The Connectivity of NK Landscapes’ Basins: A Network Analysis
Sébastien Verel, Gabriela Ochoa, Marco Tomassini

To cite this version:
Sébastien Verel, Gabriela Ochoa, Marco Tomassini. The Connectivity of NK Landscapes’ Basins: A Network Analysis. S. Bullock, J. Noble, R. Watson, and M. A. Bedau. Artificial Life XI, Aug 2008, Winchester, France. MIT Press, Cambridge, MA, pp.648-655, 2008. <hal-00331864>

HAL Id: hal-00331864
https://hal.archives-ouvertes.fr/hal-00331864
Submitted on 20 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The Connectivity of NK Landscapes’ Basins: A Network Analysis

Sébastien Verel\textsuperscript{1}, Gabriela Ochoa\textsuperscript{2} and Marco Tomassini\textsuperscript{3}

\textsuperscript{1}University of Nice Sophia-Antipolis / CNRS, Nice, France
\textsuperscript{2}University of Nottingham, Nottingham, UK
\textsuperscript{3}University of Lausanne, Lausanne, Switzerland
verel@i3s.unice.fr, gxo@cs.nott.ac.uk, Marco.Tomassini@unil.ch

Abstract

We propose a network characterization of combinatorial fitness landscapes by adapting the notion of inherent networks proposed for energy surfaces \cite{3, 4}. We use the well-known family of NK landscapes as an example. In our case the inherent network is the graph where the vertices represent the local maxima in the landscape, and the edges account for the transition probabilities between their corresponding basins of attraction. We exhaustively extracted such networks on representative small NK landscape instances, and performed a statistical characterization of their properties. We found that most of these network properties can be related to the search difficulty on the underlying NK landscapes with varying values of K.

Introduction

Local optima are the very feature of a landscape that makes it rugged. Therefore, an understanding of the distribution of local optima is of utmost importance for the understanding of a landscape. Combinatorial landscapes refer to the finite search spaces generated by important discrete problems such as the traveling salesman problem and many others. A property of some combinatorial landscapes, which has been often observed, is that on average, local optima are much closer to the optimum than are randomly chosen points, and closer to each other than random points would be. In other words, the local optima are not randomly distributed, rather they tend to be clustered in a "central massif" (or "big valley" if we are minimizing). This globally convex landscape structure has been observed in the NK family of landscapes \cite{2}, and in other combinatorial optimization problems, such as the traveling salesman problem \cite{3}, graph bipartitioning \cite{4}, and flowshop scheduling \cite{5}.

In this study we seek to provide fundamental new insights into the structural organization of the local optima in NK landscapes, particularly into the connectivity of their basins of attraction. Combinatorial landscapes can be seen as a graph whose vertices are the possible configurations. If two configurations can be transformed into each other by a suitable operator move, then we can trace an edge between them. The resulting graph, with an indication of the fitness at each vertex, is a representation of the given problem fitness landscape. A useful simplification of the graphs for the energy landscapes of atomic clusters was introduced in \cite{3, 4}. The idea consists in taking as vertices of the graph not all the possible configurations, but only those that correspond to energy minima. For atomic clusters these are well-known, at least for relatively small assemblages. Two minima are considered connected, and thus an edge is traced between them, if the energy barrier separating them is sufficiently low. In this case there is a transition state, meaning that the system can jump from one minimum to the other by thermal fluctuations going through a saddle point in the energy hyper-surface. The values of these activation energies are mostly known experimentally or can be determined by simulation. In this way, a network can be built which is called the "inherent structure" or "inherent network" in \cite{3, 4}. We propose a network characterization of combinatorial fitness landscapes by adapting the notion of inherent networks described above. We use the well-known family of NK landscapes as an example because they are a useful tunable benchmark that can provide interesting information for more realistic combinatorial landscapes. In our case the inherent network is the graph where the vertices are all the local maxima and the edges account for transition probabilities between their corresponding basins of attraction. We exhaustively extract such networks on representative small NK landscape instances, and perform a statistical characterization of their properties. Our analysis was inspired, in particular, by the work of \cite{3, 4} on energy landscapes, and in general, by the field of complex networks \cite{8}. The study of networks has exploded across the academic world since the late 90’s. Researchers from the mathematical, biological, and social sciences have made substantial progress on some previously intractable problems, bringing new techniques, reformulating old ideas, and uncovering unexpected connections between seemingly different problems. We aim here at bringing the tools of network analysis for the study of problem hardness in combinatorial optimization.

