Extrema Graphs: Fitness Landscape Analysis to the Extreme!

Sophie Sadler
Swansea University
Swansea, United Kingdom

Alma Rahat
Swansea University
Swansea, United Kingdom

David J. Walker
University of Plymouth
Plymouth, United Kingdom

Daniel Archambault
Swansea University
Swansea, United Kingdom

ABSTRACT

Fitness landscape analysis often relies on visual tools to provide insight to a search space, allowing for reasoning before optimisation. Currently, the dominant approach for visualisation is the local optima network, where the local structure around a potential global optimum is visualised using a network with the nodes as local minima and the edges as transitions between those minima through an optimiser. In this paper, we present an approach based on extrema graphs, originally used for isosurface extraction in volume visualisation, where transitions are captured between both maxima and minima embedded in two dimensions through dimensionality reduction techniques (multidimensional scaling in our prototype). These diagrams enable evolutionary computation practitioners to understand the entire search space by incorporating global information describing the spatial relationships between extrema. We demonstrate the approach on a number of continuous benchmark problems from the literature and highlight that the resulting visualisations enable the observation of known problem features, leading to the conclusion that extrema graphs are a suitable tool for extracting global information about problem landscapes.

CCS CONCEPTS

• **Human-centered computing → Information visualization**.

ACM Reference Format:
Sophie Sadler, Alma Rahat, David J. Walker, and Daniel Archambault. 2023. Extrema Graphs: Fitness Landscape Analysis to the Extreme!. In *Genetic and Evolutionary Computation Conference Companion (GECCO '23 Companion), July 15–19, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3583133.3596343

1 INTRODUCTION

Recent years have seen substantial growth in *fitness landscape analysis* (FLA) studies. The fitness landscape is a functional mapping between any given solution to an optimisation problem and its corresponding objective value. The fitness landscape can be thought of as a topological representation of each region of a problem's search space, and its analysis yields important information about the likely performance of a search algorithm when applied to it. Recent work has offered a plethora of techniques for exploring a problem's fitness landscape, but without loss of generality we denote by $\mathbf{x}_i$ the i-th sample from the search space, and given an objective function $f(\cdot)$, which is to be either minimised or maximised, we describe the corresponding fitness of $\mathbf{x}_i$ with $y_i = f(\mathbf{x}_i)$.

The features of a problem landscape have considerable impact on the likely performance of a metaheuristic when applied to optimise the problem. For example, it is well known that search algorithms can be trapped in local optima - regions of the search space that are optimal within the local area, and are thus difficult to escape from, but are not the global optimum that is sought by the algorithm. More complex characteristics such as *funnels*, disparate clusters of local optima, and flat regions within the search space can also prove problematic, and therefore identifying their presence can inform the choice of a suitable algorithm and operators.

A range of techniques have been proposed for FLA (see [15] for a recent review), and a prominent approach is the *local optima network* (LON), specifically for *visual analysis*. LONs were first proposed for the analysis of NK landscapes [18], and have since gone on to be used in a wide range of combinatorial settings [19], as well in the continuous [1] and multi-objective [5] domains. The principle behind a LON is to represent the search landscape as a directed graph, with nodes representing local optima and edges representing possible transitions between the optima that might be taken by a search algorithm. A key concept in the construction of a LON is that of *neighbourhood*, and a number of approaches to this have been taken depending on the specific representation at hand. A key motivation behind constructing a LON is so that it can be visualised. Visualisation is done using a technique such as a force-directed layout [11], and provides an intuitive way of understanding the landscape of an arbitrary problem. Providing visual information behind the structure of a fitness landscape is a highly intuitive approach to informing a problem owner or algorithm operator about the type of problem they are dealing with.

In this work, we aim to develop a novel *visualisation approach to capture the general characteristics of function landscapes with input dimensions of three or higher (similar to contour plots for two-dimensional input spaces).* We expect that this will assist the design process of problem-specific optimisation algorithms. To this end, we propose the use of *extrema graphs* for FLA. These have some key similarities with LONs, providing a graphical representation of a search space that can then be visualised, but include all extrema (minima and maxima), to provide a more complete characterisation of the search space. Regions in which the distance between the minima and maxima is small can be identified, such that the ensuing small headroom makes passing between them difficult for

GECCO '23 Companion, July 15–19, 2023, Lisbon, Portugal

ACM ISBN 979-8-4007-0120-7/23/07.
https://doi.org/10.1145/3583133.3596343

solutions. In visualising the resulting network by projecting it into two dimensions with multidimensional scaling (MDS) we also preserve the distances between extrema, a characteristic which is lost in other layouts such as force directed graphs.

The remainder of the paper is structured as follows. Section 2 describes related LON and visualisation research, before extrema graphs are introduced in detail in Section 3. Section 4 analyses the benefits of extrema graphs on a set of benchmark visualisations, while Section 5 provides further evaluation in the form of an expert interview. Discussion (including that of future work) and concluding remarks are made in Section 6.

2 BACKGROUND

Visual Landscape Analysis Fitness landscape analysis has become a considerable topic of research within evolutionary computation. In this section, we review some of the relevant background material. We note that, given the focus of this paper on a novel visual technique for FLA, we restrict ourselves to a description of techniques for visualising fitness landscapes.

One of the principal techniques for visualisation within FLA is the *local optima network* [18]. LONs are designed to illustrate the structure of a landscape as a graph $G = (V, E)$ such that V is a set of nodes representing basins of attraction (local optima) connected by edges $e_{ij} \in E$ wherein a node v_i is connected to v_j if they are deemed to be neighbours. In the original work by Ochoa et al. (2008) [18], which dealt with binary NK landscapes, a neighbourhood was defined to be any two solutions whereby one could transform to the other by means of a single bit flip. Since that work, LONs have been extended to a wide range of solution representations and problem types (including multi-objective [5]). In the case of continuous problems, which are the focus of this work, a variant of the discrete LON formulation was proposed in which the distance between two nodes was used to identify neighbours [1]. Figure 1 illustrates a LON constructed for a 5D instance of the Rastrigin problem. Concepts from LONs have since been extended to formulate *search trajectory networks* (STNs) [17]. Since mapping the search process of an evolutionary algorithm through the space is beyond the scope of this work, we do not consider STNs further.

