A Theoretical Analysis of the k-Satisfiability Search Space*

Andrew M. Sutton, Adele E. Howe, and L. Darrell Whitley

Department of Computer Science, Colorado State University,
Fort Collins CO, USA
{sutton,howe,whitley}@cs.colostate.edu

Abstract. Local search algorithms perform surprisingly well on the k-satisfiability (k-SAT) problem. However, few theoretical analyses of the k-SAT search space exist. In this paper we study the search space of the k-SAT problem and show that it can be analyzed by a decomposition. In particular, we prove that the objective function can be represented as a superposition of exactly k elementary landscapes. We show that this decomposition allows us to immediately compute the expectation of the objective function evaluated across neighboring points. We use this result to prove previously unknown bounds for local maxima and plateau width in the 3-SAT search space. We compute these bounds numerically for a number of instances and show that they are non-trivial across a large set of benchmarks.

1 Introduction

Local search methods for k-satisfiability (k-SAT) problems have received considerable attention in the AI search community. Though these methods are incomplete, they are usually able to quickly solve difficult problems that lie beyond the grasp of conventional complete solvers [1] and have been found to exhibit superior scaling behavior on soluble problems at the phase transition [2].

The behavior of local search algorithms closely depends on the underlying structure of the search space. A number of researchers have conducted empirical investigations on certain structural features of the k-SAT problem. Hoos and Stützle [3] introduced several metrics for measuring structure and presented an empirical examination of the characteristics of plateaus and their influence on the performance of local search. Clark et al. [4] studied the the relationship between problem hardness and the expected number of solutions on random problems. Frank et al. [5] analyzed the topology of the search space and experimentally probed the nature of local optima and plateaus. Yokoo [6] investigated the dependency of search cost on search space characteristics by studying how cost for local algorithms is related to the size of certain plateaus.

* This research was sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number FA9550-08-1-0422. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

T. Stützle, M. Birattari, and H.H. Hoos (Eds.): SLS 2009, LNCS 5752, pp. 46–60, 2009.

In this paper, we take an analytical view of the k-SAT search space by formalizing it as a *landscape* [7] which captures the relationship between the *objective function* associated with the problem and a *neighborhood operator*. We use the landscape formalism to analyze the search space of the k-SAT problem. We show that the search landscape can be decomposed into k *elementary* components. We prove that this decomposition provides an equation that gives the expectation of a random variable that models the objective function value of states in a given neighborhood. This quantity is equal to the average objective function value of the neighbors of a given state.

Furthermore, we use the decomposition to prove bounds for two prominent search space features: local maxima and plateaus. We show local maxima do not exist below a certain objective function value. Plateaus are regions of the search space consisting of states that are interconnected by a neighborhood operator and share an objective function value. Hoos and Stützle [3] define the *width* of a plateau P: the minimal length path between any state in P and one not in P. For many SAT instances, empirical results suggest that plateaus of width greater than one do not exist, or are at least very rare [3]. We prove there are regions of the search space that *cannot* contain plateaus of width greater than 1 and show empirically that these regions encompass the majority of the range of the objective function value. To our knowledge, there are no analytical results on the existence (or non-existence) of plateaus of particular width. Our results apply to local search on k-SAT and MAX-k-SAT where the count of unsatisfied clauses is the state evaluation function.

1.1 The Landscape Formalism

Before we specialize the discussion to k-SAT problems, we begin by introducing the landscape formalism. A combinatorial search problem is characterized as a finite but very large set X of states (complete candidate solutions) and an objective function $f : X \to \mathbb{R}$ that assigns a measure of value $f(x)$ to each state x. The objective of a search algorithm is to quickly locate a state $x^* \in X$ that extremizes f. Since f is a function over a discrete domain, we can characterize it as a vector $f \in \mathbb{R}^{|X|}$.

Local search algorithms perform local perturbations on states to move through the search space toward more promising regions. The space explored by such local methods thus requires additional structure by imposing a connectivity on X that consists of pairs of states that are separated by a *move*. We can define a second function on X denoted $N : X \to 2^X$ where $N(x)$ represents the set of all possible states that can be derived from x by applying the move operator exactly once. We refer to this set as the *neighborhood* of x. The tuple (X, N, f) is called the *landscape* of the combinatorial search problem and encompasses both the objective function values and the connectivity of states via the neighborhood.