The next section describes how combinatorial landscapes are mapped onto networks, and includes the relevant def-
Definitions and algorithms used in our study. The empirical network analysis of our selected \( NK \) landscape instances is presented next, followed by our conclusions and ideas for future work.

**Landscapes as Networks**

To model a physical energy landscape as a network, one needed to decide first on a definition both of a state of the system and how two states were connected. The states and their connections will then provide the nodes and edges of the network. For systems with continuous degrees of freedom, the author achieved this through the ‘inherent structure’ mapping. In this mapping each point in configuration space is associated with the minimum (or ‘inherent structure’) reached by following a steepest-descent path from that point. This mapping divides the configuration space into basins of attraction surrounding each minimum on the energy landscape.

Our goal is to adapt this idea to the context of combinatorial optimization. In our case, the nodes of the graph can be straightforwardly defined as the local maxima of the landscape. These maxima are obtained exhaustively by running a best-improvement local search algorithm (HillClimbing, see Algorithm 1) from every configuration of the search space. The definition of the edges, however, is a much more delicate matter. In our initial attempt we considered that two maxima \( i \) and \( j \) were connected (with an undirected and unweighted edge), if there exists at least one pair of solutions at Hamming distance one \( s_i \) and \( s_j \), one in each basin of attraction \( (b_i, b_j) \). We found empirically on small instances of \( NK \) landscapes, that such definition produced densely connected graphs, with very low (\( \leq 2 \)) average path length between nodes for all \( K \). Therefore, apart from the already known increase in the number of optima with increasing \( K \), no other network property accounted for the increase in search difficulty. Furthermore, a single pair of neighbors between adjacent basins, may not realistically account for actual basin transitions occurring when using common heuristic search algorithms. These considerations, motivated us to search for an alternative definition of the edges connecting local optima. In particular, we decided to associate weights to the edges that account for the transition probabilities between the basins of attraction of the local optima. More details on the relevant algorithms and formal definitions are given below.

**Definitions and Algorithms**

**Definition:** Fitness landscape.

A landscape is a triplet \( (S, V, f) \) where \( S \) is a set of potential solutions i.e. a search space, \( V : S \rightarrow 2^S \), a neighborhood structure, is a function that assigns to every \( s \in S \) a set of neighbors \( V(s) \), and \( f : S \rightarrow R \) is a fitness function that can be pictured as the height of the corresponding solutions.

In our study, the search space is composed by binary strings of length \( N \), therefore its size is \( 2^N \). The neighborhood is defined by the minimum possible move on a binary search space, that is, the 1-move or bit-flip operation. In consequence, for any given string \( s \) of length \( N \), the neighborhood size is \( |V(s)| = N \).

The **HillClimbing** algorithm to determine the local optima and therefore define the basins of attraction, is given below:

**Algorithm 1 HillClimbing**

Choose initial solution \( s \in S \)

repeat

choose \( s' \in V(s) \) such that \( f(s') = max_{x \in V(s)} f(x) \)

if \( f(s) \leq f(s') \) then

\( s \leftarrow s' \)

end if

until \( s \) is a Local optimum

**Definition:** Local optimum.

A local optimum is a solution \( s^* \) such that \( \forall s \in V(s^*) \), \( f(s) < f(s^*) \).

The **HillClimbing** algorithm defines a mapping from the search space \( S \) to the set of locally optimal solutions \( S^* \).

**Definition:** Basin of attraction.

The basin of attraction of a local optimum \( i \in S \) is the set \( b_i = \{ s \in S \mid \text{HillClimbing}(s) = i \} \). The size of the basin of attraction of a local optima \( i \) is the cardinality of \( b_i \).

