

Quasi-elementary Landscapes and Superpositions of Elementary Landscapes

Darrell Whitley¹ and Francisco Chicano²

¹ Dept. of Computer Science, Colorado State University, Fort Collins CO, USA*

`whitley@cs.colostate.edu`

² Dept. de Lenguajes y Ciencias de la Computación, University of Málaga, Spain**

`chicano@lcc.uma.es`

Abstract. There exist local search landscapes where the evaluation function is an eigenfunction of the graph Laplacian that corresponds to the neighborhood structure of the search space. Problems that display this structure are called “Elementary Landscapes” and they have a number of special mathematical properties. The term “Quasi-elementary landscapes” is introduced to describe landscapes that are “almost” elementary; in quasi-elementary landscapes there exists some efficiently computed “correction” that captures those parts of the neighborhood structure that deviate from the normal structure found in elementary landscapes. The “shift” operator, as well as the “3-opt” operator for the Traveling Salesman Problem landscapes induce quasi-elementary landscapes. A local search neighborhood for the Maximal Clique problem is also quasi-elementary. Finally, we show that landscapes which are a superposition of elementary landscapes can be quasi-elementary in structure.

1 Introduction

Grover [4] originally observed that there exist neighborhoods for Traveling Salesman Problem (TSP), Graph Coloring, Min-Cut Graph Partitioning, Weight Partitioning, as well as Not-All-Equal-SAT that can be modeled using a wave equation borrowed from mathematical physics. Stadler [7] named this class of problems “elementary landscapes” and showed that if a landscape is elementary, the objective function is an eigenfunction of the Laplacian matrix that describes the connectivity of the neighborhood graph representing the search space. Whitley and Sutton developed a “component” based model of elementary landscapes that makes it easy to identify elementary landscapes [13]. In many cases, the

* This research was sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number FA9550-11-1-0088. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

** It was also partially funded by the Spanish Ministry of Science and Innovation and FEDER under contract TIN2008-06491-C04-01 (the M* project) and the Andalusian Government under contract P07-TIC-03044 (DIRICOM project).

components are weights in a cost matrix. In the case of pseudo-Boolean functions, the components can also be the weight coefficients of a polynomial form of the cost function. Let the set of components be denoted by C ; a solution x will also denote the subset of components that contribute to the evaluation of x so that the sum of the components that contribute to x is the same as the evaluation of x , denoted by the evaluation function $f(x)$. Finally, let $C - x$ denote the subset of components that do not contribute to the evaluation of solution x . Note that the sum of the components in $C - x$ is computed by $(\sum_{w \in C} w) - f(x)$. Another precondition for a landscape to be elementary is:

$$\bar{f} = p_3 \sum_{w \in C} w,$$

where p_3 is the frequency of appearance of any component $w \in C$ in a random solution x .

We will denote a landscape as a triple (X, N, f) where f is the evaluation function $f : X \rightarrow \mathbb{R}$, the set of solutions X represents the discrete domain of f and $N(x)$ is the neighborhood operator that defines adjacency between elements $x \in X$ under some local search neighborhood. N can also be expressed in the form of an adjacency matrix \mathbf{A} . The elements of \mathbf{A} are such that $\mathbf{A}_{x,y} = 1$ if $y \in N(x)$ and $\mathbf{A}_{x,y} = 0$ otherwise. When a neighborhood is regular, the *Laplacian operator* is $\Delta = \mathbf{A} - d\mathbf{I}$ and it acts as a type of difference operator on the fitness function f such that:

$$\Delta f(x) = \sum_{y \in N(x)} (f(y) - f(x)).$$

A landscape is said to be *elementary* if f is an eigenvector of $-\Delta$ up to a constant, formally: $-\Delta f = k(f - b)$ for a constant b and an eigenvalue k of $-\Delta$. As a direct consequence, we can compute the neighborhood average in an elementary landscape as follows:

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) + \frac{k}{d}(\bar{f} - f(x)), \tag{1}$$

where $d = |N(x)|$ is the neighborhood size, which we assume is the same for all the solutions. Often, this same result can be expressed another way.

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1 f(x) + p_2 \left(\sum_{w \in C} w - f(x) \right), \tag{2}$$

where $p_1 = \alpha/d$ is the (sampling) rate at which “components” that contribute to $f(x)$ are removed from solution x to create a neighboring solution $y \in N(x)$, and $p_2 = \beta/d$ is the rate at which components in the set $C - x$ are sampled to create a neighboring solution $y \in N(x)$. Said in another way, in order to build

all the solutions in the neighborhood each component in x has been removed α times and each component in $C - x$ have been added β times. By simple algebra,

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1 f(x) + p_2 \left(\sum_{w \in C} w - f(x) \right) = f(x) + \frac{k}{d}(\bar{f} - f(x)), \quad (3)$$

where $k = \alpha + \beta$, $\bar{f}/p_3 = \sum_{w \in C} w$ and $p_3 = \beta/(\alpha + \beta)$ [12,13,14].