We define the $|X| \times |X|$ *Markov transition matrix* $\mathbf{T}$

$$\mathbf{T}_{xy} = \begin{cases} \frac{1}{|N(x)|} & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

This matrix quantifies the transition probabilities between states on a random walk of the graph of the state space induced by the neighborhood operator. We can also view $\mathbf{T}$ as a linear operator that acts on an arbitrary vector $g \in \mathbb{R}^{|X|}$:

$$(\mathbf{T}g) = \begin{bmatrix} \frac{1}{|N(x_1)|} \sum_{z \in N(x_1)} g(z) \\ \vdots \\ \frac{1}{|N(x_{|X|})|} \sum_{z \in N(x_{|X|})} g(z) \end{bmatrix} \tag{1}$$

where x_i is the i^{th} element of X. Intuitively, $\mathbf{T}g$ is a discrete function where $\mathbf{T}g(x)$ gives the average value of g evaluated across the neighbors of the state x.

A landscape (X, N, f) is called *elementary* if the following equation is satisfied

$$\mathbf{T}f = \lambda f + \gamma \tag{2}$$

where both λ and γ are constants [8,7]. In other words, the objective function is an *eigenfunction* of the Markov transition matrix (up to an additive constant) corresponding to eigenvalue λ.

Several well-studied combinatorial problems along with natural neighborhood operators have been shown to satisfy the above equation (e.g., traveling salesman, graph coloring, not-all-equal satisfiability). Elementary landscapes possess a number of interesting properties. For example, Grover [8] has shown that no arbitrarily poor local optima can exist on an elementary landscape and that a solution with evaluation superior to the mean objective function value can be computed in polynomial time.

Landscapes that obey Equation (2) are called elementary because they behave as building blocks of more general combinatorial search landscapes. Provided that the neighborhood operator satisfies symmetry and regularity conditions, any arbitrary landscape can be represented as a linear combination of elementary landscapes [7]. We impose in this paper the following constraints.

1. $y \in N(x) \iff x \in N(y)$
2. $|N(x)| = |N(y)| = d; \quad \forall x, y \in X$

Most "natural" operators typically satisfy these constraints. The first constraint states all neighborhood relationships are symmetric, and the second asserts that all states have exactly d neighbors. Under these conditions $\mathbf{T}$ is a real symmetric $|X| \times |X|$ matrix and thus its $|X|$ eigenvectors $\{\phi_i\}$ with corresponding real eigenvalues λ_i form an orthonormal basis.

Thus we can represent an arbitrary function f in the eigenbasis $\{\phi_i\}$ as a linear combination.

$$f = \sum_{i=0}^{|X|-1} a_i \phi_i \tag{3}$$

Each ϕ_i is an eigenvector of $\mathbf{T}$. Note that each $a_i\phi_i$ can be considered again as a function $a_i\phi_i : X \to \mathbb{R}$. Each of these component functions satisfy Equation (2) and are thus elementary with respect to the neighborhood operator N.

In the general case, an arbitrary landscape f is represented by $|X|$ elementary constituents. Clearly, $|X|$ is exponential in the problem input size for landscapes of interest in this context. Thus this property is not obviously useful. However, in some interesting cases, it has been shown that the superposition is composed of a small number of elementary components. Examples are the asymmetric traveling salesman problem [9] and the quadratic assignment problem [10], both under traditional move operators.

1.2 The Neighborhood Expectation Value

We introduce a random variable that measures the objective function value of a neighbor selected uniformly at random. Later, we will use the expectation of this random variable in a simple probabilistic argument to prove the main results of the paper. Whitley et al. [11] studied elementary landscapes in the context of this random variable by connecting Equation (2) to the first moment of its distribution. In this section, we show this analysis can be easily extended to landscapes that are superpositions of elementary components.

Let $x \in X$ be an arbitrary state. Let $y \sim N(x)$ be an element drawn uniformly at random from the neighborhood of x, i.e., y is a random move using the operator defined by N. We define the random variable $Y = f(y)$ as the objective value of the neighboring state y.

