Global Landscape Structure and the Random MAX-SAT Phase Transition

Gabriela Ochoa[1], Francisco Chicano[2], and Marco Tomassini[3]

[1] Computing Science and Mathematics, University of Stirling, Stirling, Scotland, UK
[2] ITIS Software, University of Malaga, Malaga, Spain
[3] Faculty of Business and Economics, Information Systems Department, University of Lausanne, Lausanne, Switzerland

Abstract. We revisit the fitness landscape structure of random MAX-SAT instances, and address the question:*what structural features change when we go from easy underconstrained instances to hard overconstrained ones?* Some standard techniques such as autocorrelation analysis fail to explain what makes instances hard to solve for stochastic local search algorithms, indicating that deeper landscape features are required to explain the observed performance differences. We address this question by means of local optima network (LON) analysis and visualisation. Our results reveal that the number, size, and, most importantly, the connectivity pattern of local and global optima change significantly over the easy-hard transition. Our empirical results suggests that the landscape of hard MAX-SAT instances may feature sub-optimal funnels, that is, clusters of sub-optimal solutions where stochastic local search methods can get trapped.

1 Introduction

Understanding when a specific class of problems go from being computationally easy to hard, remains a fundamental question. It is well-known that random instances of satisfiability problems exhibit a dramatic easy-to-hard phase transition with respect to the problem constrainedness [1,2,3]. Some standard fitness landscape analysis techniques such as the correlation structure fall short in explaining the performance differences of local search algorithms when solving easy and hard instances [4]. Studies of the configuration landscape of a set of local optima show that there exist big-valley structures (also called clusters) in the landscapes of 3-SAT and MAX-3SAT [5], and studies of the size and characteristics of local optima and plateaus [6] do offer interesting insight and help to explain performance differences. These studies, however, do not convey a view of the connectivity structure of local optima in MAX-SAT, as seen by stochastic local search algorithms.

Our motivating research question is: *what fitness landscape features change when we go from easy underconstrained instances to hard overconstrained ones?* Another motivation for our study is the lack of tools for visualising high-dimensional fitness landscapes, specially in the presence of high levels of neutrality as is

known to be the case for MAX-SAT. Neutrality is present in a landscape when neighbouring points have equal fitness values. A discrete landscape is regarded as neutral if a substantial fraction of adjacent pairs of solutions are neutral [7].

We conduct a detailed experimental investigation of the phase transition from underconstrained to overconstrained randomly generated instances of MAX-3SAT problems. We explore this phenomenon by means of local optima networks (LONs) [8,9] analysis and visualisation. In particular, we use the compressed LON model (CLON) [10], which allows us to deal with the high levels of neutrality observed in MAX-SAT. Our contribution is to offer a new set of fitness landscape features and visualisation tools reflecting the number, size, and, most importantly, the connectivity pattern of optima, which capture and help to explain why the computational effort of stochastic local search methods increases dramatically in the phase transition region and beyond.

2   SAT, MAX-SAT and the Phase Transition

The propositional satisfiability problem (SAT) is a prominent combinatorial decision problem with a central role in several areas of computer science. Given a Boolean formula, SAT checks if there is an assignment of variables to Boolean values such that the formula is satisfiable. The Boolean formula is commonly expressed as a conjunction of clauses (Conjunctive Normal Form). A clause is a list of literals (a Boolean variable or its negation) that is satisfied if at least one literal is true. The Boolean formula is satisfiable if all the clauses are true. MAX-SAT is the optimisation version of SAT, where the goal is to find an assignment that maximises the number of satisfied clauses.

It is well known [1,2,3] that random instances of the kSAT problem display a computational phase transition for a certain critical value α_c of the ratio $\alpha = m/n$ between the number of clauses m and the number of variables n. For fixed integers k and n, the probability that a randomly generated formula is satisfiable is a decreasing function of α. For $\alpha \to 0$ the probability that the formula is satisfiable goes to 1; it goes to 0 for $\alpha \to \infty$. Thus, a random formula is satisfiable for $\alpha < \alpha_c$ and it is unsatisfiable for $\alpha > \alpha_c$, with probability approaching 1 as $n \to \infty$. The crossover from high to low probability becomes sharper as n increases and there is a transition at a finite value of α_c which, for 3SAT, is $\alpha_c \approx 4.3$. For a given n, at low α the problems are easy but in the phase transition region they become hard and the computational effort has a peak at the SAT/UNSAT boundary.