It should also be noted that some landscapes that are not elementary can nevertheless be expressed as a superposition of a small number of elementary landscapes. For example, a MAX-3SAT landscape under the Hamming-1 neighborhood is not elementary, but it can be re-expressed as the sum of three functions. Let $M(x)$ denote the MAX-3SAT evaluation function for a Boolean string x ; there exists functions f_1, f_2 and f_3 such that the landscapes of f_1, f_2 and f_3 are elementary, and

$$M(x) = f_1(x) + f_2(x) + f_3(x).$$

This makes it possible to compute averages over the Hamming-1 neighborhood. Using Walsh analysis it is even possible to compute higher order statistical moments (variance, skew, kurtosis) in polynomial time over arbitrary Hamming balls in the landscape, even over regions that are exponentially large [9]. The same method can be applied to all k -bounded pseudo-Boolean functions, including NK-Landscapes [6].

Thus, there is a great deal that we can potentially compute about local search landscapes that is not being utilized by search algorithms. We continue to find new ways to model new problems using elementary landscape theory, and we continue to find new ways to compute statistical information even more efficiently. Quasi-elementary landscapes are another step in this direction. In the current paper we present new results for the 3-opt move operator for the Traveling Salesman Problem, as well as new more general results for a Max-Clique neighborhood. We also establish a connection between problems that are a superposition of elementary landscapes and the concept of quasi-elementary landscapes.

1.1 Quasi-Elementary Landscapes

Whitley [11] introduced the term *quasi-elementary landscape* to describe a landscape and neighborhood structure where a variant of Grover’s wave equation can be used to compute the neighborhood average by adding a correction to the usual wave equation. Thus, a landscape is quasi-elementary if:

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1 f(x) + p_2 (\bar{f}/p_3 - f(x)) + g(x). \quad (4)$$

We will refer to g as an auxiliary function. To be quasi-elementary it is critical that the computational complexity of $g(x)$ be less than the computation complexity of enumerating and evaluating the neighbors of solution x . Like the cost function $f(x)$, the auxiliary function $g(x)$ can sample from the set of components

(e.g., the cost matrix) and can compute a “correction” relative to solution x to account for the fact that the landscape is not elementary.

In principle, one might allow $g(x)$ to be computed as the sum of multiple subfunctions. We know, for example, that MAX-3SAT is a superposition of 3 elementary landscapes. In the current paper we will show the average of the neighborhood for MAX-3SAT can be computed using $g(x) = a_1f_1(x) + a_2f_2(x)$ where functions f_1 and f_2 are elementary landscapes, but $g(x)$ is not elementary. It is again critical that the complexity of $g(x)$ must be less than the complexity of enumerating and evaluating the neighbors in $N(x)$.

In some cases, we can provide additional information about the function $g(x)$. Assume that the set of components can be broken into 3 sets: $f(x)$, $f'(x)$ and $\sum_{w \in C} w - f(x) - f'(x)$ where $f'(x)$ identifies (and sums) a set of components relative to x where those components that contribute to $f'(x)$ are sampled at a different rate relative to the other two subsets of components. Assume that

$$g(x) = p_4f'(x) - p_2f'(x),$$

where p_4 is the new sampling rate for the components that contribute to $f'(x)$. Assume the complexity of computing $f'(x)$ is no greater than the complexity of computing $f(x)$. We then obtain the following result establishing a quasi-elementary landscape.

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1f(x) + p_4f'(x) + p_2 \left(\frac{\bar{f}}{p_3} - f(x) - f'(x) \right). \quad (5)$$

We use this model to show that the “shift” operator and the “3-opt” operator for the Traveling Salesman Problem induce a quasi-elementary landscape. We can also use the model to show that a local search algorithm for the Maximal Clique problem also induces a quasi-elementary landscape; this same local search algorithm can also be used to find densely connected subgraphs in larger graphs.

Finally, we can also construct a quasi-elementary landscape that samples the functions $f(x)$ and $f'(x)$ and the constant \bar{f} . Let $s_1/d = (p_1 + p_2)$, $s_2/d = (p_2 - p_4)$ and $s_3/d = (p_2/p_3)$. We will say that a landscape is quasi-elementary if:

$$\text{Avg}_{y \in N(x)}(f(y)) = \left(1 - \frac{s_1}{d}\right) f(x) - \frac{s_2}{d} f'(x) + \frac{s_3}{d} \bar{f}. \quad (6)$$

In this form, we can show that landscapes which are a superposition of elementary landscapes are in fact quasi-elementary landscapes if $f'(x)$ can be efficiently computed. In the case of a superposition of two elementary landscapes we might reasonably expect the complexity of f' to be no greater than the complexity of f since f' is a subfunction that can be use to compute f . Quasi-elementary landscapes can also result from a superposition of more than two elementary landscapes. Assume we have a superposition of elementary landscapes where $f = f_1(x) + f_2(x) + f_3(x)$. Let $g(x) = f_1(x) + f_2(x)$. For some classes of problems one can still prove that g has the same complexity as evaluating $f(x)$.