Since y is selected uniformly at random, the expectation of Y is equivalent to the average of f evaluated over all of the neighbors of x.

$$\mathbb{E}[Y] = \frac{1}{d} \sum_{z \in N(x)} f(z)$$

If the objective function can be decomposed into a small number of components, the decomposition is useful in finding the expectation of Y. For example, suppose there are only $c+1$ nonzero coefficients $a_0, a_1, \ldots, a_c$ in the decomposition shown in Equation (3).

$$\begin{aligned}
\mathbb{E}[Y] &= \frac{1}{d} \sum_{z \in N(x)} f(z) \\
&= \mathbf{T} f(x) && \text{by Eq. (1)} \\
&= \mathbf{T} \left(\sum_{i=0}^{c} a_i \phi_i(x) \right) \\
&= \sum_{i=0}^{c} \lambda_i a_i \phi_i(x) && (4)
\end{aligned}$$

Therefore, given the $c + 1$ elementary components $a_i \phi_i$ and the corresponding eigenvalues λ_i we can immediately compute $\mathbb{E}[Y]$ without computing any elements of $N(x)$.

2 Decomposition of k-SAT

We now show that the k-SAT problem (and its optimization variant MAX-k-SAT) is decomposable into k elementary components. An instance of the k-SAT problem consists of a set of n Boolean variables $\{v_1, \ldots, v_n\}$ and a set of m clauses $\{c_1, \ldots, c_m\}$. Each clause is composed of exactly k *literals* in disjunction. The objective is to find a variable assignment that maximizes the number of satisfied clauses.

In this case, a state is a complete assignment to the n variables and can be characterized as a sequence of n bits $x = (x[1], x[2], \ldots x[n])$ where

$$x[b] = \begin{cases} 1 & \text{if and only if } v_b \text{ is true} \\ 0 & \text{if and only if } v_b \text{ is false} \end{cases}$$

The *state space* X is isomorphic to the set of all sequences $x \in \{0,1\}^n$.

The objective function $f : X \to \{0, \ldots, m\}$ simply counts the number of clauses satisfied under the assignment given by x. The most natural neighborhood is the Hamming neighborhood N where $N(x)$ is the set of n states y that differ from x in exactly one bit.

Since f can be taken as a function over bit strings of length n, a natural decomposition is given by the Walsh transform. In the general case, an arbitrary pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ can be represented as a linear combination of 2^n *Walsh functions* which we will define shortly. Rana et al. [12] showed that the k-SAT objective function can be tractably decomposed into a polynomial number of such functions. We will use this result to obtain a decomposition of the k-SAT objective function into elementary components.

Given two bit strings x and y of length n, we denote the inner product $\langle x, y\rangle$ as $\sum_{b=1}^{n} x[b]y[b]$. We define the i^{th} Walsh function $i \in \{0, \ldots, 2^n - 1\}$ as

$$\psi_i(x) = (-1)^{\langle i, x\rangle}$$

Here, the i that appears in the inner product of the exponent is taken to be the *bit string representation* of the index i, that is, the binary sequence of length n that corresponds to the integer i.

The objective function f can now be written as

$$f(x) = \sum_i w_i \psi_i(x) \tag{5}$$

where each Walsh coefficient w_i is the sum of contributions from each clause.

$$w_i = \sum_{j=1}^{m} w_{i,c_j}$$

where w_{i,c_j} is the contribution to w_i from clause c_j. This is defined as follows. Let $v(c_j)$ denote a bitstring of length n where

$$v(c_j)[b] = \begin{cases} 1 & \text{if variable } v_b \text{ appears in clause } c_j \\ 0 & \text{otherwise} \end{cases}$$

Similarly, let $u(c_j)$ be a bitstring of length n where

$$u(c_j)[b] = \begin{cases} 1 & \text{if variable } v_b \text{ appears } \textit{negated} \text{ in clause } c_j \\ 0 & \text{otherwise} \end{cases}$$

If x and y are bitstrings of length n, we say

$$x \subseteq y \iff (x[b] = 1 \implies y[b] = 1)$$

for $b = \{1, \ldots, n\}$. The contribution of clause c_j to Walsh coefficient w_i is

$$w_{i,c_j} = \begin{cases} 0 & \text{if } i \not\subseteq v(c_j) \\ \frac{2^k - 1}{2^k} & \text{if } i = 0 \\ -\frac{1}{2^k}\psi_i(u(c_j)) & \text{otherwise} \end{cases} \tag{6}$$

The *order* of a Walsh coefficient w_i is the number of ones in the bitstring representation of i. This can be denoted following our notation as $\langle i, i \rangle$. Note that the order of any nonzero Walsh coefficient is bounded by k: the number of variables that appear together in a clause. Rana et al. showed it is enough to specify $f(x)$ by computing the $O(2^k m)$ non-zero Walsh coefficients and computing the superposition in Equation (5). Since k is typically taken to be $O(1)$, all nonzero Walsh coefficients can be found in polynomial time.