**Definition:** Edge weight.

Notice that for a non-neutral fitness landscapes, as are \( NK \) landscapes, the basins of attraction as defined above, produce a partition of the configuration space \( S \). Therefore, \( S = \bigcup_{i \in S} b_i \) and \( \forall i \in S \forall j \neq i, b_i \cap b_j = \emptyset \)

For each solutions \( s \) and \( s' \), let us define \( p(s \rightarrow s') \) as the probability to pass from \( s \) to \( s' \) with the bit-flip operator. In the case of binary strings of size \( N \), and the neighborhood defined by the bit-flip operation, there are \( N \) neighbors for each solution, therefore:

\[
\begin{align*}
\text{if } s \in V(s) \text{, } p(s \rightarrow s) &= \frac{1}{N} \\
\text{if } s' \notin V(s) \text{, } p(s \rightarrow s') &= 0
\end{align*}
\]

We can now define the probability to pass from a solution \( s \in S \) to a solution belonging to the basin \( b_j \), as:

\[
p(s \rightarrow b_j) = \sum_{s' \in b_j} p(s \rightarrow s')
\]

Notice that \( p(s \rightarrow b_j) \leq 1 \).

Thus, the total probability of going from basin \( b_i \) to basin \( b_j \) is the average over all \( s \in b_i \) of the transition probabilities to solutions \( s' \in b_j \):

\[
p(b_i \rightarrow b_j) = \frac{1}{2^{b_i}} \sum_{s \in b_i} p(s \rightarrow b_j)
\]
$b_i$ is the size of the basin $b_i$. We are now prepared to define our ‘inherent’ network or network of local optima.

**Definition:** Local optima network. The local optima network $G = (S^*, E)$ is the graph where the nodes are the local optima $1$, and there is an edge $e_{ij} \in E$ with the weight $w_{ij} = p(b_i \rightarrow b_j)$ between two nodes $i$ and $j$ if $p(b_i \rightarrow b_j) > 0$.

According to our definition of weights, $w_{ij} = p(b_i \rightarrow b_j)$ may be different than $w_{ji} = p(b_j \rightarrow b_i)$. Two weights are needed in general, and we have an oriented transition graph.

**Empirical Basin and Network Analysis**

The $NK$ family of landscapes $6$ is a problem-independent model for constructing multimodal landscapes that can gradually be tuned from smooth to rugged. In the model, $N$ refers to the number of (binary) genes in the genotype (i.e. the string length) and $K$ to the number of genes that influence a particular gene (the epistatic interactions). By increasing the value of $K$ from 0 to $N - 1$, $NK$ landscapes can be tuned from smooth to rugged. The $k$ variables that form the context of the fitness contribution of gene $s_i$ can be chosen according to different models. The two most widely studied models are the random neighborhood model, where the $k$ variables are chosen randomly according to a uniform distribution among the $n - 1$ variables other than $s_i$, and the adjacent neighborhood model, in which the $k$ variables that are closest to $s_i$ in a total ordering $s_1, s_2, \ldots, s_n$ (using periodic boundaries). No significant differences between the two models were found $1$ in terms of global properties of the respective families of landscapes, such as mean number of local optima or autocorrelation length. Similarly, our preliminary studies on the characteristics of the $NK$ landscape optima networks did not show noticeable differences between the two neighborhood models. Therefore, we conducted our full study on the more general random model.

In order to avoid sampling problems that could bias the results, we used the largest values of $N$ that can still be analysed exhaustively with reasonable computational resources. We thus extracted the local optima networks of landscape instances with $N = 14, 16, 18$, and $K = 2, 4, 6, \ldots, N - 2, N - 1$. For each pair of $N$ and $K$ values, 30 randomly generated instances were explored. Therefore, the networks statistics reported below represent the average behaviour of 30 independent instances.

**Basins of Attraction**

Besides the maxima network, it is useful to describe the associated basins of attraction as these play a key role in search algorithms. Furthermore, some characteristics of the basins can be related to the optima network features. The notion of the basin of attraction of a local maximum has been presented before. We have exhaustively computed the size and number of all the basins of attraction for $N = 16$ and $N = 18$ and for all even $K$ values plus $K = N - 1$. In this section, we analyze the basins of attraction from several points of view as it is described below.