Visualisation Inspiration for our approach comes from existing literature in the visualisation community on extrema graphs [7, 8]. Extrema graphs, used in volume visualisation, provide a structure to extract isosurfaces automatically from volumes. Two-dimensional surfaces are defined in three dimensions, similar to the to one-dimensional contour lines on a map, that separate areas of higher density from areas of lower density in the volume. In isosurface extraction, the graphs are usually embedded in three dimensions, but that is not always the case for visual landscape analysis. Therefore, for the purposes of landscape analysis, we visualise the calculated extrema graph directly through dimensionality reduction of the maxima/minima along with samples along the edges between them.

Dimensionality reduction takes data in high dimensional spaces and maps them down to lower dimensional spaces which for visualisation is often two dimensions. Multidimensional scaling approaches [2, 3, 14, 20, 22] are one type of these methods which optimises distances between points in the low dimensional space so that they are representative of distances in the high dimensional space. In our approach, we use multidimensional scaling to embed the extrema and the sample points along the edges (those which make up the "edge nodes") for visualisation in 2D. MDS has previously been used in other approaches to embed search spaces for visualisation, such as in the work by Michalak [16]. However, our work differs in the use of a graphical representation. For this reason, we position our method as a complementary approach to LONs rather than comparing it directly to other less similar approaches.

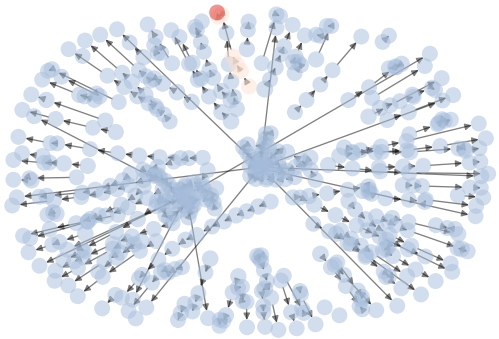

Figure 1: An example of a LON for the Rastrigin function with 5-dimensional decision space using the default settings from Adair et al. [1]. The dark red node represents the global minimum. The series of red nodes show the basin of attraction for the global minimum, while other blue nodes represent basins for local minima. Arrows are also included showing the direction of travel in the basin-hopping algorithm.

While in this work we use MDS to visualise the solution space, it has also been used to visualise high-dimensional spaces in many-objective optimisation, presenting trade-off surfaces for problems comprising four or more conflicting objectives [24, 25]. Later work considered the visualisation of a multi-objective optimiser's route through the search space in a way that enabled landscape characteristics to be inferred from optimiser behaviour.

3 EXTREMA GRAPHS

Similarly to a LON, extrema graphs use a graphical representation where nodes consist of extrema (in the case of LONs, only the minima) and edges represent some metric of distance between these extrema. However, extrema graphs differ from LONs in both the construction and the visualisation of this graph, for example, in the inclusion of both maxima and minima. The incorporation of all extrema, however, is infeasible for large landscapes with many extrema due to the computational expense in identifying them. To generate our extrema graphs, we have therefore used the *Niching Migratory Multi-Swarm Optimiser* (NMMSO) [4] to sample extrema to include in the graph. NMMSO is a good multimodal optimiser capable of finding many, if not all, of the extrema in our benchmark problems, however could be substituted for another approach. In this work, we focus on developing a novel visualisation technique, rather than optimising the process for identifying extrema.

3.1 Graph Construction

The graph is then constructed from these extrema in the following steps, where, at the first stage, each node in the graph corresponds to one of the extrema. A full overview of the graph construction process is contained in Algorithm 1.

Firstly, an edge is added between two nodes where the Euclidean distance between the corresponding extrema is less than a threshold. The threshold, t is given by a hyperparameter ρ multiplied by the Euclidean distance between the bounds of the search space, such that $t = \rho\|\mathbf{b}_u - \mathbf{b}_l\|$ where $\mathbf{b}_u$ and $\mathbf{b}_l$ vectors are the upper and lower bounds of the search space (line 4 of Algorithm 1). We call the hyperparameter ρ the *radius percentage*. This hyperparameter can take values between 0 and 1, thus determining the proportion of edges which are included, and can be chosen by the problem owner based on their requirements when designing an optimiser (for example, if it depends on a certain neighbourhood value) as well as their own personal perception of the visualisation.

Once the edges have been determined, they are then replaced with a sequence of nodes, which we call *edge nodes*. In the visualisation there are therefore no edges in the graph, however the original extrema nodes have a much larger visual representation than the new edge nodes, such that these take the appearance of edges between the extrema nodes. The benefits of drawing the edges as a sequence of nodes are to vary the colour along the edge to visualise changes in fitness value, as will be described in the subsequent section 3.2, as well as to visually indicate how the search space is folded by the dimensionality reduction. These benefits offset the downside of additional visual clutter. Each of these edge nodes also has an associated location in the search landscape, such that they are linearly spaced along the line connecting the two relevant extrema. The associated locations in the search landscape of both the extrema nodes and the new edge nodes are stored in a location matrix, L, as described in lines 5 and 10 of Algorithm 1.

This process produces the final graph topology, where we have a number of extrema nodes and a number of edge nodes, each having a corresponding location in the search landscape. The final step before we can visualise the extrema graph is to translate these locations in high-dimensional space to locations in 2-dimensional space. In our extrema graph prototype, we use MDS [14] to perform this translation as it aims to preserve relative distances, though there are many other relevant dimensionality-reduction techniques which could be applied, including Landmark MDS [20] or t-SNE [22], which has previously been applied for visualising LONs [23]. The chosen dimensionality-reduction method should be applied to all nodes - both extrema nodes, and edge nodes, to determine their location in the extrema graph visualisation. It should be noted that, while MDS aims to preserve pairwise distances when projecting to lower-dimensional space, it is impossible to do so perfectly and therefore there may be some unintuitive distortion of the space. This motivates the use of edge nodes to aid understanding.