Zhang [11] investigated the relationship between the phase transitions of 3-SAT and MAX-3SAT, and found a near linear correlation between these two phase transitions. The computational cost of MAX-3SAT envelops the computational cost peaks of 3-SAT. It is worth noticing that in decision problems there is an easy-hard-easy phase transition, while in optimisation problems the transition is easy-hard. That is, the computational cost remains high after the transition.

3 Local Optima Networks

The local optima network (LON) model [8] is a tool to capture the global structure of fitness landscape as seen by stochastic local search algorithms. In this paper, we use a coarse variant of LONs, the compressed LONs (CLONs) [10], which helps us to model the structure of landscapes with high amounts of neutrality. We describe below the LON model, before introducing the compressed model (CLON).

3.1 LON Model

A fitness landscape [12] is a triplet (S, N, f) where S is a set of potential solutions i.e., a search space, $N : S \longrightarrow 2^S$, is a neighbourhood structure, a function that assigns a set of neighbours $N(s)$ to every solution $s \in S$, and $f : S \longrightarrow \mathbb{R}$ is a fitness function that can be pictured as the *height* of the corresponding solutions.

In our study, the search space is $\{0, 1\}^n$, i.e., the space of binary strings of length n, so its size is 2^n. As neighbourhood, we consider the standard Hamming distance 1 neighbourhood, that is, the set of all solutions at a maximum Hamming distance of 1 from the current solution.

Local optima. A local optimum, which in MAX-SAT is a maximum, is a solution l such that $\forall s \in N(l)$, $f(l) \geq f(s)$. Notice that the inequality is not strict, in order to allow the treatment of neutrality (local optima of equal fitness), which is known to widely occur on MAX-SAT. The set of local optima, which corresponds to the set of nodes in the network model, is denoted by L.

Perturbation edges. Edges are directed and based on the perturbation operator (p bit flips). There is an edge from local optimum l_1 to local optimum l_2, if l_2 can be obtained after applying a random perturbation (p bit flips) to l_1 followed by hill climbing. Edges are weighted with estimated frequencies of transition. We determined the edge weights in a sampling process. The weight is the number of times a transition between two local optima occurred. The set of edges is denoted by E.

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

3.2 Compressed LON Model

When the number of local optima is high in a LON it is difficult to visualise the structure of the landscape. A natural way of reducing the model size in landscapes with high-levels of neutrality is to redefine the nodes of the model. The compressed LON model joins the local optima that are connected and have the same fitness value.

Compressed local optima. A compressed local optimum is a set of connected nodes (a connected component) in the LON with the same fitness value. This set of compressed optima, denoted by CL, corresponds to the set of nodes in the Compressed LON model.

Compressed Perturbation edges. The set of perturbation edges is defined as above for the LON model. However, after compression, there are no edges between nodes with the same fitness, as connected components with the same fitness become a single node. The set of edges among compressed nodes are also aggregated and their weights summed. We call this set compressed edges, CE.

Compressed LON. Is the directed graph $\text{CLON} = (CL, CE)$, where nodes are the compressed local optima CL and edges the compressed perturbation edge set CE.

4 Methodology

4.1 Benchmark Instances

For the computational experiments, we used unweighted MAX3-SAT instances from the well-studied Uniform Random 3-SAT distribution [2]. Our instances have a relatively low number $n = 50$ of variables to allow us to compute the global optimum in all of them. Instances with $n = 50$ are not a challenge for state-of-the-art MAX-SAT solvers. However, $n = 50$ is large enough to well capture the performance transition region [2,3], thus suitable for our study. Our goal is to gain a deeper understanding of the fitness landscape structure. Since we are interested in studying instances around the phase transition, we generated random instances with $\alpha \in [3.0, 5.0]$ in steps of 0.2. For each value of α we generated 10 instances with different random seeds. The source code of the instance generator and the instructions to replicate the experiments are available at: `https://github.com/jfrchicanog/EfficientHillClimbers`.