2 Examples of Quasi-Elementary Landscape

We will first look at examples of quasi-elementary landscapes which can be described using the following equation.

$$\text{Avg}(f(y)) = f(x) - p_1 f(x) + p_4 f'(x) + p_2 (\bar{f}/p_3 - f(x) - f'(x)). \quad (7)$$

$y \in N(x)$

2.1 The Shift Operator for the TSP

We will assume a permutation representation is used. A shift operator works by deleting one vertex from the permutation, then that vertex is re-inserted at every other possible position in the permutation. When done at every possible position, this yields some duplicate neighbors. These duplicates can be eliminated by 1) shifting the vertex to be deleted to the beginning of the permutation, and 2) doing insertion to the next $n - 3$ possible positions. Therefore, the neighborhood size is $n(n - 3)$.

One might assume that the “shift” operator is not a commonly used TSP operator. However, the “shift” operator can also be modeled as a special 3-opt move where one of the segments is a single city. If the tour is then broken into a segment of 1 city, and then broken into two segments of 2 or more cities, reversing the two longer segments exactly yields a move under the “shift” operator.

Let $f'(x)$ denote an auxiliary function to $f(x)$. When a vertex is deleted in solution x the deletion removed 2 edges, and introduces 1 new edge: $f'(x)$ ignores the deleted edges, but counts the cost associated with the new edges. Since n vertices are deleted, $f'(x)$ is the sum of the n new edges. For example, consider the tour: 1 2 3 4 5. Then the cost function $f(x) = w_{1,2} + w_{2,3} + w_{3,4} + w_{4,5} + w_{5,1}$ while the auxiliary function $f'(x) = w_{1,3} + w_{2,4} + w_{3,5} + w_{4,1} + w_{5,2}$. In the case of the “shift” operator, the edges in $f'(x)$ appear with greater frequency in the neighborhood $N(x)$.

When a vertex is deleted and reinserted, this cuts 3 edges: it cuts the edge to the left and to the right of the deleted vertex, and it cuts the edge where the deleted vertex is reinserted. Because every vertex will be deleted and it will be inserted into all possible (non-duplicate) positions all the edges in x will be removed at the same frequency. We can group the neighbors into subsets of size $n - 3$ where the same vertex is deleted from x , but the deleted vertex is reinserted into all feasible positions. In this subset of $n - 3$ neighbors, the edge to the right of the deleted vertex is reinserted once (e.g., if AB represent the two consecutive vertices in the tour and A is deleted and re-inserted after B, the edge (A, B) is recovered). This happens 1 time for every edge in solution x across all neighbors.

We can calculate the rate at which edges are removed from x by assuming the symmetry case as a baseline, then correct for the 1 neighbor where a specific edge from x is reinserted. This means that the rate with which edges are removed from x is given by

$$-\frac{3}{n}f(x) + \frac{1}{n(n-3)}f(x) = \frac{-3n+10}{n(n-3)}f(x).$$

The calculation of \bar{f} is neighborhood independent; therefore we can use the calculation associated with the 2-opt neighborhood:

$$\sum_{w \in C} w - f(x) = \left(\frac{\beta + \alpha}{\beta} \bar{f} \right) - f(x) = \left(\frac{n-1}{2} \bar{f} \right) - f(x).$$

We next consider the edges in the set $C - x$. All of the edges are sampled in a symmetric fashion, but a subset of edges receive additional samples.

The symmetric case can be described as follows. New edges are created by inserting the deleted vertex in a new position. Since the insertion occurs in every position that does not produce a redundant neighbor, this case is symmetric. This can happen in only 4 ways: vertex P has been deleted and it is inserted before and after Q ; or vertex Q is deleted and it is inserted before and after vertex P . Thus, the symmetric sampling rate over the entire neighborhood is $4/(n(n-3))$.

The non-symmetric case derives from the fact that when a vertex is deleted, the deletion also creates a new edge. Furthermore, the *same* edge is created $n-3$ times for each of the $n-3$ cases where the *same* vertex is deleted. However, one of these $n-3$ cases is also one of the symmetric cases previously counted. Removing this one case, there remains $n-4$ cases where the same edge is created when the same vertex is deleted.

The set of edges that are sampled an additional $n-4$ times can be found starting the with current solution x and then deleting each vertex in x to create a permutation (circuit) of $n-1$ vertices. As each vertex is deleted, one new edge is created: thus, there are n new edges that are created. Let $f'(x)$ be the sum of the n new edges that are created by deleting each vertex one at a time.