3.2 Visual Encoding

The MDS drawing of the extrema graph is visualised directly. However, in order to distinguish minima, maxima, and edge nodes, as well as to provide other useful information about the search landscape topology, we use additional visual channels in our encoding.

Firstly, extrema nodes are sized significantly larger than edge nodes to distinguish them as local extrema. All nodes are then coloured according to an isoluminant colour scheme representing the fitness value at the corresponding location in the search landscape. In our prototype, we use Matplotlib's Viridis colour scheme, which varies from yellow for the highest values through green, blue and finally purple to the lowest values. Thus minima are usually represented in purple and maxima in yellow. Edge nodes are also coloured according to this scheme, allowing for interpretation of the shape of the landscape between two extrema. The global minima are marked separately in red to distinguish them from other minima, which may be globally non-optimal despite appearing similar in colour when using the Viridis colour scheme.

None of the visual features described here are present in LONs, and are therefore a key differentiator of the two methods.

Algorithm 1 Extrema Graph Construction[1]

Input: For a problem with dimensionality m: function $f : \mathbb{R}^m \rightarrow \mathbb{R}$; upper and lower bounds on the search region, $\mathbf{b}_u, \mathbf{b}_l \in \mathbb{R}^m$; number of fitness evaluations, $2N$; radius percentage, ρ; number of edge nodes per edge, e

Output: Extrema graph and 2D node locations

1: Run NMMSO algorithm for N fitness evaluations to obtain n_{max} maxima
2: Run NMMSO algorithm for N fitness evaluations to obtain n_{min} minima
3: Initialise graph $G = (V, E)$ with $|V| = n_{max} + n_{min}$ and $E = \emptyset$
4: Set the threshold, $t = \rho\|\mathbf{b}_u - \mathbf{b}_l\|$
5: Initialise location matrix $L \in \mathbb{R}^{n \times |V|}$ such that for every node $i \in V$, $L_{\_,i} = o_i$ where o_i is the location of optimum i in the landscape
6: **for** pairs of extrema, o_i, o_j **do**
7: If $\|o_i - o_j\| < t$, add edge (i, j) to E
8: **end for**
9: **for** edges (i, j) in G **do**
10: Add e nodes to V and corresponding locations in the search landscape to L such that these locations are evenly spaced along the line connecting o_i and o_j
11: Remove (i, j) from E
12: **end for**
13: Perform MDS on L to reduce to 2 dimensions
14: Output $G = (V, \emptyset)$ and L
15: **Note:** $\|\cdot\|$ denotes the Euclidean norm in n dimensions

4 EXPERIMENTAL RESULTS AND ILLUSTRATIONS

In order to demonstrate the efficacy of extrema graphs for understanding search landscapes, we demonstrate them as applied to a number of well-known, continuous benchmark functions: Sphere, Rastrigin, Schwefel, Ackley, Griewank and Rosenbrock. Formulations of these functions can be found in Appendix A, while the 2D surface plots are shown in Figures 2 and 3 below.

[1] An implementation of the algorithm in Python is available at https://github.com/sophiefsadler/extrema_graphs.

We have produced an extrema graph for each function in 2, 3 and 5 dimensions to demonstrate the performance and variation across these differing numbers of dimensions. Additionally, the extrema graph for the Sphere function in 10 dimensions can be found in Appendix B. In the graph construction, we used 20, 000, 30, 000 and 50, 000 fitness evaluations respectively to generate extrema graphs in 2, 3 and 5 dimensions. The radius percentage hyperparameter, ρ was chosen for each function individually; these values are as in table 1. As this value controls the number of edges included in the extrema graph, using a larger value of this hyperparameter allows more information about the fitness landscape to be captured and included. However, for functions with large numbers of maxima and minima, a large radius percentage can lead to an extrema graph with too much visual clutter to be understood, and also drastically increases computation time. For our evaluation, we chose this value experimentally, by testing a range of values and selecting the one for which the resulting extrema graph contained the best quantity of valuable information in the eyes of our interpretation. These values were then fixed across all dimensions so that all extrema graphs for a given function could be directly compared. Note that the NMMSO runs need not be rerun in order to test new values of ρ, so these were performed only once.

Radius Percentage Values		
	ρ	Search Bounds
Sphere	0.75	$[-5.12, 5.12]^D$
Rastrigin	0.08	$[-5.12, 5.12]^D$
Schwefel	0.1	$[-500, 500]^D$
Ackley	0.05	$[-32.768, 32.768]^D$
Griewank	0.25	$[-5, 5]^D$
Rosenbrock	0.75	$[-5, 10]^D$

Table 1: The value of the "radius percentage" hyperparameter ($\rho \in (0, 1]$) used during graph construction, alongside the bounds of the search region explored.

4.1 Benchmark Visualisations and Interpretation

The first set of extrema graphs is shown in Figure 2 where we visualise 2D, 3D and 5D landscapes for Sphere, Rastrigin and Schwefel. The top row shows a conventional surface plot of a 2D fitness landscape for each, while the second, third and fourth row show the 2D, 3D and 5D extrema graph for the corresponding problem.

4.1.1 Sphere. Considering the first column, we see the results for the Sphere problem – the problem with the simplest fitness landscape considered herein. Figure 2(d) shows an extrema graph of a 2D landscape. The graph consists of five extrema nodes: four nodes shown in yellow, with the central node being shown in red. As described in section 3.2, the red node corresponds to the global minimum, while yellow nodes have high fitness value (so are maxima). The larger extrema nodes are connected by smaller edge nodes, also coloured according to fitness. In the case of the Sphere function, this simple visualisation provides substantially more information than would be visible in a LON representing the same landscape. In that case, the graph would comprise of a single node (representing the global minimum) with no edges. Here it is possible to observe the maxima, which correspond to the corners of the landscape shown in Figure 2(a); the scale of the fitness in those nodes is determined by the bounding box selected when instantiating the problem. The colour of the internal edges, connecting the maxima to the minimum, show a steady decrease in fitness (the colour gradually progresses from yellow to purple). This represents the smooth slope of the landscape, which has no local minima. The outer edges decrease to a trough before climbing back towards the adjoining maxima. This is a graphical representation of how the surface cuts the landscape's bounding box, which is the U-shape that can be seen at the edge of the landscape in Figure 2(a).