4.2 Sampling and Construction of the Network Models

To construct the network models for a given instance, we aggregate the (unique) local optima and transition edges obtained by 20 runs of a fast iterated local search (ILS) algorithm [13] that incorporates Grey-Box optimisation techniques [14]. An outline can be found in Algorithm 1. Iterated local search is a simple yet powerful metaheuristic that combines two steps: one for reaching local optima efficiently, and the other for escaping local optima (known as the perturbation step). The stopping condition was set as a fixed running time (60 s). Weights are added to edges indicating the number of times they appear in the sampling process. In our ILS implementation, the perturbation step flips 10% of the variables selected at random (which corresponds to 5 bit flips for $n = 50$). The local search operator is a first improvement local search applied

Algorithm 1 Iterated Local Search

```
 1: x ← generateRandomSolution();
 2: x ← applyLocalSearch(x);
 3: while not stopping condition do
 4:     y ← perturb (x);
 5:     y ← applyLocalSearch(y);
 6:     reportEdge(x,y);
 7:     if f(y) > f(x) then
 8:         x ← y;
 9:     end if
10: end while
11: return  x;
```

in the 1-flip neighbourhood. That is, if a flip in one bit of the current solution increases the number of satisfied clauses, the current solution is replaced by the new one. The local search is applied until a local optimum is reached (no neighbour can increase the number of satisfied clauses). A new local optimum is only accepted in Line 7 if it improves the incumbent solution. However, we report all the edges encountered between local optima in Line 6, which includes neutral and worsening edges. All the local optima (and edges) visited after finding the global optimum are removed from the LON. The reason is that they would bias the metrics, since they generate only worsening edges and they do no not provide additional information on the difficulty of the search process.

4.3   Determining the Global Optimum

For all the instances, we computed the global optimum using exact methods. Instead of using an exhaustive enumeration (which could take a long time) we used minisat[4] to prove that there is no better solution than the one provided by ILS. The approach to prove optimality changes depending on the solution provided by ILS:

- If ILS finds an assignment satisfying all the clauses, the formula is satisfiable and the global optimum is found.
- If the optimal assignment found by ILS leaves one single clause unsatisfied, we apply minisat to check if the formula is satisfiable. If it is unsatisfiable, then the global optimum was found by ILS, otherwise (formula satisfiable) ILS didn't find a global optimum.
- If the optimal assignment found by ILS leaves u clauses unsatisfied, then we generate all the MAX-3SAT instances derived from the original instance where exactly $u - 1$ different clauses are removed and apply minisat to all of them. If minisat finds all of them unsatisfiable, then ILS found the global optimum. If minisat finds at least one of them satisfiable, then a better assignment can be found with $u-1$ unsatisfied clauses at most, which means that ILS didn't find the global optimum.

---

[4] http://minisat.se

The previous procedure to figure out the global optimum can be costly when the number of unsatisfied clauses in the optimal assignment is high. We expected this to happen when α is high. However, in the range of values of α used in the experiments the maximum number of unsatisfied clauses was 2, which made it possible to apply minisat in the way described above to certify that the global optimum was reached.

5 Results

5.1 Performance and Network Metrics

Our aim is to identify fitness landscape features that correlate with and help to explain the search difficulty of stochastic search algorithms. In order to measure search difficulty, we selected WalkSAT, a well-known local search algorithm for SAT and MAX-SAT [15]. WalkSAT has a wide influence among modern local search algorithms and is known to be very efficient in solving random 3-SAT and MAX-3SAT instances. We ran WalkSAT 10 000 times per instance and counted the number of bit flips (steps) it needs to find a global optimum. We have 10 instances per α value, thus a total of 100 000 measures per α. The distribution of this metric is shown in Fig. 1 (steps) with log scale, indicating a large variability, but a clear exponential increase in the search cost with increasing α.

Table 1: Description of Metrics.

Performance Metric	
steps	Number of bit flips (steps) by WalkSAT before reaching the global optimum.
LON Metrics	
nodes	Number of nodes (local and global optima).
global	Proportion of nodes that are global optima.
$edges_e$	Proportion of edges between nodes with equal fitness.
CLON Metrics	
cnodes	Number of compressed nodes.
subopt-size	Size of the largest suboptimal compressed node.
path-length	Average length of the shortest paths from start nodes to the global optimum.
$cedges_w$	Proportion of edges to compressed nodes with worse fitness.