Lemma 1. *The Walsh function ψ_i of order $\langle i, i \rangle = p$ is an eigenvector of the Markov transition matrix $\mathbf{T}$ with eigenvalue $\left(1 - \frac{2p}{n}\right)$*

Proof. Let x be an arbitrary state.

$$\mathbf{T}\psi_i(x) = \frac{1}{n} \sum_{z \in N(x)} \psi_i(z) \qquad \text{by Eq. (1)}$$

A Hamming neighbor $z \in N(x)$ differs from x in exactly one bit position b. By definition, $\psi_i(z) = (-1)^{\langle i,z \rangle}$. Consider $i[b]$, that is, the bit located at position b in the bitstring representation of i. If $i[b] = 0$ then $\langle i, z \rangle = \langle i, x \rangle$ and $\psi_i(z) = \psi_i(x)$. On the other hand, if $i[b] = 1$ then $|\langle i, z \rangle - \langle i, x \rangle| = 1$ and $\psi_i(z) = -\psi_i(x)$.

Since each Hamming neighbor differs from x in each of the n possible bit positions, there are p elements z of $N(x)$ that satisfy the second condition and $n - p$ that satisfy the first. Thus we have

$$\begin{aligned} \frac{1}{n} \sum_{z \in N(x)} \psi_i(z) &= \frac{1}{n} \left((n - p)\, \psi_i(x) - p\psi_i(x) \right) \\ &= \left(1 - \frac{2p}{n}\right) \psi_i(x) \end{aligned}$$

Since we chose x arbitrarily,

$$\mathbf{T}\psi_i = \left(1 - \frac{2p}{n}\right) \psi_i$$

and ψ_i is an eigenfunction of $\mathbf{T}$. □

We define $\varphi^{(p)}$ as the *Walsh span* of order p.

$$\varphi^{(p)}(x) = \sum_{i:\langle i,i\rangle=p} w_i \psi_i(x)$$

Intuitively, $\varphi^{(p)}$ is an element of the linear space spanned by the Walsh functions of order p. Now we can write the objective function as a sum over Walsh spans of each order p (recall p is bounded by k).

$$f(x) = \sum_{p=0}^{k} \varphi^{(p)}(x) \tag{7}$$

We now show that this is a superposition of elementary components.

Proposition 1. *The* p^{th} *Walsh span is an elementary landscape.*

Proof. We show that $\varphi^{(p)}$ is an eigenfunction of $\mathbf{T}$. Consider

$$\begin{aligned}
\mathbf{T}\varphi^{(p)} &= \mathbf{T}\left[\sum_{i:\langle i,i\rangle=p} w_i \psi_i\right] \\
&= \sum_{i:\langle i,i\rangle=p} w_i \left(1 - \frac{2p}{n}\right) \psi_i && \text{by Lemma 1} \\
&= \left(1 - \frac{2p}{n}\right)\left[\sum_{i:\langle i,i\rangle=p} w_i \psi_i\right] \\
&= \left(1 - \frac{2p}{n}\right)\varphi^{(p)}
\end{aligned}$$

thus $\varphi^{(p)}$ is an eigenfunction of $\mathbf{T}$ corresponding to eigenvalue $\left(1 - \frac{2p}{n}\right)$. □

We can use the decomposition from the previous section to compute the expectation of Y.

Corollary 1. *On any k-SAT instance, the expectation of the random variable* Y *is a linear combination of the* $k+1$ *Walsh spans evaluated at* x.

$$\mathbb{E}[Y] = \sum_{p=0}^{k} \left(1 - \frac{2p}{n}\right) \varphi^{(p)}(x)$$

This follows directly from the proposition along with Equations (4) and (7).

The following two lemmas will be useful in the next section. First, we show that the Walsh span of order zero is always a constant that is equal to the mean objective function value over X.