Figure 2(g) shows the corresponding 3D Sphere landscape. The characteristics highlighted in the 2D case can be seen here. In this case, there are eight nodes representing maxima. This is to be expected, as any instance of this problem will feature a hypercube comprising 2^D corners, which form the maxima. Each maximum is again directly connected to the minimum via an edge, which again exhibits the steady decrease we would expect from the smooth landscape. Each adjacent maximum is again connected by an edge that shows the U-shape formed by the bounding box, however in this case there are additional edge nodes. This is an artefact of the compression of a 3D landscape into two dimensions, and shows connections between maxima in the third dimension. The same topology is visible in the 5D graph shown in Figure 2(j). There are again $2^5 = 32$ nodes representing maxima, though the number of edge connections in this case makes it difficult to observe direct connections between them. It is, however, still possible to identify the smooth progression from maxima to minimum by following the colour of the edges as they move toward the centre of the graph.

4.1.2 Rastrigin. The second column of Figure 2 illustrates the Rastrigin examples. Again, the top row shows a the conventional landscape view and the second row (Figure 2(e)) shows the corresponding 2D extrema graph. In some ways, the visualisation is not dissimilar to the Sphere graph shown in Figure 2(d). There are four maxima, again shown in yellow, though we note that the scale is different as it depends on the fitness values in a specific graph. The principal difference is the presence of a number of local minima, represented as nodes within the graph. This is an intuitive way of viewing the graph, since the Rastrigin objective function is formed of the sum of squared decision variables (accounting for the similarities to the Sphere case) which is offset by a cosine term introducing regular local minima (accounting for the additional minima nodes). The regularity of the minima can be seen by the way in which the nodes have been placed into a grid arrangement. Clearly, the graph is missing regions of the landscape; this is because of the stochastic nature of the multimodal optimiser that has been used to sample the space. One option that could be used to address this and provide a more complete map would be to run multiple repeats of the optimiser in the way that a LON is constructed, however, this would increase the computational cost which is currently minimised by the use of only two optimiser runs, one for the maxima and one for the minima. An interesting characteristic visible from the edge nodes is that the problem is non-convex. This is shown by the regular progression of colour between nodes.

The 3D case shown in Figure 2(h) exhibits similar characteristics to the 2D case, albeit in a less uniform layout because of the compression of the additional dimension. The minimum is still shown at the centre of the graph, however the graph shows fewer connections between the inner core and the outer maxima. This again is likely to be an artefact of the way in which the multimodal optimiser has traversed the space while sampling. This effect is even more pronounced in the 5D case (Figure 2(k)).

4.1.3 Schwefel. The Schwefel problem is similar in construction to Rastrigin. Figure 2(f) indicates this in that the graph has a similar regular structure. Noise in this case is introduced by a sine term rather than the cosine, and taking the square root of the absolute value of each decision variable introduces local minima at varying heights, rather than the steadily decreasing minima of Rastrigin. In Figure 2(e), the 2D Rastrigin extrema graph, the further a minimum is toward the global minimum the lower the fitness value. In Figure 2(f), for Schwefel, the minima increase and decrease as they progress to the global minimum. Another interesting aspect of this problem compared to Sphere and Rastrigin is that it only has one maximum. This is placed at the opposite end of the graph from the minimum, which is again shown in red. While these characteristics are, as has been the case in the other problems, visible in the 3D case Figure 2(i), they are more difficult to observe in the 5D case Figure 2(l). We theorise that this is because of the effect of the more rugged landscape on the sampling process, leading to fewer connected edges which in turn leads to a less precise embedding with MDS.

4.1.4 Ackley. The left-hand column of Figure 3 illustrates extrema graphs for the Ackley problem. Again, Figure 3(a) shows the 2D landscape in a surface plot, and 3(d) shows the corresponding extrema graph. This graph follows the typical layout we have seen in previous cases, with maxima arranged around a central minimum (again, shown in red). The regions between the external maxima and the minimum are sparse. This, in combination with the high concentration of minima identified near the global minimum, indicates that the optimiser has not struggled to locate the problem's central funnel. Those maxima and local minima that have been found are arranged in a somewhat regular fashion, which is as would be expected for the problem. When considering the 3D case (Figure 3(g)) the number of local minima nodes found near the global minimum has reduced considerably. This indicates that the funnel is becoming steeper, and the optimiser is less likely to be trapped by local minima once it locates solutions within the funnel. Having increased to 5D in Figure 3(j), the distance between the funnel and the maxima nodes is increased. This is due to the MDS compression, which has placed nodes towards the corners of the projection as it seeks to keep nodes that are distant in the 5D space distant in the corresponding 2D embedding.

4.1.5 Griewank. As can be seen in Figure 3(b), the Griewank landscape is highly regular. This regularity has resulted in a distinct extrema graph (Figure 3(e)). As seen previously, the global minimum is placed toward the centre of the graph. The local minima and maxima are placed around it, organised into lines. These lines cut across the fitness landscape, and the peaks and troughs visible in the 2D landscape (Figure 3(a)). This effect is visible in the 3D case, shown in Figure 3(h), and the 5D case, shown in Figure 3(k). In both cases, compressing the extra dimensions into a two-dimensional extrema graph has resulted in curved edge nodes. While in the 2D and 3D cases all of the maxima and local minima have been located, the optimiser has not sampled the entire 5D landscape.