**Global optimum basin size versus $K$.** In Fig. 1 we plot the average size of the basin corresponding to the global maximum for $N = 16$ and $N = 18$, and all values of $K$ studied. The trend is clear: the basin shrinks very quickly with increasing $K$. This confirms that the higher the $K$ value, the more difficult for an stochastic search algorithm to locate the basin of attraction of the global optimum.

![Figure 1: Average of the relative size of the basin corresponding to the global maximum for each $K$ over 30 landscapes.](image1)

**Number of basins of a given size.** Fig. 2 shows the cumulative distribution of the number of basins of a given size (with regression line) for a representative instances with $N = 18$, $K = 4$. Table 1 shows the average (of 30 independent landscapes) correlation coefficients and linear regression coefficients (intercept $\alpha$ and slope $\beta$) between the number of nodes and the basin sizes for instances with $N = 16, 18$. Notice that distribution decays exponentially.

![Figure 2: Cumulative distribution of the number of basins of a given size with regression line. A representative landscape with $N = 18$, $K = 4$ is visualized. A lin-log scale is used.](image2)
or faster for the lower $K$ and it is closer to exponential for the higher $K$. This could be relevant to theoretical studies that estimate the size of attraction basins (see for example \textcite{5}). These studies often assume that the basin sizes are uniformly distributed, which is not the case for the $NK$ landscapes studied here. From the slopes $\beta$ of the regression lines (table \ref{table:1}) one can see that high values of $K$ give rise to steeper distributions (higher $\beta$ values). This indicates that there are fewer basins of large size for large values of $K$. Basins are thus broader for low values of $K$, which is consistent with the fact that those landscapes are smoother.

**Fitness of local optima versus their basin sizes.** The scatter-plots in Fig. \ref{fig:3} illustrate the correlation between the basin sizes of local maxima (in logarithmic scale) and their fitness values. Two representative instances for $N = 18$ and $K = 4, 8$ are shown. Notice that there is a clear positive correlation between the fitness values of maxima and their basins’ sizes. In other words, the higher the peak the wider tend to be its basin of attraction. Therefore, on average, with a stochastic local search algorithm, the global optimum would be easier to find than any other local optimum. This may seem surprising. But, we have to keep in mind that as the number of local optima increases (with increasing $K$), the global optimum basin is more difficult to reach by a stochastic local search algorithm (see Fig. \ref{fig:4}). This observation offers a mental picture of $NK$ landscapes: we can consider the landscape as composed of a large number of mountains (each corresponding to a basin of attraction), and those mountains are wider the taller the hilltops. Moreover, the size of a mountain basin grows exponentially with its height.

**General Network Statistics**

We now briefly describe the statistical measures used for our analysis of maxima networks.

The standard clustering coefficient \textcite{8} does not consider weighted edges. We thus use the *weighted clustering measure* proposed by \textcite{9}, which combines the topological information with the weight distribution of the network:

$$e^w(i) = \frac{1}{s_i (k_i - 1)} \sum_{j,h} w_{ij} + w_{ih} - 2 a_{ij} a_{jh} a_{hi}$$

where $s_i = \sum_{j \neq i} w_{ij}$, $a_{nm} = 1$ if $w_{nm} > 0$, $a_{nm} = 0$ if $w_{nm} = 0$ and $k_i = \sum_{j \neq i} a_{ij}$.

For each triple formed in the neighborhood of the vertex $i$, $e^w(i)$ counts the weight of the two participating edges of the vertex $i$. $C^w$ is defined as the weighted clustering coefficient averaged over all vertices of the network.

The standard topological characterization of networks is obtained by the analysis of the probability distribution $p(k)$ that a randomly chosen vertex has degree $k$. For our weighted networks, a characterization of weights is obtained