Note that $|C - x| = n(n-3)/2$. Each of the n edges that contributes to $f'(x)$ appears $n-4$ times across the entire neighborhood. Therefore, the symmetric and non-symmetric sampling from $C - x$ is given by

$$\frac{4}{n(n-3)} \left(\frac{n-1}{2} \bar{f} - f(x) \right) + \frac{n-4}{n(n-3)} f'(x).$$

Using the sample rate from x and $C - x$ yields the combined effect. Note that the operator is only well-defined when $n > 3$. Therefore:

$$\text{Avg}(f(y)) = f(x) - \frac{3n-10}{n(n-3)} f(x) + \frac{4}{n(n-3)} \left(\frac{n-1}{2} \bar{f} - f(x) \right) + \frac{n-4}{n(n-3)} f'(x).$$

This can be rearranged to yield:

$$\text{Avg}(f(y)) = f(x) - \frac{3n-10}{n(n-3)} f(x) + \frac{n}{n(n-3)} f'(x) + \frac{4}{n(n-3)} \left(\frac{n-1}{2} \bar{f} - f(x) - f'(x) \right).$$

While computing $f'(x)$ with no prior knowledge requires $O(n)$ time, if we currently know the evaluation of $f(x)$ and $f'(x)$ and we move to a point y such that $y \in N(x)$, then both $f(y)$ and $f'(y)$ can be computed as a partial update to $f(x)$

and $f'(x)$ respectively, and both partial updates can be computed in constant time. Thus, the resulting landscape is quasi-elementary.

In special cases (usually when n is small) the quasi-elementary landscape is actually elementary. For example, when $n = 4$ we find that

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) + \frac{3}{2}(\bar{f} - f(x)).$$

2.2 The 3-opt Quasi-Elementary Landscape

The next example is a landscape for the classic 3-opt operator. This version of 3-opt does not include the “shift” operator and all segments must be of length 2 or greater. However, we will find that the results presented in this paper can be combined to characterize a more general 3-opt neighborhood that allows one of the segments to include the single city (shift operator) case.

Stattenberger et al. [8] give a general formula for counting the number of ways that a Hamiltonian Circuit can be cut into k segments corresponding to those used by a Lin-Kernighan k -opt operator. For $k = 3$ this quantity is $n(n - 4)(n - 5)/3!$. In principle, one could then reconfigure the tour by reversing one segment, two segments, or all three segments. However, note that reversing one segment results in a 2-opt move because the two segments that are not reversed can be concatenated into one segment. Thus, there are four patterns of reversal where either two segments are reversed or all three segments are reversed. Therefore:

$$d = 4n(n - 4)(n - 5)/3! = 2n(n - 4)(n - 5)/3.$$

In the following we will use $f(x)$ and $f'(x)$ with the same meaning as in the case of the “shift” neighborhood. However, the values of d , p_1 , p_2 , p_3 and p_4 could change. We have already presented the new value of d . Next we will present the new values of the p_i constants. We are searching for an expression of the following form:

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1 f(x) + p_4 f'(x) + p_2 \left(\sum_{w \in C} w - f(x) - f'(x) \right).$$

In a 3-opt move, exactly 3 edges are removed from the current solution. Thus, $p_1 = 3/n$ as in the case of the “shift” neighborhood. Let us now consider the edges in $C - x$ that are included in a new neighbor.

Let P and Q be two cities that are distance 2 apart in the current solution x . In order to bring together these 2 cities (and include the edge $e_{P,Q} \in C - x$ in the neighbor) a segment of length 2 must be reversed. There are two ways this segment of length two can be chosen. It can be chosen to include the first city P , which is then moved adjacent to the second city when the segment is reversed, or it can be chosen to include the second city Q which is moved adjacent to the first city when the segment is reversed. The location of the third cut does not matter, and there are $n - 5$ possible locations for the third cut to occur. Two

segments must be reversed to be a legal 3-opt move: one must be the segment of length 2, the other must be the segment that does not contain either P or Q . Three segments cannot be reversed. Thus, there are $2(n - 5)$ ways to segment the tour to yield the desired result, and there is only one reversal pattern in each case that reverses two segments to yield the desired result (see Figure 1). The sum of the weights of the edges of cities that are distant 2 apart in x is exactly $f'(x)$. Thus, $f'(x)$ must be summed $2(n - 5)$ times in the whole neighborhood, yielding $p_4 = 2(n - 5)/d$. Note that the auxiliary function is *exactly* the same as the one used for the “shift” operator. If the tour is: 1 2 3 4 5 then the auxiliary function is $f'(x) = w_{1,3} + w_{2,4} + w_{3,5} + w_{4,1} + w_{5,2}$. This means the results for the two neighborhoods can be easily combined.

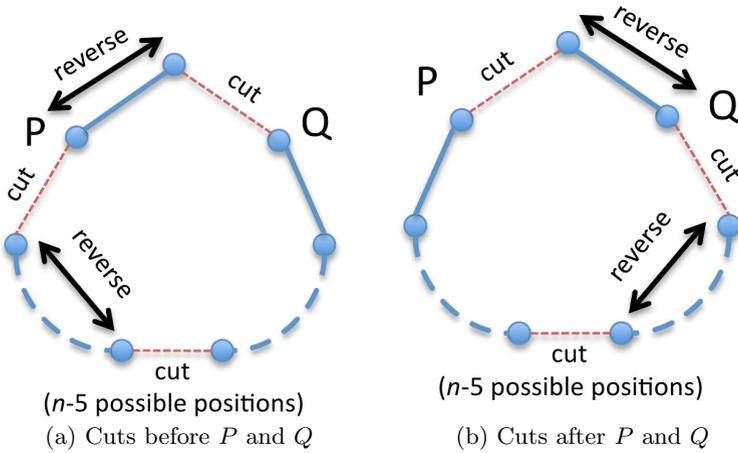


Fig. 1. The two ways cuts can be placed to bring cities P and Q together if they are distance 2 apart