4.1.6 Rosenbrock. The final problem we consider is Rosenbrock. As shown in Figure 3(c), this problem is characterised by a deep trench of reasonable fitness, only a specific portion of which is actually optimal. Figure 3(f) illustrates the 2D extrema graph. Here, there is a single maxima, due to a slight tilt in the fitness landscape that biases the value of one of the maxima to be marginally higher than the other corners (however, by expanding the bounds of the problem the opposite side of the landscape in the x_2 axis becomes visible). In the lower right-hand corner of the graph are a set of near-optimal local minima connected to each other, as well as the global minimum. This indicates the trench, and this characteristic can be seen in Figures 3(i) for the 3D case and 3(l) for the 5D case. In both the 3D and 5D cases the optimiser has located maxima in other corners, and in both cases these are connected to the trench (and, in some cases, the global minimum itself) by decreasing edge nodes as was the case in the Sphere example.

5 EXPERT INTERVIEW

We also conducted an interview with an FLA expert to ascertain whether extrema graphs as we have presented them can be easily understood, and whether they fulfil the intended purpose. Expert interviews are a common assessment of visualisation techniques, usually consisting of studies with a small number of experts[2]. In this case, we perform such an evaluation with one expert and will expand this evaluation in future work.

Experimental Design. The interview took the following format. We began by introducing the premise of our approach and the aims of our research. This was followed by some technical discussion of the methodology, and finally by provision of examples in the form of the Sphere function extrema graphs in 2D, 3D, 5D and 10D.

In the second stage, we presented randomly ordered extrema graphs and randomly ordered surface plots for the remaining benchmark functions in 2D (Ackley, Griewank, Rastrigin, Rosenbrock and Schwefel), and asked the expert to match each extrema plot with one of the surface plots. At this stage, we did *not* tell the expert whether the answers were correct, but immediately presented randomly ordered extrema graphs for the benchmark functions in 3D and asked the expert to again match these to the surface plots. Although the expert was *not* informed of the correct answers for the 2D plots, we cannot rule out that a learning effect could take place to improve the expert's performance on the subsequent 3D plots. This approach was chosen to assess the expert's understanding as it gives us a quantitative evaluation of performance, in contrast to other qualitative elements of the expert interview. The interview was concluded with an exploration of the benefits and drawbacks of our approach as perceived by the expert.

Participant Response. For the 2D functions, the expert correctly matched the Ackley, Rastrigin and Schwefel graphs to their surface plots without difficulty, but deliberated over the matching of the

[2]Five or fewer experts can find the majority of issues with a visualisation [21]. In line with standard ethical practice, anonymity of participants/experts is protected.

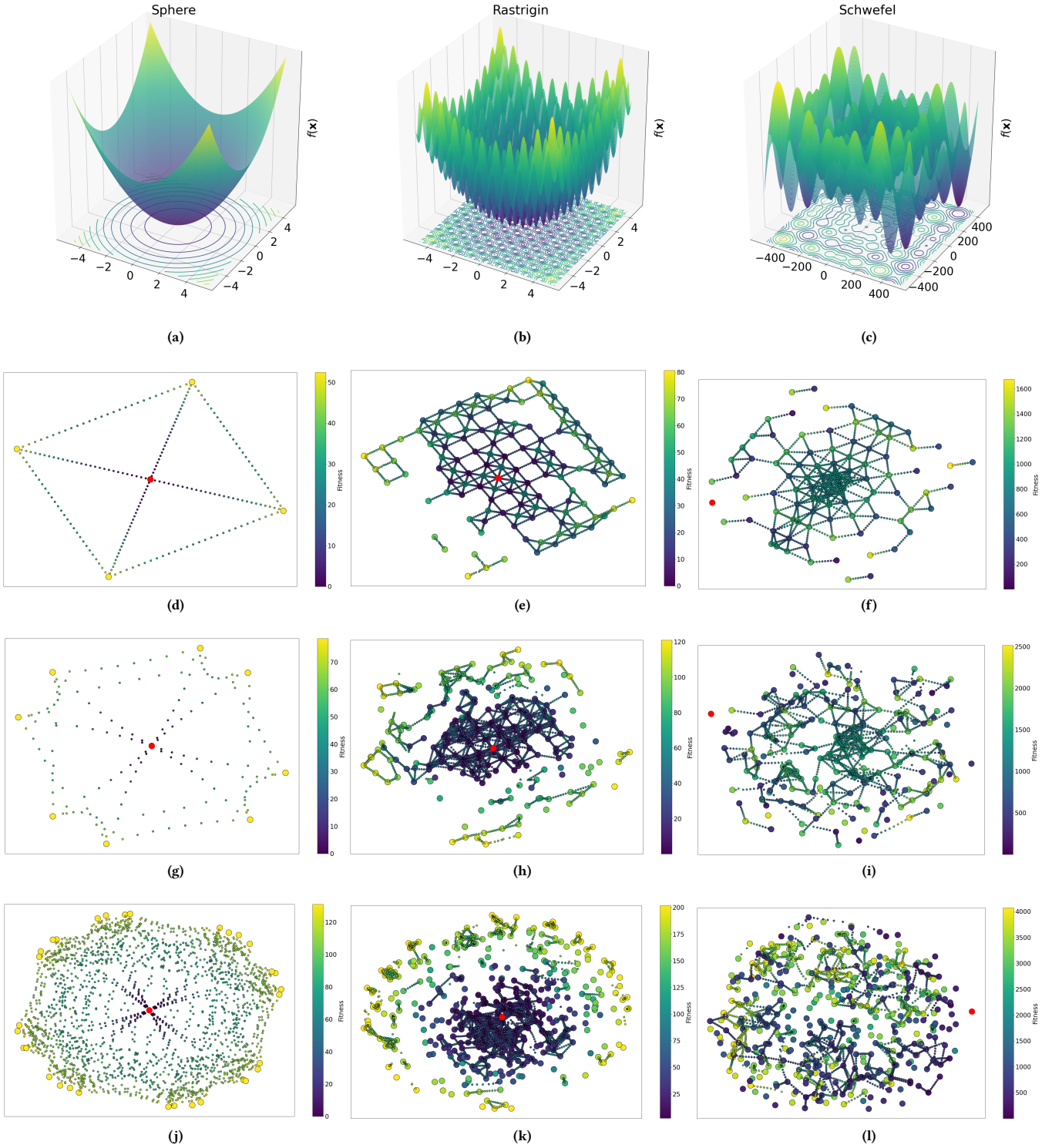

Figure 2: Surface plots for the Sphere (a), Rastrigin (b) and Schwefel (c) functions. Extrema graphs for the Sphere function in 2D, 3D and 5D are shown in (d), (g) and (j); for the Rastrigin function in 2D, 3D and 5D in (e), (h) and (k); and for the Schwefel function in 2D, 3D and 5D in (f), (i), and (l).