For each of the 10 instances per α value, we extracted the LON and CLON models and computed the measurements described in Table 1. Network metrics are shown as the distribution of values over the 10 generated instances (Figs. 1 and 2); a large variability is observed across the 10 instances, meaning that

different instances will have different structure and performance, this is known to be the case for randomly generated instances. Results are shown as box-plots with instances grouped by the value of α in order to analyse the variation in the metrics as the value of α changes.
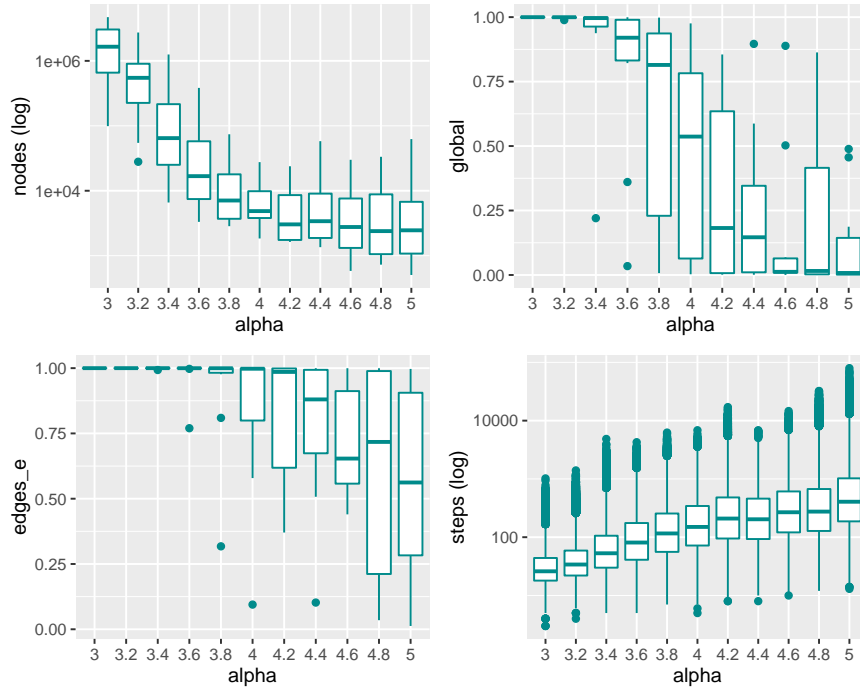

Fig. 1: Distribution of the LON metrics and the performance metric, as described in Table 1, for all values of alpha.

As Figure 1 indicates, the number of nodes (unique local and global optima) visited decreases exponentially with increasing α. For low values of α, over one million nodes are visited. This number drops to a few thousands and even several hundreds for high values of α. The proportion of global optima (*global* in Fig. 1) reveals that for $\alpha < 4.0$, the vast majority of nodes (above 75%) are global optima, and this proportion drastically decreases for larger values of α. This is a known result [1]; in the underconstrained region instances are satisfiable and the density of solutions is high, thus making it relatively easy to find a global optimum. For high values of α, instances are overconstrained, making it unlikely to find a solution (when the formula is satisfiable), or an assignment with the maximum possible number of true clauses (for unsatisfiable formulas). Our network analysis complements this finding by revealing that most of the search transitions in the underconstrained region are among candidate solutions

with equal fitness (traversing plateaus). This is revealed by the high proportion of equal edges ($edges_e$ in Fig. 1), which is almost 100% for $\alpha < 4.0$. Our network models capture search transitions that are either improving, deteriorating or neutral. The proportion of deteriorating transitions also correlates with search difficulty as is revealed by our analysis of the compressed LONs (Fig. 2).
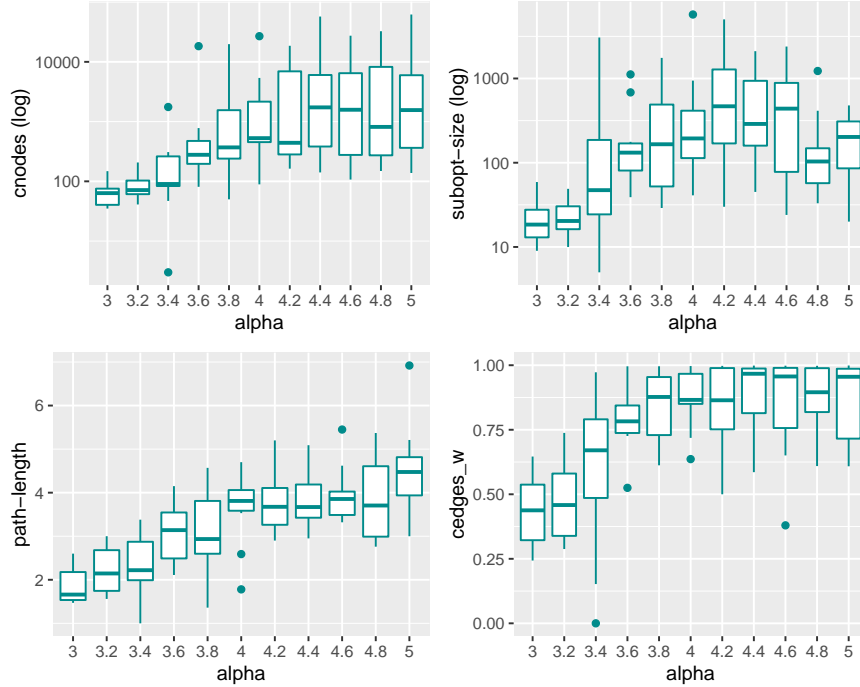