Lemma 2. *Let $\bar{f}$ be the mean objective value over X,*

$$\bar{f} = \frac{1}{|X|} \sum_{x \in X} f(x)$$

For all $x \in X$, the zeroth Walsh span is the constant function

$$\varphi^{(0)}(x) = \bar{f}$$

Proof. Let $x \in X$. There is only one Walsh function of order zero: $\psi_0(x) = 1$. We have $\varphi^{(0)}(x) = w_0 \psi_0(x) = w_0$. Note that for $p \neq 0$ we have

$$\frac{1}{|X|} \sum_{x \in X} \varphi^{(p)}(x) = 0 \tag{8}$$

because of the parity of bitstrings of order p. By some algebraic manipulation,

$$\begin{aligned}
w_0 &= \left(\frac{1}{|X|} \sum_{x \in X} w_0 \right) \\
&= \frac{1}{|X|} \sum_{x \in X} \varphi^{(0)}(x) \\
&= \frac{1}{|X|} \sum_{x \in X} \varphi^{(0)} + \frac{1}{|X|} \sum_{x \in X} \sum_{p=1}^{k} \varphi^{(p)}(x) && \text{by Eq. (8)} \\
&= \frac{1}{|X|} \sum_{x \in X} \sum_{p=0}^{k} \varphi^{(p)}(x) \\
&= \frac{1}{|X|} \sum_{x \in X} f(x) && \text{by Eq. (7)}
\end{aligned}$$

□

Corollary 2. *The objective function f for any k-SAT or MAX-k-SAT instance is a superposition of k elementary landscapes*

$$f(x) = \bar{f} + \sum_{p=1}^{k} \varphi^{(p)}(x)$$

In the next section, we will need to bound the value of $\varphi^{(p)}$ over all states $x \in X$. We use the absolute values of the Walsh coefficients w_i to do so.

Lemma 3. *For all $x \in X$,*

$$\sum_{\langle i,i \rangle = p} -|w_i| \leq \varphi^{(p)}(x) \leq \sum_{\langle i,i \rangle = p} |w_i|$$

Proof. Let x be an arbitrary state in X. By definition we have

$$\varphi^{(p)}(x) = \sum_{\langle i,i\rangle=p} w_i\psi_i(x) = \sum_{\langle i,i\rangle=p} \pm|w_i|$$

since $\psi_i(x) = \pm 1$ and $w_i = \pm|w_i|$. Clearly, the smallest that each term could be is $-|w_i|$ and the largest is $|w_i|$. □

3 Some Bounds for 3-SAT

Two structural search space characteristics that directly affect the performance of local heuristic search algorithms are *local maxima* and *plateaus*. In this section we will use the results from the previous section to prove some bounds on the evaluation of states that are local maxima or belong to plateaus of width greater than 1.

Before we continue we prove the following lemma that provides an identity for a series expansion that will allow for some algebraic manipulation in the theorems below.

Lemma 4. *On 3-SAT we have the following identity.*

$$\sum_{p=0}^{3} p\varphi^{(p)}(x) = 2f(x) - 2\bar{f} - \varphi^{(1)}(x) + \varphi^{(3)}(x)$$

Proof. The series is equal to

$$\sum_{p=0}^{3} p\varphi^{(p)}(x) = \varphi^{(1)}(x) + 2\varphi^{(2)}(x) + 3\varphi^{(3)}(x)$$

We can group the terms on the right hand side as follows

$$\left[\varphi^{(1)}(x) + \varphi^{(2)}(x) + \varphi^{(3)}(x)\right] + \left[\varphi^{(2)}(x) + 2\varphi^{(3)}(x)\right]$$

By the decomposition in Equation (7),

$$\left[f(x) - \varphi^{(0)}(x)\right] + \left[f(x) - \varphi^{(0)}(x) - \varphi^{(1)}(x) + \varphi^{(3)}(x)\right]$$

By Lemma 2,

$$\left[f(x) - \bar{f}\right] + \left[f(x) - \bar{f} - \varphi^{(1)}(x) + \varphi^{(3)}(x)\right]$$

and simplifying gives the result. □

A state x is said to be a *local maximum* if, for all $y \in N(x)$, $f(y) \leq f(x)$. We point out that this definition is distinct from studies that allow for multi-state local maxima (e.g., [5]). Our single-state definition coincides with Hoos and Stützle [3]. Furthermore, every *global maximum* is also a local maximum.

Grover [8] showed on *elementary* landscapes no local maxima (minima) lie below (above) the mean value of the objective function over X. This will not necessarily hold for arbitrary functions. However, we show here that the knowledge of the elementary components and their properties also allow us to bound the evaluation of local maxima on 3-SAT.