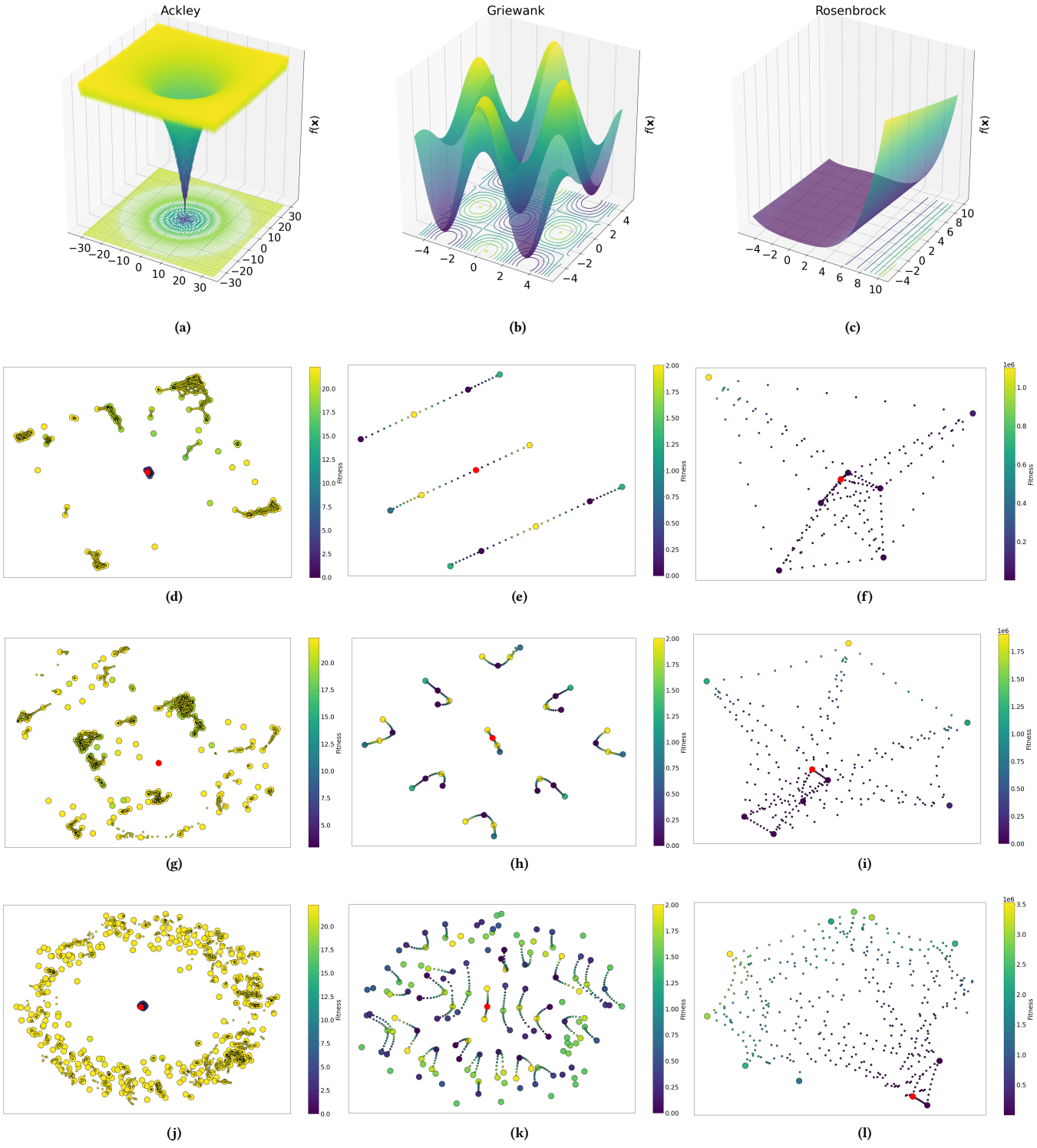

Figure 3: Surface plots for the Ackley (a), Griewank (b) and Rosenbrock (c) functions. Extrema graphs for the Ackley function in 2D, 3D and 5D are shown in (d), (g) and (j); for the Griewank function in 2D, 3D and 5D in (e), (h) and (k); and for the Rosenbrock function in 2D, 3D and 5D in (f), (i), and (l).

Griewank and Rosenbrock functions and ultimately matched these incorrectly. For the 3D functions, the expert once again deliberated over the matching of the Griewank and Rosenbrock functions, but this time got the matching of all 5 functions correct. This indicates that the expert was able to gain enough insight from the extrema graphs to identify the corresponding functions. Increased exposure may have led to the improved performance in the second round.

During the discussion in the final stage of the interview, the expert identified the ability of extrema graphs to capture the ruggedness of the landscape, which is derived from the concise representation of the spatial relationship of the extrema, as their main benefit. These approaches can provide an overview of the functions in higher dimensions, in particular with respect to their function variability. Furthermore, the expert highlighted that this method may be useful in visualising paths of optimisers for analysing their performance on specific problems.

Regarding drawbacks, the expert noted that the appearance of the extrema graphs may be sensitive to methodological details. The expert highlighted with the following comments some examples of possible sensitivities: the use of edge nodes may introduce bias into the MDS procedure towards regions where there are more edges (i.e. where there are more maxima and minima closer together in Euclidean space); interpolating edge nodes between extrema "as the crow flies" restricts the view of the landscape to the regions directly connecting extrema; and finally, the NMMSO algorithm expends some proportion of its work towards finding the global optimum, so it may ignore certain optima.

