. When set at 0.50 (*FD1*), dimension is positively associated with success rate for all three algorithms. In all cases, there appears to be statistical validity. Looking at the runtime rows, we can see that there is a positive connection with SA runtime, but negative ones

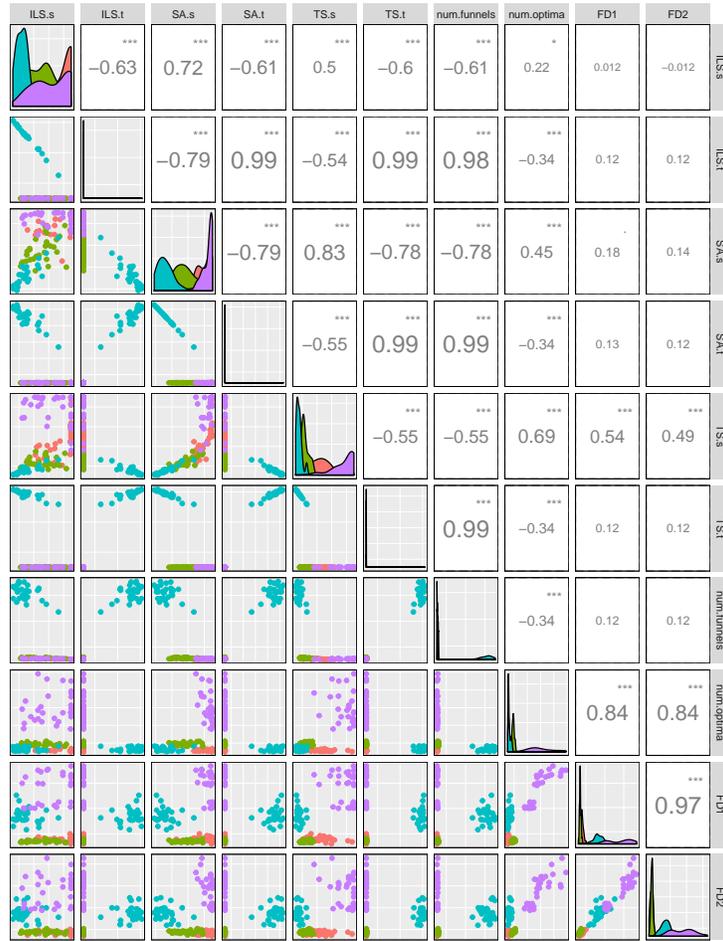


Fig. 4. Correlation matrices of performance metrics and landscape features (see facet titles). Lower triangle: pairwise scatter plots. Diagonal: density plots. Upper triangle: pairwise Spearman's rank correlation, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Colour represents instances split into different levels of epistasis ($K \in \{2, 4, 6, 8\}$).

with ILS and TS. All relationships between this particular FD and performance have ** $p < 0.01$.

If we now consider the other FD, where ϵ is set leniently at 1.0 (column *FD2*), there are quite different results. For the two algorithms based on single-strength mutation (SA; TS), there appears to be a statistically significant connection with success rate. For all algorithms, the correlation with runtime is positive; however, of these, only that of SA is significant, and indeed this is one of the strongest relationships on the matrix.

Regression Models. Using linear models which include the FD as a predictor and a performance metric as the dependent variable allows us to further examine connections with search difficulty. We can normalise the predictors used as $p = \frac{p-E(p)}{sd(p)}$, where $E(p)$ is the expected predictor value and $sd(p)$ is the standard deviation. In this way, we can examine the strengths of contribution from problem features and gain a view of how much variance seen in the response variable is attributable to those we are interested in — the fractal measures.

Table 1. Predictor variables used in the regression models.

Notation	Description
<i>Optima</i>	Number of local optima
<i>Funnels</i>	Number of basins of attraction occurring in the space of local optima
<i>Fractal</i> ¹	FD with fitness difference threshold ϵ set at 0.25 (strictest)
<i>Fractal</i> ²	FD with fitness difference threshold ϵ set at 0.50
<i>Fractal</i> ³	FD with fitness difference threshold ϵ set at 0.75
<i>Fractal</i> ⁴	FD with fitness difference threshold ϵ set at 1.00 (most lenient)

Table 2. Linear mixed models. The dependent variables are respective metaheuristic success rates. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	ILS	SA	TS
Predictor	Estimate	Estimate	Estimate
Fractal Dim ¹	0.605 (0.271)*	0.483 (0.169)**	0.298(0.202)
Fractal Dim ²	-0.313(0.202)	-0.531 (0.126)***	0.575 (0.151)***
Fractal Dim ³	-0.098(0.145)	-0.122(0.090)	0.164(0.108)
Fractal Dim ⁴	0.023(0.083)	0.174 (0.052)***	0.306 (0.062)***
Optima	0.000(0.000)	0.000(0.000)	0.000(0.000)*
Funnels	-0.003(0.001)***	-0.004(0.000)***	-0.004(0.001)***
R^2	0.380	0.709	0.685

Table 2 summarises three linear models: one for the success rate of each metaheuristic considered. The predictors used are introduced in Table 1.