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
$K$ & $N = 16$ & $\beta$ \\
\hline
2 & -0.9440.0454 & 2.890.673 \\
4 & -0.9590.0310 & 4.190.554 \\
6 & -0.9670.0280 & 5.090.504 \\
8 & -0.9820.0110 & 5.970.321 \\
10 & -0.9850.0161 & 6.740.392 \\
12 & -0.9900.0088 & 7.470.346 \\
14 & -0.9940.0059 & 8.080.241 \\
15 & -0.9950.0044 & 8.370.240 \\
\hline
$N = 18$ & & \\
\hline
2 & -0.9990.0257 & 3.180.696 \\
4 & -0.9600.0409 & 4.570.617 \\
6 & -0.9700.0283 & 5.500.520 \\
8 & -0.9770.0238 & 6.440.485 \\
10 & -0.9850.0141 & 7.240.372 \\
12 & -0.9890.0108 & 7.080.370 \\
14 & -0.9940.0072 & 8.690.276 \\
16 & -0.9950.0056 & 9.330.249 \\
17 & -0.9950.0113 & 9.490.386 \\
\hline
\end{tabular}
\caption{Correlation coefficient ($\beta$), and linear regression coefficients (intercept ($\alpha$) and slope ($\beta$)) of the relationship between the basin size of optima and the cumulative number of nodes of a given (basin) size (in logarithmic scale: $\log(p(s)) = \alpha + \beta s + \epsilon$). The average and standard deviation values over 30 instances, are shown.}
\end{table}
by the connectivity and weight distribution \( p(w) \) that any
given edge has weight \( w \).

In our study, for each node \( i \), the sum of weights from the
node \( i \) is equal to 1. So, an important measure is the weight
\( w_{ii} \) of self-connecting edges (remaining in the same node).
We have the relation: \( w_{ii} + s_i = 1 \). The vertex strength, 
\( s_i \), is defined as \( s_i = \sum_{j \in V(i) - \{i\}} w_{ij} \), where the sum is
over the set \( V(i) - \{i\} \) of neighbors of \( i \). The strength of
a node is a generalization of the node’s connectivity giving
information about the number and importance of the edges.

Another network measure we report here is disparity \( Y_2(i) \), which measures how heterogeneous is the contributions
of the edges of node \( i \) to the total weight (strength):

\[
Y_2(i) = \sum_{j \neq i} \left( \frac{w_{ij}}{s_i} \right)^2
\]

The disparity could be averaged over the node with the
same degree \( k \). If all weights are nearby of \( s_i/k \), the disparity
for nodes of degree \( k \) is nearby \( 1/k \).

Finally, in order to compute the average distance (shortest
path) between two nodes on the optima network of a given
landscape, we considered the expected number of bit-flip
mutations to pass from one basin to the other. This expected
number can be computed by considering the inverse of the
transition probabilities between basins. In other words, if
we attach to the edges the inverse of the transition proba-
bilities, this value would represent the average number of ran-
dom mutations to pass from one basin to the other. More
formally, the distance (expected number of bit-flip muta-
tions) between two nodes is defined by \( d_{ij} = 1/w_{ij} \) where
\( w_{ij} = p(b_i \rightarrow b_j) \). Now, we can define the length of a
path between two nodes as being the sum of these distances along
the edges that connect the respective basins.

Detailed Study of Network Features

In this section we study in more depth some network fea-
tures which can be related to stochastic local search diffi-
culty on the underlying fitness landscapes. Table \( 2 \) reports
the average (over 30 independent instances for each \( N \) and
\( K \)) of the network properties described. \( \bar{n}_v \) and \( \bar{n}_e \) are,
respectively, the mean number of vertices and the mean num-
ber of edges of the graph for a given \( K \) rounded to the next
integer. \( C^w \) is the mean weighted clustering coefficient. \( \bar{Y} \)
is the mean disparity, and \( \bar{d} \) is the mean path length.

Clustering Coefficients. The fourth column of table \( 2 \)
lists the average values of the weighted clustering coefficients
for all \( N \) and \( K \). It is apparent that the clustering
coefficients decrease regularly with increasing \( K \) for all \( N \).
For the standard unweighted clustering, this would mean that
the larger \( K \) is, the less likely that two maxima which are
connected to a third one are themselves connected. Taking
weights, i.e. transition probabilities into account this means
that either there are fewer transitions between neighboring
basins for high \( K \), and/or the transitions are less likely to
occur. This confirms from a network point of view the com-
mon knowledge that search difficulty increases with \( K \).