Theorem 1. *On any 3-SAT instance with n variables and m clauses, there exists a positive real number τ such that for any state x, if $f(x) < \bar{f} - \tau$, then x cannot be a local maximum.*

Proof. We begin by showing if $f(x) < \mathbb{E}[Y]$, it cannot be a local maximum. We will then use the previous results to bound the inequality. Let x be a state such that $f(x) < \mathbb{E}[Y]$. There exists some point y in the neighborhood of x that has an evaluation $f(y) > f(x)$. Thus x cannot be a local maximum. By Corollary 1 we thus have

$$f(x) < \mathbb{E}[Y]$$
$$f(x) < \sum_{p=0}^{3} \left(1 - \frac{2p}{n}\right) \varphi^{(p)}(x)$$
$$f(x) < \sum_{p=0}^{3} \varphi^{(p)}(x) - \frac{2}{n} \sum_{p=0}^{3} p\varphi^{(p)}(x)$$

The first term on the right hand side is simply the decomposition of $f(x)$ given by Equation (7). Thus we can make the following substitution.

$$f(x) < f(x) - \frac{2}{n} \sum_{p=0}^{3} p\varphi^{(p)}(x)$$

By Lemma 4,

$$f(x) < f(x) - \frac{2}{n} \left(2f(x) - 2\bar{f} - \varphi^{(1)}(x) + \varphi^{(3)}(x)\right)$$

Simplifying, we have

$$f(x) < \bar{f} + \frac{1}{2} \left(\varphi^{(1)}(x) - \varphi^{(3)}(x)\right) \tag{9}$$

Inequality (9) describes a threshold that depends on $\varphi^{(1)}(x)$ and $\varphi^{(3)}(x)$ such that if $f(x)$ is less than this threshold, x cannot be locally maximum. We now give a threshold that holds *over the entire search space.*

By Lemma 3, we have for *any* $x \in X$,

$$\left(\varphi^{(1)}(x) - \varphi^{(3)}(x)\right) \geq \left(\sum_{\langle i,i\rangle=1} -|w_i| - \sum_{\langle i,i\rangle=3} |w_i|\right)$$

and letting

$$\tau = \frac{1}{2}\left(\sum_{\langle i,i\rangle=1} |w_i| + \sum_{\langle i,i\rangle=3} |w_i|\right) \tag{10}$$

we now have the following bound on the r.h.s. of Inequality (9).

$$\bar{f} - \tau \leq \bar{f} + \frac{1}{2}\left(\varphi^{(1)}(x) - \varphi^{(3)}(x)\right)$$

and thus, for all $x \in X$, if $f(x) < \bar{f} - \tau$, then x cannot be a local maximum. The threshold $\bar{f} - \tau$ is simply computed (in polynomial time) by summing the absolute Walsh coefficients of order 1 and 3 and holds over the entire search space. □

In a similar manner, we can bound the function value at which plateaus of width greater than one can appear. A plateau is a maximal set P of states such that for all $x, y \in P$ there is a path $(x = x_1, x_2, \ldots, x_t = y)$ of length $t \geq 1$ with $f(x) = f(x_i)$ for $i = 1, 2, \ldots, t$ and, if $t > 1$, $x_{i+1} \in N(x_i)$. The *level* of a plateau P is the evaluation $f(x_p), \forall x_p \in P$.

We say the neighborhood of a state x is *flat* if, for all $y \in N(x)$, $f(y) = f(x)$, that is, x has the same value as all the states in its neighborhood. We show that flat neighborhoods cannot exist at certain levels of the objective function.

Theorem 2. *On any 3-SAT instance with n variables and m clauses, there exists a positive real number τ such that for any state x, if $f(x) < \bar{f} - \tau$ or $f(x) > \bar{f} + \tau$, then x cannot have a flat neighborhood.*

Proof. We prove the equivalent contrapositive. Let x be a state with a flat neighborhood. We have

$$\begin{aligned}
f(x) &= \mathbb{E}[Y] \\
&= \sum_{p=0}^{3}\left(1 - \frac{2p}{n}\right)\varphi^{(p)}(x) \\
&= \sum_{p=0}^{3}\varphi^{(p)}(x) - \frac{2}{n}\sum_{p=0}^{3} p\varphi^{(p)}(x) \\
&= f(x) - \frac{2}{n}\sum_{p=0}^{3} p\varphi^{(p)}(x) && \text{by Eq. (7)}
\end{aligned}$$

Therefore, at such a point x we must have

$$\begin{aligned}
\sum_{p=0}^{3} p\varphi^{(p)}(x) &= 0 \\
2f(x) - 2\bar{f} - \varphi^{(1)}(x) + \varphi^{(3)}(x) &= 0 && \text{by Lemma 4}
\end{aligned}$$

thus if x has a flat neighborhood, the following must hold.

$$f(x) = \bar{f} + \frac{1}{2}\left(\varphi^{(1)}(x) - \varphi^{(3)}(x)\right) \tag{11}$$

Using Lemma 3 we can bound the terms $\varphi^{(1)}(x)$ and $\varphi^{(3)}(x)$ giving the following

$$\bar{f} - \tau \leq f(x) \leq \bar{f} + \tau$$

where τ is given by Equation (10) in Theorem 1. □

Recall the width of a plateau P is the minimal length path between any state in P and one not in P. We have the following corollary.