Fig. 2: Distribution of the CLON metrics, as described in Table 1, for all values of alpha.

Figure 2 shows the metrics distribution for the compressed LON models. Compressed nodes aggregate connected local optima with the same fitness. The number of compressed nodes (*cnodes* in Fig. 2) shows an opposite tendency than the number of nodes (*nodes* in Fig. 1), that is, the number of compressed nodes increases exponentially with increasing α. This is related to the proportion of equal edges in the LON model ($edges_e$ in Fig. 1); if there are many equal edges, there will be many local optima per compressed node, reducing their number while increasing their size. This means that, for low α values, CLONs will have a few nodes, indicating an easy to traverse fitness landscape. We can expect in this case a big node in the CLON capturing all the global optima, this is what our network visualisations reveal (Figs. 3a, and 3b). The average path lengths towards the global optimum are also correlated with search difficulty (*path-length*

in Fig. 2); shorter paths (of 1 to 4 hops) are observed in the underconstrained region. The size of the largest sub-optimal compressed node (*subopt-size* in Fig. 2) increases exponentially with increasing α, which helps to explain the increased search cost in the overconstrained region, search processes spend time traversing sub-optimal plateaus. The CLON models only have improving and deteriorating edges, as the neutral edges are collapsed in the compressed nodes. When looking at the proportion of worsening edges ($edges_w$ in Fig. 2), we can observe that it increases noticeably in the overconstrained region; for $\alpha > 4.0$, over 80% of the search transitions are deteriorating, indicating that search processes spend a long time finding exits from sub-optimal compressed nodes. These observations are supported by our network visualisations (Figs. 3c - 3f), where larger sub-optimal nodes, higher proportion of deteriorating (blue) edges, and longer path lengths to the global optima can be observed.

5.2   Visualisation

The network visualisations in Figure 3 capture the compressed local optima networks (CLONs) on representative instances with different values of α. Plots were produced with the R statistics language and the igraph library [16]. Network plots are decorated to reflect features relevant to search dynamic. Single optima are visualised as circles of fixed size. When more than one local optimum is compressed in a local optimum plateau, the node is represented as a rectangle with length proportional to its size (ie. number of local optimum configurations). Red nodes are global optima, green nodes indicate the start configuration of trajectories in the sampling process; recall that our implementation combines 20 independent ILS runs to construct the networks; thus up to 20 green starting nodes can be visualised in the figures. Grey identifies the intermediate nodes, with grey edges representing improving transitions. Blue edges indicate deteriorating transitions, with blue nodes indicating the end of these transitions. The edges width is proportional to the sampled frequency of transitions, thus thick edges represent common search paths. In order to have manageable images, only 5% of the worsening transitions are shown. That is, the networks are pruned before visualisation by removing 95% of the deteriorating edges selected uniformly at random. The CLON visualisations in Figure 3 support what the metrics indicate. Instances get harder to search as α increases, because more intermediate nodes appear and the trajectories get longer. The size of the global optimum node decreases drastically with increasing α. For low values of α (plots 3a, and 3b) a large "central" global optimum node is observed, which attracts all the search trajectories. As α increases, (plots 3c - 3f), the size of the global optimum decreases, and additional disconnected global optima may appear. Clearly, with increasing α, large sub-optimal compressed nodes emerge, and the proportion of deteriorating edges (visualised in blue) is larger. This indicates that the search process gets trapped in large sub-optimal nodes, requiring several escape attempts before finding an exit towards the global optimum. With increasing α, the path lengths of the trajectories to the global optimum (optima), measured as the number of edges they contain, tends to increase.