We next compute the value of p_2 . Let P and Q be two cities in x which are distance $l > 2$ apart. We pick $l \leq n/2$ so we consider the shortest path between P and Q in the tour. There are four ways cuts can be placed to bring P and Q together:

- Before P and before Q ($n - 6$ other cuts are possible with only one reversal pattern). See Figure 2(a).
- After P and after Q ($n - 6$ other cuts are possible with only one reversal pattern). See Figure 2(b).
- After P and before Q ($n - l - 2$ cuts are possible with 2 reversal patterns). See Figure 2(c).
- Before P and after Q ($l - 2$ cuts are possible with 2 reversal patterns). See Figure 2(d).

Therefore $p_2 = (n - 6)/d + (n - 6)/d + 2(n - l - 2 + l - 2)/d = 4(n - 5)/d$.

Finally, by substitution:

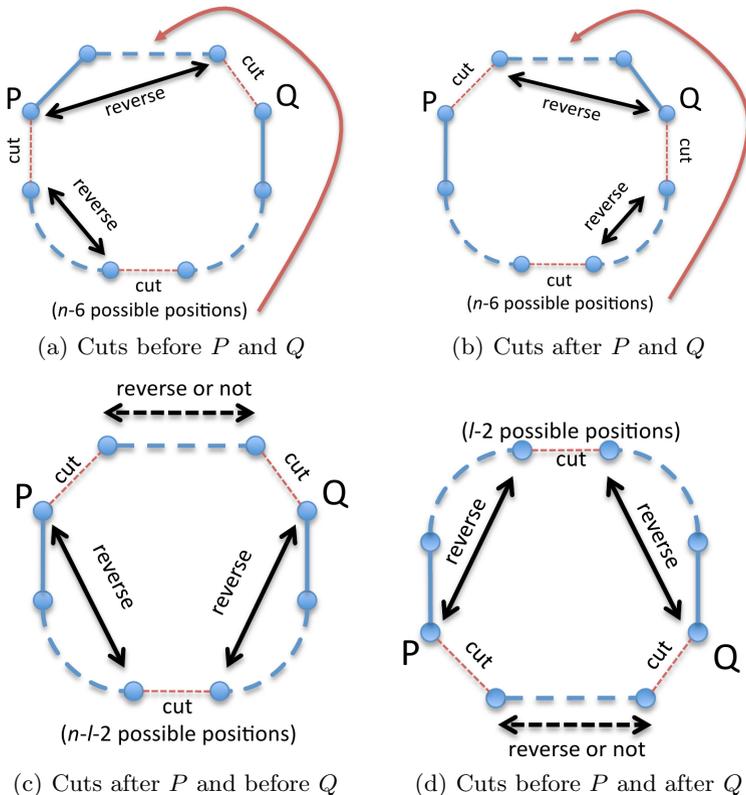


Fig. 2. The four ways cuts can be placed to bring cities P and Q together in a neighboring solution

$$\begin{aligned}
 \text{Avg}(f(y)) &= f(x) - \frac{3}{n}f(x) + \frac{2(n-5)}{d}f'(x) + \frac{4(n-5)}{d} \left(\sum_{w \in C} w - f(x) - f'(x) \right) \\
 &= f(x) - \frac{3}{n}f(x) + \frac{2}{2n(n-4)/3}f'(x) + \frac{4}{2n(n-4)/3} \left(\sum_{w \in C} w - f(x) - f'(x) \right) \\
 &= f(x) - \frac{3}{n}f(x) + \frac{3}{n(n-4)}f'(x) + \frac{6}{n(n-4)} \left(\sum_{w \in C} w - f(x) - f'(x) \right). \tag{8}
 \end{aligned}$$

Recall that computing both $f(y)$ and $f'(y)$ can be done as a partial update of $f(x)$ and $f'(x)$, respectively.

While we could combine the results for the “shift” neighborhood and this 3-opt neighborhood result, it is more useful to leave them separated, since it provides information about which subset of the larger more general 3-opt neighborhood yields better moves on average. It is also notable that “shift” increases the relative sampling of $f'(x)$ while this form of 3-opt decreases the relative sampling of $f'(x)$.

2.3 The Maximal Clique Problem

Katayama *et al.* [5] propose a variable neighborhood “k-opt” move operator¹ for the maximal clique problem. Katayama *et al.* report that this operator is competitive with other heuristic search methods for generating solutions to the maximal clique problem. While the operator has been applied to the maximal clique problem, it can also be used to search for subgraphs with maximal density independent of whether the subgraph is a clique or not.

The operator breaks a graph G with vertices V and edges E into two subgraphs, one with a set of vertices denoted by Z , and the other with the remaining vertices, $V - Z$. Under the “k-opt” move operator, k vertices in Z are removed and k vertices from $V - Z$ are added to Z . We will say that an edge $e_{i,j}$ belongs to Z if vertices v_i and v_j are in the subgraph Z . All subsets of k vertices that are not currently in Z are moved into subset Z ; thus all edges that do not belong to Z come into Z .