Shortest Path to the Global Optimum. The average
shortest path lengths \( \bar{d} \) are listed in the sixth column of ta-
ble \( 2 \). Fig. \( 4 \) (top) is a graphical illustration of the average
shortest path length between optima for all the studied \( NK \)
landscapes. Notice that the shortest path increases with \( N \),
this is to be expected since the number of optima increases
exponentially with \( N \). More interestingly, for a given \( N \) the
shortest path increases with \( K \), up to \( K = 10 \), and then it
stagnates and even decreases slightly for the \( N = 18 \). This
is consistent with the well known fact that the search diffi-
culty in \( NK \) landscapes increases with \( K \). However, some
paths are more relevant from the point of view of a stochastic
local search algorithm following a trajectory over the max-
ima network. In order to better illustrate the relationship of
this network property with the search difficulty by heuristic
local search algorithms, Fig. \( 4 \) (bottom) shows the shortest
path length to the global optimum from all the other optima
in the landscape. The trend is clear, the path lengths to the
optimum increase steadily with increasing \( K \).

Weight Distribution. Here we report on the weight distri-
butions \( p(w) \) of the maxima network edges. Fig. \( 5 \) shows
the empirical probability distribution function for the cases
\( N = 16 \) and \( N = 18 \) (logarithmic binning has been used on

Figure 4: Average distance (shortest path) between nodes
(top), and average path length to the optimum from all the
other basins (bottom).
Table 2: \( N \times K \) landscapes network properties. Values are averages over 30 random instances, standard deviations are shown as subscripts. \( n_v \) and \( n_e \) represent the number of vertexes and edges (rounded to the next integer), \( C^w \), the mean weighted clustering coefficient. \( Y \) represent the mean disparity coefficient, \( d \) the mean path length (see text for definitions).

| \( K \) | \( n_v \) | \( n_e \) | \( C^w \) | \( Y \) | \( d \) |
|-------|-------|-------|-------|-------|-------|
| \( N = 14 \) |
| 2 | 146 | 200 | 0.980,0153 | 0.367,0094 | 76,194 |
| 4 | 70 | 3163 | 0.920,0139 | 0.148,0010 | 89,6 |
| 6 | 184 | 12327 | 0.790,0149 | 0.095,0031 | 119,4 |
| 8 | 350 | 25828 | 0.660,0153 | 0.070,0020 | 132,3 |
| 10 | 585 | 41680 | 0.540,0094 | 0.058,0010 | 139,1 |
| 12 | 890 | 57420 | 0.460,0048 | 0.052,0006 | 140,1 |
| 13 | 1085 | 65287 | 0.420,0045 | 0.050,0006 | 139,1 |
| \( N = 16 \) |
| 2 | 33 | 516358 | 0.960,0245 | 0.326,0079 | 56,14 |
| 4 | 178 | 9129930 | 0.920,0171 | 0.137,0011 | 126,4 |
| 6 | 460 | 417914690 | 0.790,0154 | 0.084,0028 | 170,3 |
| 8 | 940 | 933844934 | 0.650,0102 | 0.062,0011 | 194,2 |
| 10 | 1,470 | 1621394502 | 0.530,0070 | 0.050,0006 | 206,1 |
| 12 | 2,254 | 2279123670 | 0.440,0031 | 0.043,0003 | 207,1 |
| 14 | 3,354 | 2907320556 | 0.380,0022 | 0.040,0003 | 203,3 |
| 15 | 3,868 | 3212030861 | 0.350,0022 | 0.039,0004 | 200,0 |
| \( N = 18 \) |
| 2 | 50 | 157918854 | 0.950,0291 | 0.307,0030 | 73,15 |
| 4 | 330 | 2626607056 | 0.920,0137 | 0.127,0081 | 174,9 |
| 6 | 994 | 14644118858 | 0.780,0155 | 0.076,0044 | 237,5 |
| 8 | 2,093 | 3540993772 | 0.640,0097 | 0.056,0012 | 272,3 |
| 10 | 3,619 | 62052120318 | 0.529,0071 | 0.044,0007 | 292,1 |
| 12 | 5,657 | 899742140011 | 0.430,0037 | 0.038,0003 | 297,7 |
| 14 | 8,352 | 116366501995 | 0.360,0023 | 0.034,0002 | 293,1 |
| 16 | 11,786 | 140687006222 | 0.320,0012 | 0.032,0001 | 285,1 |
| 17 | 13,795 | 15247304818 | 0.300,0009 | 0.032,0001 | 277,1 |

The x-axis. The case \( N = 14 \) is similar but is not reported here because it is much more noisy for \( K = 2 \) and \( 4 \) due to the small size of the graphs in these cases (see table 2).