(a) $m = 150$, $\alpha = 3.0$

(b) $m = 170$, $\alpha = 3.4$

(c) $m = 200$, $\alpha = 4.0$

(d) $m = 210$, $\alpha = 4.2$

(e) $m = 220$, $\alpha = 4.4$
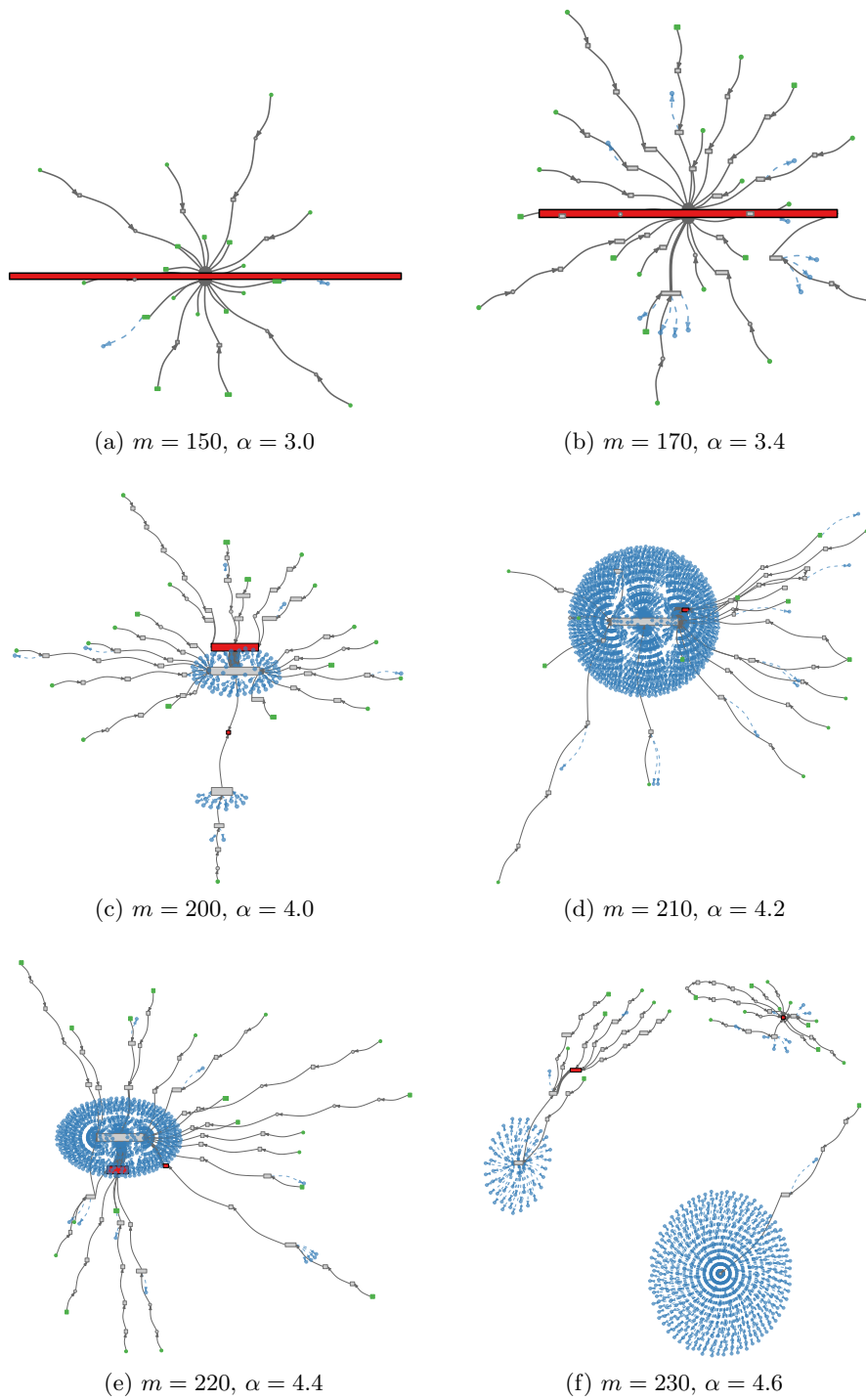
(f) $m = 230$, $\alpha = 4.6$

Fig. 3: CLONs for representative benchmark instances. The global optimum (optima) are indicated in red, while the start nodes in green. Edge widths are proportional to their weight. Blue dashed edges indicate deteriorating transitions.