We must also compute \bar{f} as a uniform sample over the cost components. Assume we are attempting to maximize the number of edges in subgraph $Z \subset V$. Let $f(Z)$ count the number of edges of E in Z . Let $r = |Z|$, $q = |V - Z|$ and $n = |V|$. Using counting arguments, one can prove that the average of f over all possible assignments of vertices to Z is given by:

$$\bar{f} = \frac{r(r-1)}{n(n-1)} \sum_{w \in C} w. \quad (9)$$

To have an elementary landscape, we need to divide “components” into those that are in the solution, and those that are not in the solution. And each set needs to be sampled at a uniform frequency corresponding to p_1 or p_2 . But when counting (and maximizing) the number of edges in Z and the number of edges not in Z , we can classify edges into 3 different types that are uniformly sampled.

- Type 1: Edges that belong to Z and contribute to $f(Z)$.
- Type 2: Edges not in Z that connect two vertices not in Z .
- Type 3: Edges not in Z that connect a vertex in Z to a vertex in $(V - Z)$.

Therefore we can compute p_1 and p_3 , but there are two sampling rates instead of one for the edges normally accounted for p_2 . As in the previous examples we will denote the two different sampling rates by p_2 and p_4 .

We should also note here that Katayama *et al.* limit their neighborhood moves to those involving only the Type 1 and Type 2 edges. If a vertex in $(V - Z)$ did not connect to a vertex in Z , then moving that vertex was not considered in the Katayama implementation. However, we will analysis the full neighborhood. To do this we generalize an early result by Whitley [11].

Theorem 1. *The “k-opt” operator for maximizing the density of edges in a subgraph Z of graph G induces a quasi-elementary landscape. Let $r = |Z|$, $q =$*

¹ This is not the k -opt operator used for the Traveling Salesman Problem.

$|V - Z|$ and $n = |V| = r + q$. For $r > k$ and $q > k$ the neighborhood average for the “ k -opt” operator is given by:

$$\begin{aligned} \text{Avg}_{Y \in N(Z)} (f(Y)) &= f(Z) - \frac{k(r + q + 2q^2r + k(q - q^2 - r + r^2) - (r + q)^2)}{q(q - 1)r(r - 1)} f(Z) \\ &\quad + \frac{k(qr + k(1 - r - q))}{rq(q - 1)} f'(Z) + \frac{k(k - 1)n(n - 1)}{r(r - 1)q(q - 1)} \bar{f}, \end{aligned} \tag{10}$$

where the function $f'(Z)$ is the weighted sum of the edges having one vertex in Z and the other in $V - Z$.

Proof. We already know that $p_3 = \frac{r(r-1)}{n(n-1)}$ by Eq. (9).

We need to count 3 kinds of edges:

1. $f(Z)$ is the weighted sum of edges of E that are included in Z .
2. $f'(Z)$ is the weighted sum of edges with one vertex in Z and one in $V - Z$.
3. $\sum_{w \in C} w$ is the total weighted sum of edges.

Putting these together we obtain

$$\text{Avg}_{Y \in N(Z)} (f(Y)) = f(Z) - p_1 f(Z) + p_4 f'(Z) + p_2 \left(\sum_{w \in C} w - f(Z) - f'(Z) \right). \tag{11}$$

We first compute p_1 , the probability that edge $e_{i,j}$ moves out in Z . If two vertices are randomly drawn from Z for exchange, there are $(r - k)/r$ ways the first vertex can stay in Z and $(r - k - 1)/(r - 1)$ that the second vertex can stay in Z . The probability of an edge that is currently in Z moving out of Z is

$$p_1 = 1 - \frac{(r - k)(r - k - 1)}{r(r - 1)}.$$

We define p_2 to be the probability that an edge “contained” within $V - Z$ moves to Z . There are k/q ways to select the first vertex, and $(k - 1)/(q - 1)$ ways that the second vertex can be selected. Thus,

$$p_2 = \frac{k(k - 1)}{q(q - 1)}.$$

Finally, we will compute p_4 . Consider an edge $e_{i,j}$ such that $v_i \in Z$ and $v_j \in V - Z$. Then p_4 corresponds to the probability that v_i stays in Z and v_j moves to Z when any random neighbor is considered.

$$p_4 = \frac{(r - k)}{r} \cdot \frac{k}{q} = \frac{k(r - k)}{rq}.$$

Substituting into (11) we obtain:

$$\begin{aligned} \text{Avg}_{Y \in N(Z)}(f(Y)) &= f(Z) - \left(1 - \frac{(r-k)(r-k-1)}{r(r-1)}\right) f(Z) + \frac{k(r-k)}{rq} f'(Z) \\ &+ \frac{k(k-1)}{q(q-1)} \left(\frac{n(n-1)}{r(r-1)} \bar{f} - f(Z) - f'(Z)\right), \end{aligned} \tag{12}$$

which reduces to (10).