We agree that our method is highly dependent on the efficacy of the multimodal optimiser in finding distinct extrema. Nonetheless, we believe the sensitivity due to this is minimal in our demonstration, especially due to the high number of function evaluations used for the known synthetic problems explored in this paper, and focus here on evaluation of the visualisation technique itself.

During this informal discussion of benefits and drawbacks, the expert provided insights that may lead to potential further work, detailed in section 6 below. We conclude based on the expert's feedback that extrema graphs provide the benefit of capturing a realistic idea of landscape ruggedness, while potentially being sensitive to details in the methodology used to generate these visualisations.

6 DISCUSSION AND CONCLUSIONS

In this work, we have presented a novel methodology for visualising fitness landscapes through the use of extrema graphs. We have shown that the incorporation of both minima and maxima can provide valuable information, and that appropriate use of dimensionality reduction is key for visualising ruggedness of the landscape. The use of sampled "edge nodes" replacing edges can then be used to visualise the folding of the search space.

6.1 Limitations and Future Work

Some limitations of extrema graphs are accepted here as a trade-off for associated benefits. For example, the extrema included in the graph are limited to those identified by NMMSO, however the use of just two runs of this optimiser keeps the computational complexity low. Similarly, the visual clutter introduced by edge nodes is accepted for the benefit of the information gained about the folding of the search space, the general characteristics of the landscape, and how this changes for a given value of ρ.

Some interesting directions for areas of future work were identified during the expert interview segment of our evaluation, based on suggestions made by the expert. These included:

- Replacing edge nodes with edges with coloured line segments. Where these lines intersect with different colours would then give some information about how the search space has been folded during dimensionality reduction.
- Investigating a methodology for automatically determining the value of the radius percentage hyperparameter.

One of the commonly cited reasons for a visualisation, especially in the field of *explainable AI* (XAI), is that we seek to represent a complex concept to a lay audience. The expert interview conducted in this study has shown that extrema graphs can convey useful information characterising problem landscapes to experts, however it would be beneficial to scale the human-centered evaluation by interviewing more experts. Since access to experts is limited, this is an inherent limitation to expert interviews.

Additionally, while we have not tested whether our visualisations can be understood by a non-expert, it would likely be more difficult for a non-expert to perform such analysis. Two avenues of future work are therefore to explore the extent to which these are accessible to non-expert users, and – in the event that they are not intuitive – to perform an investigation into approaches by which the technical barrier to entry could be reduced.

Faster methods for multidimensional scaling could be used to draw our extrema graphs, similar to stressed-based layout methods in graph drawing [9] which have been accelerated via many methods [6, 10, 13, 26]. Using a more scalable MDS algorithm would allow extrema graphs to scale to larger data sets.

Beyond enhancing the visualisation, a larger range of problems should be explored, including those with discrete representations, as well as those with multi-objective search spaces. A key consideration will be how to represent solution quality, and whether trade-offs between solutions that define such problems are usefully visualised in an extrema graph.

One of the primary motivations behind developing the method was to investigate the fitness landscape of a maximum likelihood estimation (MLE) function. When fitting a mathematical model, e.g. a compartmental model for disease progression in COVID-19, to real-world observations, we often query whether the parameters are identifiable [12]. Specifically, this relates to the identification of the convexity of the relevant MLE function, and there are common statistical procedures in biomathematics for this purpose. Current work focuses on deploying extrema graphs to the parameter identifiability problem, to complement standard parameter identifiability approaches.

Acknowledgement

This work was supported by the Engineering and Physical Science Research Council [grant numbers EP/S023992/1 and EP/W01226X/1]. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission. All underlying data to support the conclusions are provided within this paper.

REFERENCES

[1] Jason Adair, Gabriela Ochoa, and Katherine M. Malan. 2019. Local Optima Networks for Continuous Fitness Landscapes. In *Proc. Genetic and Evolutionary Computation Conference (GECCO '19 Companion)*. 1407–1414.

[2] Ulrik Brandes and Christian Pich. 2007. Eigensolver Methods for Progressive Multidimensional Scaling of Large Data. In *Graph Drawing*, Michael Kaufmann and Dorothea Wagner (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 42–53.

[3] Michael A. A. Cox and Trevor F. Cox. 2008. *Multidimensional Scaling*. Springer Berlin Heidelberg, Berlin, Heidelberg, 315–347. https://doi.org/10.1007/978-3-540-33037-0_14

[4] Jonathan E. Fieldsend. 2014. Running Up Those Hills: Multi-modal search with the niching migratory multi-swarm optimiser. In *2014 IEEE Congress on Evolutionary Computation (CEC)*. 2593–2600. https://doi.org/10.1109/CEC.2014.6900309

[5] Jonathan E. Fieldsend and Khulood Alyahya. 2019. Visualising the Landscape of Multi-Objective Problems using Local Optima Networks. In *Proc. Genetic and Evolutionary Computation Conference (GECCO '19 Companion)*. 1421–1429.

[6] Emden R. Gansner, Yehuda Koren, and Stephen North. 2005. Graph Drawing by Stress Majorization. In *Graph Drawing*, János Pach (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 239–250.

[7] T. Itoh and K. Koyamada. 1994. Isosurface generation by using extrema graphs. In *Proceedings Visualization '94*. 77–83. https://doi.org/10.1109/VISUAL.1994.346334

[8] T. Itoh and K. Koyamada. 1995. Automatic isosurface propagation using an extrema graph and sorted boundary cell lists. *IEEE Transactions on Visualization and Computer Graphics* 1, 4 (1995), 319–327. https://doi.org/10.1109/2945.485619

[9] Tomihisa Kamada and Satoru Kawai. 1989. An algorithm for drawing general undirected graphs. *Inform. Process. Lett.* 31, 1 (1989), 7–15. https://doi.org/10.1016/0020-0190(89)90102-6

[10] Mirza Klimenta and Ulrik Brandes. 2013. Graph Drawing by Classical Multidimensional Scaling: New Perspectives. In *Graph Drawing*, Walter Didimo and Maurizio Patrignani (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 55–66.

[11] Stephen G. Kobourov. 2013. Force-directed drawing algorithms. In *Handbook of Graph Drawing and Visualization*, Roberto Tamassia (Ed.). CRC Press, Chapter 12, 383–408.