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." [17]. It has been suggested that the energy landscape of proteins is characterised by a single deep funnel, a feature that underpins their ability to fold to their native state. In contrast, some shorter polymer chains (polypeptides) that misfold, are expected to have other funnels that can act as traps. We follow this loose definition here, by a funnel we mean a grouping of local optima in a coarse-grained structure around a high quality solution. According to this definition, most of the instances we analysed across the different values of α reveal a single funnel. This is consistent with the results reported in [5], showing that there exist big valley structures in the landscapes of 3-SAT and MAX-3SAT. However, in some of the hardest overconstrained instances, our results reveal multiple funnel structures. An example is visualised in Fig. 3f, where three funnels can be identified as separate connected components. The two structures at the top converge to a global optimum (red node), so we can consider them as global funnels, whereas the structure at the bottom is a sub-optimal funnel. It is worth stressing that the global structures here encountered are approximations, as our approach is based on sampling, and on a particular value of the perturbation strength of the ILS sampling process. We argue that this is still an interesting observation, as it suggests that multiple funnels may exist in hard to solve MAX-SAT instances. When sub-optimal funnels exist, search can get trapped and fail to reach the global optimum despite a large computational effort.

6   Related Work

Some detailed analyses of random MAX-SAT fitness landscapes, using standard techniques, concentrate on the over-constrained region [18,19]. More relevant to this research is the work presented in [6,4,20], where the landscape structure at and around the phase transition of random MAX-3SAT is explicitly examined. Frank et al. [6] identity several interesting features of plateaus (such as their size and number of exits) that impact the performance of local search algorithms. Sutton et al. [20] also studied the neutral regions by establishing theoretical bounds on plateau sizes, and assessing their accuracy on sampled problem instances. Sutton et al. [4] computed the exact correlation structure of random MAX-3SAT instances, showing that the correlation structure is oblivious to the phase transition, that is landscapes before and after the phase transition show the same correlation structure. This last result indicates that alternative techniques for studying the global structure of fitness landscapes, such as our proposal in this article, are required to gain a deeper understanding of search difficulty.

7 Discussion and Conclusion

We revisited the global structure of random MAX-3SAT fitness landscapes across the transition from underconstrained to overconstrained instances, using local optima networks analysis and visualisation. Our results confirm some previous findings, but also bring new structural metrics that correlate and help to explain the increased search cost incurred by stochastic local search algorithms in the phase transition region and beyond. The compressed LON model proved very valuable as a tool to analyse and visualise the landscapes' global structure. Before the phase transition, a large global optimal node is observed that is easy to reach (after a few hops) by local search algorithms. During the phase transition and beyond, however, the size of the global optimum drastically decreases, while the size of sub-optimal compressed nodes increases. The proportion of transitions to deteriorating solutions greatly increases, as well as the length of the trajectories towards the global optimum.

Under our empirical setting, most of the instances we studied revealed a single 'valley' or 'funnel'. A rigorous and well established definition of funnels is still lacking in evolutionary computation. We take the term here to loosely refer to a grouping of local optima, forming a coarse-level gradient towards a high quality solution at the end. Some of the hardest instances we considered (for high values of α) showed several disconnected groups of local optima. We suggest that these groupings may be related to the notion of sub-optimal funnels. Multiple funnels have been empirically observed in other hard combinatorial landscapes [21,10,22], contributing to our understanding of why some instances are harder to solve than others. To the best of our knowledge, multiple funnels have not been documented in the study of MAX-SAT fitness landscapes. This observation deserves further investigation. The funnel structure, as studied by local optima networks, depends on the amount of perturbation defining the transition edges [22]. Moreover, when sampling is involved, the identification of any global structure is approximated. A more precise characterisation of these structures, as well as the study of larger and different classes of MAX-SAT instances, is left as future work.