One can see that the weights, i.e. the transition probabilities between neighboring basins are small. The distributions are far from uniform and, for both \( N = 16 \) and \( N = 18 \), the low \( K \) have longer tails. For high \( K \) the decay is faster. This seems to indicate that, on average, the transition probabilities are higher for low \( K \).

**Disparity** Fig. 8 depicts the disparity coefficient as defined in the previous section for \( N = 16, 18 \). An interesting observation is that the disparity (i.e. dishomogeneity) in the weights of a node’s outcoming links tends to decrease steadily with increasing \( K \). This reflects that for high \( K \) the transitions to other basins tend to become equally likely, which is another indication that the landscape, and thus its representative maxima network, becomes more random and difficult to search.

When \( K \) increases, the number of edges increases and the number of edges with a weight over a certain threshold increases too (see fig. 8). Therefore, for small \( K \), each node is connected with a small number of nodes each with a relative high weight. On the other hand, for large \( K \), the weights become more homogeneous in the neighbourhood, that is, for each node, all the neighboring basins are at similar distance.

If we suppose that edges with higher weights are likely to be connected to nodes with larger basins (an intuition that we need to confirm in future work). Then, as the larger basins tend to have higher fitness (see Fig. 3), the path to higher fitness values would be easier to find for lower \( K \) than for larger \( K \).

**Boundary of basins.** Fig. 6 shows the averages, over all the nodes in the network, of the weights \( w_{ij} \) (i.e. the probabilities of remaining in the same basin after a bit-flip mutation). Notice that the weights \( w_{ij} \) are much higher when compared to those \( w_{ij} \) with \( j \neq i \) (see Fig. 5). In particular, for \( K = 2, 50\% \) of the random bit-flip mutations will produce a solution within the same basin of attraction. These average probabilities of remaining within the same basin, are above 12% for the higher values of \( K \). Notice that the averages are nearly the same regardless the value of \( N \), but decrease with the epistatic parameter \( K \).

The exploration of new basins with the random bit-flip mutation seems to be, therefore, easier for large \( K \) than for low \( K \). But, as the number of basins increases, and the fitness correlation between neighboring solutions decreases...
with increasing $K$, it becomes harder to find the global maxima for large $K$. This result suggests that the dynamic of stochastic local search algorithms on $NK$ landscapes with large $K$ is different than that with lower values of $K$, with the former engaging in more random exploration of basins.

The boundary of a basin of attraction can be defined as the set of configurations within a basin that have at least one neighbor's solution in another basin. Conversely, the interior of a basin is composed by the configurations that have all their neighbors in the same basin. Table 3 gives the average number of configurations in the interior of basins (this statistic is computed on 30 independent landscapes). Notice that the size of the basins’ interior is below 1% (except for $N = 14$, $K = 2$). Surprisingly, the size of the basins’ boundaries is nearly the same as the size of the basins themselves. Therefore, the probability of having a neighboring solution in the same basin is high, but nearly all the solutions have a neighbor solution in another basin. Thus, the interior basins seem to be “hollow”, a picture which is far from the smooth standard representation of landscapes in 2D with real variables where the basins of attraction are visualized as real mountains.