Corollary 3. *A plateau P with level less than $\bar{f} - \tau$ or greater than $\bar{f} + \tau$ cannot have width greater than 1.*

Proof. This follows directly from the fact that no flat neighborhoods exist outside of the range $\bar{f} - \tau$ to $\bar{f} + \tau$. Thus, for these points, every state on a plateau P must have at least one neighbor outside P and the width of P is at most 1. □

4 Derived Values in Practice

We have shown how the average value of the neighborhood can be obtained analytically for any particular state and that a region (τ from $\bar{f}$) can be defined outside of which plateaus of width greater than one cannot exist and certain local optima cannot be found. We illustrate the proved properties in Figure 1. In this section, we show empirically that the expectation value computation is informative and that the region is non-trivial in benchmark problem instances.

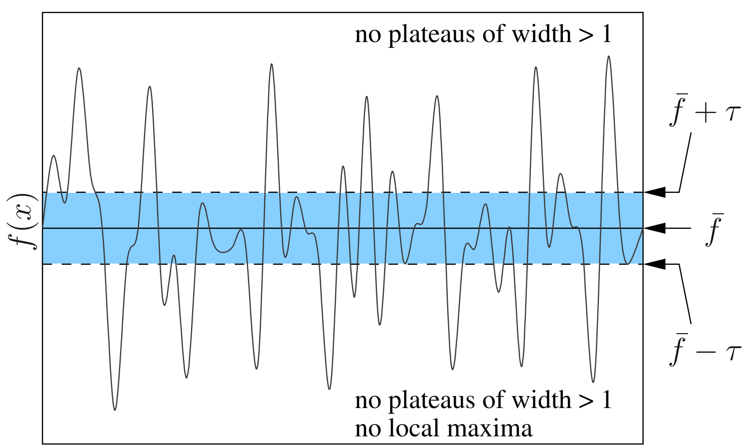

Fig. 1. An illustration of the proved properties. No plateaus of width strictly greater than one can lie outside the interval. No local maxima can lie below the interval.

4.1 Empirical Values of Neighborhood Expectation Value

The neighborhood expectation value computed in Equation (4) is useful because it can potentially provide algorithms with higher resolution information about states than the objective function. For example, given two states x and y with $f(x) = f(y)$, it is not necessarily the case that the neighborhood expectation values are equal for both x and y.

Stochastic local search algorithms applied to k-SAT problems often must select a neighboring state from a large set of moves with equal evaluation. This presents a problem for such algorithms due to the lack of gradient information in the neighborhood [3]. A collection of states at the same evaluation level are indistinguishable in terms of objective function value. However, we conjecture the expectation value can serve as a predictor of the number of *improving moves* that exit a particular state.

To illustrate this concept, we sampled 100 states at a particular objective function level ($f(x) = 390$) on each of 1000 instances that make up the uf100-430 benchmark set in SATLIB (100 vars, 430 clauses). For each point we calculated the correspondence between the expectation value given by Equation (4) and the actual number of improving moves in the neighborhood of the state. These data are plotted in Figure 2. A correlation test gives a strong positive correlation value of 0.51 with $p < 2.2 \times 10^{-16}$ indicating that better expectation leads to more potential for improvement. These data are preliminary indicators that the neighborhood expectation value can provide useful information about the neighborhoods of points even if they are equal in objective function value.

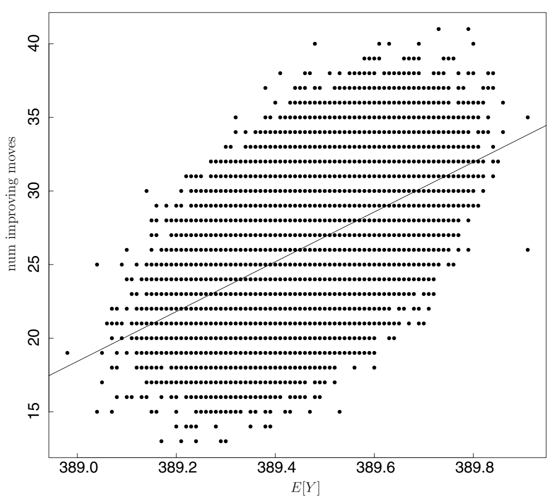

Fig. 2. Number of improving moves vs $\mathbb{E}[Y]$ at $f(x) = 390$ for 100 points each on 1000 instances of SATLIB benchmark set uf100-430. Line indicates linear best fit.