References

1. Cheeseman, P., Kanefsky, B., Taylor, W.M.: Where the really hard problems are. In: International Joint Conference on Artificial Intelligence (IJCAI), Morgan

Kaufmann (1991) 331–337

2. Mitchell, D., Selman, B., Levesque, H.: Hard and easy distributions of SAT problems. In: National Conference on Artificial intelligence (AAAI). (1992) 459–465
3. Kirkpatrick, S., Selman, B.: Critical behavior in the satisfiability of random boolean expressions. Science **264**(5163) (1994) 1297–1301
4. Sutton, A.M., Whitley, L.D., Howe, A.E.: A polynomial time computation of the exact correlation structure of k-satisfiability landscapes. In: Genetic and Evolutionary Computation Conference (GECCO), ACM (2009) 365–372
5. Zhang, W.: Configuration landscape analysis and backbone guided local search. Part I: Satisfiability and maximum satisfiability. Artificial Intelligence **158** (2004) 1–26
6. Frank, J., Cheeseman, P., Stutz, J.: When gravity fails: Local search topology. Journal of Artificial Intelligence Research **7** (1997) 249–281
7. Reidys, C.M., Stadler, P.F.: Neutrality in fitness landscapes. Applied Mathematics and Computation **117**(2) (2001) 321 – 350
8. Ochoa, G., Tomassini, M., Verel, S., Darabos, C.: A study of NK landscapes' basins and local optima networks. In: Genetic and Evolutionary Computation Conference (GECCO), ACM (2008) 555–562
9. Verel, S., Ochoa, G., Tomassini, M.: Local optima networks of NK landscapes with neutrality. IEEE Transactions on Evolutionary Computation **15**(6) (2011) 783–797
10. Ochoa, G., Veerapen, N., Daolio, F., Tomassini, M.: Understanding phase transitions with local optima networks: Number partitioning as a case study. In: European on Conference Evolutionary Computation in Combinatorial Optimization (EvoCOP). Volume 10197 of Lecture Notes in Computer Science. (2017) 233–248
11. Zhang, W.: Phase transitions and backbones of 3-SAT and maximum 3-SAT. In: Principles and Practice of Constraint Programming (CP). Lecture Notes in Computer Science, Springer (2001)
12. Stadler, P.F.: Fitness landscapes. Appl. Math. and Comput **117** (2002) 187–207
13. Lourenço, H.R., Martin, O.C., Stützle, T.: Iterated Local Search. Handbook of Metaheuristics (2003) 320–353
14. Chicano, F., Whitley, L.D., Ochoa, G., Tinos, R.: Optimizing one million variable NK landscapes by hybridizing deterministic recombination and local search. In: Genetic and Evolutionary Computation Conference, ACM (2017) 753–760
15. Selman, B., Kautz, H.A., Cohen, B.: Local search strategies for satisfiability testing. In: DIMACS Series in Discrete Mathematics and Theoretical Computer Science. (1996) 521–532
16. Csardi, G., Nepusz, T.: The igraph software package for complex network research. InterJournal, Complex Systems **1695**(5) (2006) 1–9
17. Doye, J.P.K., Miller, M.A., Wales, D.J.: The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. Journal of Chemical Physics **110**(14) (1999) 6896–6906
18. Hoos, H.H., Smyth, K., Stützle, T.: Search space features underlying the performance of stochastic local search algorithms for MAX-SAT. In: Parallel Problem Solving from Nature (PPSN), Springer (2004) 51–60
19. Prugel-Bennett, A., Tayarani-Najaran, M.H.: Maximum satisfiability: Anatomy of the fitness landscape for a hard combinatorial optimization problem. IEEE Transactions on Evolutionary Computation **16**(3) (2011) 319–338
20. Sutton, A.M., Howe, A.E., Whitley, L.D.: Estimating bounds on expected plateau size in MAXSAT problems. In: International Workshop on Engineering Stochastic Local Search Algorithms, Springer (2009) 31–45

21. Hains, D., Whitley, L.D., Howe, A.E.: Revisiting the big valley search space structure in the TSP. JORS **62**(2) (2011) 305–312
22. Ochoa, G., Herrmann, S.: Perturbation strength and the global structure of QAP fitness landscapes. In: Parallel Problem Solving from Nature - PPSN XV. Volume 11102 of Lecture Notes in Computer Science., Springer (2018) 245–256