Computing both $f(Z)$ and $f'(Z)$ can be done by enumerating the vertices in Z and checking the edges that are incident on vertices in Z . Therefore $f'(Z)$ can be computed as a side-effect of computing $f(Z)$. And for $Y \in N(Z)$, both $f(Y)$ and $f'(Y)$ can be computed as a partial update to $f(Z)$ and $f'(Z)$ respectively. Therefore computing $f'(Z)$ has complexity less than or equal to computing $f(Z)$ and the maximal clique problem is a quasi-elementary landscape.

These results could be expressed with respect to the neighborhood size. The total size of the neighborhood is given by

$$d = \binom{r}{k} \binom{q}{k}.$$

When $p_4 = p_2$ and $k \geq 2$ the landscape becomes elementary. This happens when $r = 3, q = 4$ and $k = 2$ for example. But this is a special case, and this does not happen in general.

3 Superpositions of Two Elementary Landscapes

Assume we have a landscape that is a superposition of two elementary landscapes, such that

$$f(x) = f_1(x) + f_2(x).$$

In this case, the evaluation function is not an elementary landscape, but f_1 and f_2 are elementary landscapes. A number of problems have been shown to be a superposition of two elementary landscapes. These include the asymmetric frequency assignment problem [2,12] and all pseudo-Boolean functions with 2-bounded complexity, like MAX-2SAT, Unconstrained Quadratic Optimization (UQO) [3], the Subset Sum [1] as well as all the NK-Landscapes when $K=1$ [9].

This means that

$$\begin{aligned} \text{Avg}_{y \in N(x)}(f(y)) &= \text{Avg}_{y \in N(x)}\{f_1(y)\} + \text{Avg}_{y \in N(x)}\{f_2(y)\} \\ &= f_1(x) + \frac{k_1}{d} (\bar{f}_1 - f_1(x)) + f_2(x) + \frac{k_2}{d} (\bar{f}_2 - f_2(x)). \end{aligned} \tag{13}$$

Given that we know $f(x) = f_1(x) + f_2(x)$ and $f_2(x) = f(x) - f_1(x)$ we obtain:

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) + \frac{k_1}{d} (\bar{f}_1 - f_1(x)) + \frac{k_2}{d} (\bar{f}_2 - f_2(x)),$$

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - \frac{k_1}{d} f_1(x) - \frac{k_2}{d} (f(x) - f_1(x)) + \left\{ \frac{k_1}{d} \bar{f}_1 + \frac{k_2}{d} \bar{f}_2 \right\},$$

where $\frac{k_1}{d} \bar{f}_1 + \frac{k_2}{d} \bar{f}_2$ is a constant. Note that we can chose to eliminate either f_1 or f_2 and normally would select the simpler of the two functions to include in the computation. This means that every problem which is a superposition of two elementary landscapes is also a quasi-elementary landscape as long as the computational complexity of either f_1 or f_2 is less than the cost of enumerating the neighborhood. Generally, we would expect the computational cost of f_1 or f_2 to be less than the cost of computing the full evaluation function, $f = f_1 + f_2$.

We have previously shown in this paper that the “shift” operator and the “3-opt” operators for the TSP, as well as the Katayama’s “k-opt” operator for Max-Clique can be captured by an equation of the following form:

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) - p_1 f(x) + p_4 f'(x) + p_2 \left(\frac{\bar{f}}{p_3} - f(x) - f'(x) \right). \quad (14)$$

In all of these problems the auxiliary function $f'(x)$ is the linear sum of a subset of components drawn from the set $C - x$. This means that in these problems there are exactly 3 distinct sampling rates over component’s in the set C . Having this knowledge also makes it easier to search for a superposition of elementary landscapes, because it limits the number of equivalence classes that must be constructed when attempting to construct a superposition of elementary landscapes.

4 MAX-3SAT: A Superposition of Three Elementary Landscapes

Now assume that we have a superposition of three elementary landscapes: $f(x) = f_1(x) + f_2(x) + f_3(x)$. Then the average can be computed as:

$$\begin{aligned} \text{Avg}_{y \in N(x)}(f(y)) &= \text{Avg}_{y \in N(x)}\{f_1(y)\} + \text{Avg}_{y \in N(x)}\{f_2(y)\} + \text{Avg}_{y \in N(x)}\{f_3(y)\} \\ &= f_1(x) + \frac{k_1}{d} (\bar{f}_1 - f_1(x)) + f_2(x) + \frac{k_2}{d} (\bar{f}_2 - f_2(x)) \\ &\quad + f_3(x) + \frac{k_3}{d} (\bar{f}_3 - f_3(x)). \end{aligned} \quad (15)$$

But for MAX-3SAT, we can include \bar{f} in f_1 so that $\bar{f}_2 = \bar{f}_3 = 0$ and

$$\text{Avg}_{y \in N(x)}(f(y)) = f(x) + \frac{k_1}{d} (\bar{f}_1 - f_1(x)) - \frac{k_2}{d} (f(x) - f_1(x) - f_3(x)) - \frac{k_3}{d} f_3(x).$$

One can express f, f_1, f_2 and f_3 as Walsh functions. Assume f_1 captures the linear interactions, f_2 pairwise, and f_3 the order-3 interactions [10]. There is

only one 3-way interaction per clause. And there are only n linear terms, the number of variables of the instance. We remove f_2 because it is larger and less regular in structure than f_1 and f_3 . If there are n bits and m clauses, then the number of components needed to compute the average (with f_2 out of the picture) is at most $n + m + 1$ since there are at most m f_3 coefficients and n linear f_1 coefficients and $\bar{f} = \bar{f}_1$.