Conclusions

We have proposed a new characterization of combinatorial fitness landscapes using the family of $NK$ landscapes as an example. We have used an extension of the concept of inherent networks proposed for energy surfaces in order to abstract and simplify the landscape description. In our case the inherent network is the graph where the nodes are all the local maxima and the edges accounts for transition probabilities (using the bit-flip operator) between the local maxima basins of attraction. We have exhaustively obtained these graphs for $N = \{14, 16, 18\}$, and for all even values of $K$, plus $K = N - 1$, and conducted a network analysis on them. Our guiding motivation has been to relate the statistical properties of these networks, to the search diffic-
Table 3: Average (on 30 independent landscapes for each $N$ and $K$) of the mean sizes of the basins interiors.

| K  | $N = 14$ | $N = 16$ | $N = 18$ |
|----|----------|----------|----------|
| 2  | 0.01167 | 0.00500 | 0.00280 |
| 4  | 0.00250 | 0.00120 | 0.00060 |
| 6  | 0.00290 | 0.00140 | 0.00070 |
| 8  | 0.00430 | 0.00220 | 0.00110 |
| 10 | 0.00550 | 0.00310 | 0.00180 |
| 12 | 0.00610 | 0.00400 | 0.00250 |
| 13 | 0.00590 | 0.00450 | 0.00310 |
| 14 | 0.00500 | 0.00440 | 0.00350 |
| 15 |          |          | 0.00340 |
| 16 |          |          |          |
| 17 |          |          |          |

The distribution of the basin sizes is approximately exponential for all $N$ and $K$, but the basin sizes are larger for low $K$. The disparity coefficients reflect that for high $K$ the transition to other basins tend to become equally likely, which is an indication of the randomness of the landscape.

The construction of the maxima networks requires the determination of the basins of attraction of the corresponding landscapes. We have thus also described the nature of the basins, and found that the size of the basin corresponding to the global maximum becomes smaller with increasing $K$. The distribution of the basin sizes is approximately exponential for all $N$ and $K$, but the basin sizes are larger for low $K$, another indirect indication of the increasing randomness and difficulty of the landscapes when $K$ becomes large. Furthermore, there is a strong positive correlation between the basin size of maxima and their degrees. Finally, we found that the size of the basins boundaries is roughly the same as the size of basins themselves. Therefore, nearly all the configurations in a given basin have a neighbor solution in another basin. This observation suggests a different landscape picture than the smooth standard representation of 2D landscapes where the basins of attraction are visualized as hilltops.

This study is our first attempt towards a topological and statistical characterization of combinatorial landscapes, from the point of view of complex networks analysis. Much remains to be done. The results should be confirmed for larger instances of $NK$ landscapes. This will require good sampling techniques, or theoretical studies since exhaustive sampling becomes quickly impractical. Other landscape types should also be examined, such as those containing neutrality, which are very common in real-world applications. Finally, the landscape statistical characterization is only a step towards implementing good methods for searching it. We thus hope that our results will help in designing or estimating efficient search techniques and operators.

Acknowledgements. Thanks to Professor Edmund Burke for useful advice and support. This work was partially funded by EPSRC grant number EP/D061571/1.

References

Barthélemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. (2005). Characterization and modeling of weighted networks. *Physica A*, 346:34–43.

Boese, K. D., Kahng, A. B., and Muddu, S. (1994). A new adaptive multi-start technique for combinatorial global optimizations. *Operations Research Letters*, 16:101–113.

Doye, J. P. K. (2002). The network topology of a potential energy landscape: a static scale-free network. *Phys. Rev. Lett.*, 88:238701.

Doye, J. P. K. and Massen, C. P. (2005). Characterizing the network topology of the energy landscapes of atomic clusters. *J. Chem. Phys.*, 122:084105.

Garnier, J. and Kallel, L. (2001). Efficiency of local search with multiple local optima. *SIAM Journal on Discrete Mathematics*, 15(1):122–141.

Kauffman, S. A. (1993). *The Origins of Order*. Oxford University Press, New York.

Merz, P. and Freisleben, B. (1998). Memetic algorithms and the fitness landscape of the graph bi-partitioning problem. In *Parallel Problem Solving from Nature V*, volume 1498 of *Lecture Notes in Computer Science*, pages 765–774. Springer-Verlag.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45:167–256.

Reeves, C. R. (1999). Landscapes, operators and heuristic search. *Annals of Operations Research*, 86:473–490.

G. Ochoa, M. Tomassini, and S. Vérel. A study of NK landscapes basins’ and local optima networks. In *Genetic and Evolutionary Computation Conference, GECCO 2008, Proceedings*. ACM, 2008.