The model summaries include coefficient estimates, standard errors (shown in parentheses), indication of p-value (see caption), and the adjusted R^2 value for the model.

Noteable results are shown in boldened text. Looking at the R^2 values, we can see that the ILS model is considerably weaker than those of the SA and TS.

Table 3. Linear mixed models. The dependent variables are respective metaheuristic run-times. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	ILS	SA	TS
Predictor	Estimate	Estimate	Estimate
Fractal Dim ¹	-0.183(0.481)	-0.281(0.549)	-0.409(0.504)
Fractal Dim ²	-0.228(0.358)	- 1.994 (0.409)***	- 0.854 (0.376)*
Fractal Dim ³	-0.025(0.257)	-0.055(0.293)	-0.133(0.269)
Fractal Dim ⁴	0.402 (0.147)**	1.512 (0.168)***	0.538 (0.154)***
Optima	0.000(0.000)**	-0.001(0.000)***	0.001(0.000)***
Funnels	0.049(0.001)***	0.000(0.001)	0.056(0.001)***
R^2	0.9743	0.918	0.979

Nonetheless, there is one fractal metric which appears to have statistical significance as a predictor in this model: Fractal Dim¹, which is the FD when using the ‘strictest’ of the four values for ϵ . This predictor represents the dimensionality of the network object if box-counting — the method for calculating FD — can only box together nodes which are both sufficiently close in links *and* within 0.25 of each other in (normalised) fitness. In this model, the coefficient is positive; this implies a higher value should be associated with a raised success rate by the ILS.

The SA model, in the middle column, is the strongest of the three, with over 70% of variance in the success rate explained. Here, three of the four FDs have statistical significance. Two are positive: the strictest value for ϵ , and the most lenient. The dimension with ϵ set at 0.5, however, has a negative coefficient — which is larger (in absolute value) than the other two.

If we now look at the TS summary, we can see from the R^2 that this is a very strong model. The fractal coefficients are positive without exception here. Two are noteworthy: the dimension at the most lenient value for ϵ , and the dimension with the value set at 0.5.

Table 3 summarises mixed models similar to those in Table 2, but this time with run-time as the dependent, instead of success rate. Again the predictors are described in Table 1. We can see from the R^2 that these are better models than when considering success rate, with over 90% of variance explained in all three cases.

The ILS model is of particular interest, with an R^2 of over 97%. The coefficients for the fractal predictors show the critical effect of the parameter ϵ . The coefficients are all negative, except for the most lenient value, which is also the only one with statistical significance. This implies a slower search in the case of a LON with high FD if considering mostly link-distance.

If we now look to the middle column in Table 3, we can see that the SA model is of very good fit. A similar trend to the ILS is seen: the fractal coefficients are

negative, except for the one with largest value for ϵ , which is significant, alongside the dimension with it set at 0.5.

The final summary in the Table shows that the TS model is a extremely strong, with $\sim 98\%$ of variance explained. Similarly to ILS and SA, the coefficients for the fractal predictors are negative in the case of three out of four. Again, the only exception is the dimension when using the largest value for ϵ . This predictor has a p-value which indicates significance; as does the dimension when using 0.5, which is larger in absolute value.

An important point in Tables 2 and 3 is that whenever fractal predictors had p-values indicating significance, their coefficients were orders of magnitude larger than those of the number of optima, or the number of funnels, even when these also had acceptable p-values.

5 Discussion

5.1 The Fractal Shape of Local Optima Networks

In Section 4.1, we saw that the instances under study have a varied dimension in their local optima networks.

The local optima networks which were extracted from instances with the lowest level of epistasis (and therefore ruggedness) had a FD interquartile range of ~ 1.008 to ~ 1.379 . Clearly, most of these local optima networks are between one and two dimensional in the way they fill space. Having a dimension just above one implies a linear sort of structure, but with some extra pattern, such that it cannot be classified as a one-dimensional line. An interpretation of this for local optima networks is that the networks comprise a somewhat linear sequence of local optima, with a little bit of deviation or detail.

The networks from the $K=4$ and 6 problems generally had higher dimensions. Some of these were complex shapes between one and two dimensions (for example, $FD=1.89$), implying a line with a substantial amount of extra detail. This could mean that the networks are winding convoluted sequences, with spokes leading off the main path. Some networks from these categories had dimensions which were just above two. A fractional dimension just above an integer signifies a low-complexity object. For our local optima networks, this could possibly mean a low number of paths of local optima, with a bit of detail over and above 2d scaling.

The local optima networks which were extracted from the most rugged problems ($K=8$) are virtually impossible to properly envision, as the majority of them had a FD greater than three. This being said, most were just above three, indicating the detail in the objects scales mostly in the manner of a three-dimensional object. It is possible that the three dimensions represent connecting or intersecting paths of local optima.