However, evaluations can also be done by partial evaluation. Let y_b be a neighbor of solution x generated by flipping bit b . It is then easy to prove that the cost of the partial evaluation is constant on average for $f(y)$, $f_1(y)$ and $f_3(y)$ given $f(x)$, $f_1(x)$ and $f_3(x)$. If y_b is a neighbor of x , only one Walsh coefficient (w_p) changes in f_1 . Thus, $f_1(y_b) = f_1(x) - 2(\psi_p(x)w_p)$. And on average, only a constant number of order-3 Walsh coefficients changes sign in f_3 ; the coefficients exactly map to those clauses that contain bit b . In expectation, a bit appears in a clause with probability $3/n$ and across all clauses a bit appears $3m/n = O(1)$ times. Thus, evaluating f_3 also has a partial update that is almost identical to the partial update for evaluating f .

Hence, $\text{Avg}(f(y))_{y \in N(x)}$ can also be computed in $O(1)$ time on average, and the MAX-3SAT landscape is quasi-elementary. This generalizes to all MAX-kSAT problems, as well as NK-Landscapes.

5 Conclusions

In this paper we present three examples of quasi-elementary landscapes: TSP with the “shift” and 3-opt neighborhoods and Maximal Clique with Katayama’s k -opt neighborhood. We also show that functions which are a superposition of elementary landscapes can also be quasi-elementary landscapes. A direct application of the concept of quasi-elementary landscapes is the generalization of the Grover’s wave equation to landscapes which are not elementary. In future work we plan to continue to explore the relationship between the elementary landscape decomposition of the problems and the quasi-elementary property.

References

1. Chicano, F., Whitley, D., Alba, E.: A methodology to find the elementary landscape decomposition of combinatorial optimization problems. *Evolutionary Computation* (2011)
2. Chicano, F., Whitley, D., Alba, E., Luna, F.: Elementary landscape decomposition of the frequency assignment problem. *Theoretical Computer Science* 412(43), 6002–6019 (2011)
3. Chicano, F., Alba, E.: Elementary landscape decomposition of the 0-1 unconstrained quadratic optimization. *Journal of Heuristics* xx, xx–xx (2011), doi:10.1007/s10732-011-9170-6
4. Grover, L.K.: Local search and the local structure of NP-complete problems. *Operations Research Letters* 12, 235–243 (1992)
5. Katayama, K., Hamamoto, A., Harihisa, H.: An effective local search for the maximum clique problem. *Information Processing Letters* 95(5), 503–511 (2005)

6. Kauffman, S., Levin, S.: Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* 128, 11–45 (1987)
7. Stadler, P.F.: Toward a theory of landscapes. In: López-Peña, R., Capovilla, R., García-Pelayo, R., Waelbroeck, H., Zertruche, F. (eds.) *Complex Systems and Binary Networks*, pp. 77–163. Springer (1995)
8. Stattenberger, G., Dankesreiter, M., Baumgartner, F., Schneider, J.J.: On the neighborhood structure of the traveling salesman problem generated by local search moves. *Journal of Statistical Physics* 129, 623–648 (2007)
9. Sutton, A., Whitley, D., Howe, A.: Computing the moments of k -bounded pseudo-boolean functions over hamming spheres of arbitrary radius in polynomial time. *Theoretical Computer Science* (2011), doi:10.1016/j.tcs.2011.02.006
10. Sutton, A.M., Howe, A.E., Whitley, L.D.: A Theoretical Analysis of the k -Satisfiability Search Space. In: Stützle, T., Birattari, M., Hoos, H.H. (eds.) *SLS 2009*. LNCS, vol. 5752, pp. 46–60. Springer, Heidelberg (2009)
11. Whitley, D.: Quasi-elementary landscapes. In: *Multi-Interdisciplinary Scheduling: Theory and Applications*, MISTA (2011)
12. Whitley, D., Chicano, F., Alba, E., Luna, F.: Elementary landscapes of frequency assignment problems. In: *GECCO 2010*, pp. 1409–1416. ACM Press (2010)
13. Whitley, D., Sutton, A.M.: Partial neighborhoods of elementary landscapes. In: *Proceedings of the Genetic and Evolutionary Computation Conference*, Montreal, Canada, pp. 381–388 (July 2009)
14. Whitley, L.D., Sutton, A.M., Howe, A.E.: Understanding elementary landscapes. In: *Proceedings of the Genetic and Evolutionary Computation Conference*, Atlanta, GA, pp. 585–592 (July 2008)