[12] Clemens Kreutz, Andreas Raue, Daniel Kaschek, and Jens Timmer. 2013. Profile likelihood in systems biology. *The FEBS journal* 280, 11 (2013), 2564–2571.

[13] J. F. Kruiger, P. E. Rauber, R. M. Martins, A. Kerren, S. Kobourov, and A. C. Telea. 2017. Graph Layouts by t-SNE. *Computer Graphics Forum* 36, 3 (2017), 283–294. https://doi.org/10.1111/cgf.13187

[14] J.B. Kruskal. 1964. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29 (1964), 115–129.

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

[16] Krzysztof Michalak. 2019. Low-Dimensional Euclidean Embedding for Visualization of Search Spaces in Combinatorial Optimization. *IEEE Transactions on Evolutionary Computation* 23, 2 (2019), 232–246. https://doi.org/10.1109/TEVC.2018.2846636

[17] Gabriela Ochoa, Katherine M Malan, and Christian Blum. 2021. Search trajectory networks: A tool for analysing and visualising the behaviour of metaheuristics. *Applied Soft Computing* 109 (2021), 107492.

[18] Gabriela Ochoa, Marco Tomassini, Sebástien Vérel, and Christian Darabos. 2008. A Study of NK Landscapes' Basins and Local Optima Networks. In *Proc. Genetic and Evolutionary Computation Conference (GECCO 2008)*. 555–562.

[19] Gabriela Ochoa and Nadarajen Veerapen. 2016. Additional dimensions to the study of funnels in combinatorial landscapes. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*. 373–380.

[20] Vin Silva and Joshua Tenenbaum. 2004. Sparse Multidimensional Scaling using Landmark Points. *Technology* (01 2004).

[21] M. Tory and T. Moller. 2005. Evaluating visualizations: do expert reviews work? *IEEE Computer Graphics and Applications* 25, 5 (2005), 8–11. https://doi.org/10.1109/MCG.2005.102

[22] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9 (11 2008), 2579–2605.

[23] Nadarajen Veerapen and Gabriela Ochoa. 2018. Visualising the global structure of search landscapes: genetic improvement as a case study. *Genetic Programming and Evolvable Machines* 19 (09 2018), 1–33. https://doi.org/10.1007/s10710-018-9328-1

[24] David J. Walker, Richard M. Everson, and Jonathan E. Fieldsend. 2013. Visualizing Mutually Nondominating Solution Sets in Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation* 17, 2 (2013), 165–184.

[25] Mathew J. Walter, David J. Walker, and Matthew J. Craven. 2020. Visualising evolution history in multi-and many-objective optimisation. In *Proc. Parallel Problem Solving from Nature (PPSN)*. 299–312.

[26] J. X. Zheng, S. Pawar, and D. M. Goodman. 2019. Graph Drawing by Stochastic Gradient Descent. *IEEE Transactions on Visualization and Computer Graphics* 25, 9 (2019), 2738–2748. https://doi.org/10.1109/TVCG.2018.2859997

A FUNCTION DEFINITIONS

Here, we present the mathematical formulations of the benchmark functions used in our paper for extrema graph visualisations.

Ackley Function. The Ackley function is defined in n dimensions as:

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

Our search bounds were defined by the hypercube $x_i \in [-32.768, 32, 768]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{0}$ with a function value of $f(\mathbf{0}) = 0$.

Griewank Function. The Griewank function is defined in n dimensions as:

$$f(\mathbf{x}) = \sum_{i=1}^{n}\frac{x_i^2}{4000} - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

Our search bounds were defined by the hypercube $x_i \in [-5, 5]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{0}$ with a function value of $f(\mathbf{0}) = 0$.

Rastrigin Function. The Rastrigin function is defined in n dimensions as:

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

Our search bounds were defined by the hypercube $x_i \in [-5.12, 5.12]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{0}$ with a function value of $f(\mathbf{0}) = 0$.

Rosenbrock Function. The Rosenbrock function is defined in n dimensions as:

$$f(\mathbf{x}) = \sum_{i=1}^{n-1}\left(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right)$$

Our search bounds were defined by the hypercube $x_i \in [-5, 10]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{1} = (1, \ldots, 1)$ with a function value of $f(\mathbf{1}) = 0$.

Schwefel Function. The Schwefel function is defined in n dimensions as:

$$f(\mathbf{x}) = 418.9829n - \sum_{i=1}^{n}x_i\sin(\sqrt{|x_i|})$$

Our search bounds were defined by the hypercube $x_i \in [-500, 500]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{420.9687} = (420.9687, \ldots, 420.9687)$ with a function value of $f(\mathbf{420.9687}) = 0$.

Sphere Function. The Sphere function is defined in n dimensions as:

$$f(\mathbf{x}) = \sum_{i=1}^{n}x_i^2$$

Our search bounds were defined by the hypercube $x_i \in [-5.12, 5.12]$, and in this region there is one global minimum at $\mathbf{x} = \mathbf{0}$ with a function value of $f(\mathbf{0}) = 0$.

B SUPPLEMENTARY VISUALISATIONS

In addition to the 2D, 3D and 5D examples created for the benchmark functions, we also generated one example in 10D as a proof-of-concept for higher dimensions. Included here is an extrema graph for the Sphere function in 10D.

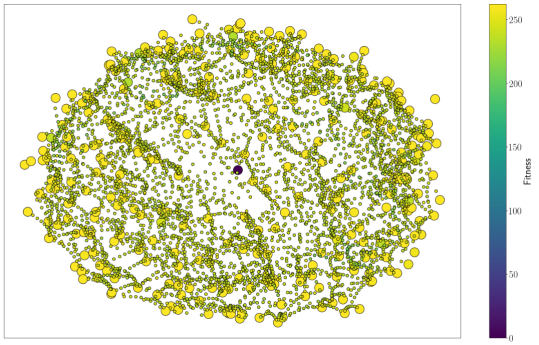

Figure 4: An extrema graph for the Sphere function in 10 dimensions