5.2 Connections with Search Difficulty

In Section 4.2 we saw that there are connections between dimension in a local optima network and search difficulty. When using $\epsilon=1.0$, the calculated dimen-

sions were linked to a raised success rate and a prolonged search time in metaheuristics. Both phenomena could be explained if we recall that many of the easier-instance LONs had dimension between one or two; in other words, the dimension implies a linear, sequence-like structure, with some degree of extra detail or pattern. The labyrinthine nature of the detail would explain the speed, and the one-dimensional base could mean that search algorithms on the underlying instance manage to get through the sequence of local optima to the global best.

Changing the value of the parameter ϵ changes the way the FDs are calculated and the results are markedly different. We saw in Table 3 that when values $\{0.25, 0.50, 0.75\}$ were used during the calculations, instead of 1.00, the corresponding network dimensions were associated with a quicker runtime in search algorithms. This is the opposite result to that obtained using $\epsilon=1.00$. The difference lies in the boxing of nodes in the box-counting algorithm. With stricter criteria for nodes being boxed together, the network is partitioned in a different way. A high fractional dimension, therefore, might be associated with ‘nearby’ nodes (in edge-distance) not satisfying the fitness condition. Nearby nodes such as these might form paths with big fitness ‘jumps’ which might explain the quicker behaviour of search algorithms.

6 Conclusions and Future Work

We have conducted an empirical and introductory study on the potential for use of fractal analysis in local optima networks. A benchmark combinatorial problem, the NK Landscape Model, was used as a case study. Various extents of ruggedness were used.

The results indicate that when we consider our local optima networks to be general complex networks, a high FD relates to increased run-time for prominent metaheuristics, but also an enhanced rate of hitting the optimum. Accordingly, fractal analysis seems to capture a unique phenomenon in a local optima network: previously-proposed metrics have been linked to lowered performance only.

Another important result is the critical effect of the parameter specific to the case of a network composed of local optima, denoted here as ϵ . In particular, networks with high dimension when considering *fitness* as part of the scale used, were shown to be associated with increased efficiency in ILS and TS. We argue that this may be due to ‘close’ local optima being associated with large fitness jumps.

While the focus was on small problems here, there is no reason that fractal analysis could not be deployed on sampled local optima networks, given a robust sampling algorithm. With this new insight into the ‘middle level’ in a search space — the space of local optima — we can proceed further down this avenue of possibility. The more traditional (agnostic of the semantics of the network) fractal analysis captures a unique element of the local optima network: a phenomenon which is linked to lowered efficiency but raised effectiveness, and which we argue merits further investigation. Furthermore, the addition of the fitness-

mandate condition for the special case of a local optima network gives valuable insight into the importance of optima connectivity and fitness distribution.

Acknowledgements. This work is supported by the Leverhulme Trust (award number RPG-2015-395) and by the UK's Engineering and Physical Sciences Research Council (grant number EP/J017515/1). We gratefully acknowledge that all network data used during this research were obtained from [2].

References

1. Weinberger, E.D., Stadler, P.F.: Why Some Fitness Landscapes are Fractal. *Journal of Theoretical Biology* 163(2), 255–275 (1993)
2. Ochoa, G., Tomassini, M., Vérel, S., Darabos, C.: A Study of NK Landscapes' Basins and Local Optima Networks. In: Proceedings of the 10th annual conference on Genetic and evolutionary computation. pp. 555–562. ACM (2008)
3. Mandelbrot, B.: How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science* 156(3775), 636–638 (1967), <http://www.sciencemag.org/content/156/3775/636.abstract>
4. Song, C., Gallos, L.K., Havlin, S., Makse, H.A.: How to Calculate the Fractal Dimension of a Complex Network: the Box Covering Algorithm. *Journal of Statistical Mechanics: Theory and Experiment* 2007(03), P03006 (2007)
5. Stadler, P.F.: Fitness landscapes. In: *Biological Evolution and Statistical Physics*, pp. 183–204. Springer (2002)
6. Weinberger, E.: Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference. *Biological cybernetics* 63(5), 325–336 (1990)
7. Locatelli, M.: On the Multilevel Structure of Global Optimization Problems. *Computational Optimization and Applications* 30(1), 5–22 (2005)
8. Zelinka, I., Zmeskal, O., Saloun, P.: Fractal Analysis of Fitness Landscapes. In: *Recent Advances in the Theory and Application of Fitness Landscapes*, pp. 427–456. Springer (2014)
9. Thomson, S.L., Daolio, F., Ochoa, G.: Comparing Communities of Optima with Funnels in Combinatorial Fitness Landscapes. In: Proceedings of the Genetic and Evolutionary Computation Conference. pp. 377–384. GECCO '17, ACM, New York, NY, USA (2017), <http://doi.acm.org/10.1145/3071178.3071211>
10. Cahon, S., Melab, N., Talbi, E.G.: Paradiseo: A Framework for the Reusable Design of Parallel and Distributed Metaheuristics. *Journal of heuristics* 10(3), 357–380 (2004)