4.2 Empirical Values of τ

To demonstrate the region outside the interval is not trivial, we computed the values for τ as a percentage of the objective function range m across 18 benchmark distributions from SATLIB and the 2008 SAT competition. In Table 1 we report the mean (μ), standard deviation (σ), minimum, and maximum of the value τ/m over all N problems in each distribution.

The mean value of τ is consistently about 10% of the range m with a relatively low standard deviation. The maximum value of τ does not exceed 13% of the total objective function range over all the problem distributions we tested.

Table 1. Computed statistics for τ/m across several benchmark distributions from SATLIB and 2008 SAT competition

set	setsize	μ	σ	min	max
SATLIB					
uf20-91	1000	0.10252	0.00707	0.08104	0.12775
uf50-218	1000	0.10467	0.00421	0.08945	0.11984
uf75-325	100	0.10487	0.00358	0.09538	0.11231
uf100-430	1000	0.10483	0.00307	0.0968	0.11483
uf125-538	100	0.10477	0.00241	0.09898	0.11245
uf150-645	100	0.10514	0.00221	0.10039	0.11027
uf175-753	100	0.10533	0.00239	0.0991	0.11155
uf200-860	100	0.10469	0.00203	0.09942	0.11047
uf225-960	100	0.10484	0.00194	0.0987	0.10898
uf250-1065	100	0.10478	0.00167	0.10082	0.10986
uuf50-218	1000	0.10131	0.00406	0.08888	0.1164
2008 SAT competition					
v360	10	0.10382	0.00146	0.10046	0.10535
v400	10	0.1037	0.00198	0.10072	0.10651
v450	10	0.10369	0.00162	0.10016	0.10571
v500	10	0.10384	0.00177	0.09947	0.10616
v550	10	0.10366	0.00113	0.10137	0.10494
v600	10	0.10404	0.00107	0.1027	0.10603
v650	10	0.104	0.00108	0.10293	0.10627

5 Conclusion

Studying the structural characteristics of combinatorial search spaces is important to understanding the behavior of stochastic search algorithms. These characteristics, along with how algorithms respond to them, define how poorly or how well the algorithm performs, in some cases determining whether a problem or problem class is easily solved or not. We have presented analytical tools for analyzing the search space of k-SAT and MAX-k-SAT.

We have shown that the landscape formalism provides insight into certain structural relationships. We have shown that the decomposition of the objective function into elementary components supplies us with the expectation value of the objective function of neighboring states. We have also proved that the objective function of k-SAT can be decomposed into k computationally efficient elementary landscape functions. We have applied this result to obtain previously unknown bounds on the objective function levels for local maxima and plateau width in the 3-SAT search space.

We have shown empirically on a large number of cases that the region for which our results hold cover the majority of the objective function range. We also have demonstrated that neighborhood expectation varies across a set of states of equal evaluation and that this expectation correlates with improvement. Clearly the relationship between expectation and improvement needs to be carefully explored as does the implications of the theoretical results to algorithm design.

References

1. Gent, I.P., Walsh, T.: Towards an understanding of hill-climbing procedures for sat. In: Proc. of AAAI 1993, pp. 28–33. MIT Press, Cambridge (1993)
2. Parkes, A.J., Walser, J.P.: Tuning local search for satisfiability testing. In: Proc. of AAAI 1996, pp. 356–362. MIT Press, Cambridge (1996)
3. Hoos, H.H., Stützle, T.: Stochastic Local Search: Foundations and Applications. Morgan Kaufmann, San Francisco (2004)
4. Clark, D.A., Frank, J., Gent, I.P., MacIntyre, E., Tomov, N., Walsh, T.: Local search and the number of solutions. In: Freuder, E.C. (ed.) CP 1996. LNCS, vol. 1118, pp. 119–133. Springer, Heidelberg (1996)
5. Frank, J., Cheeseman, P., Stutz, J.: When gravity fails: Local search topology. J. of Artificial Intelligence Research 7, 249–281 (1997)
6. Yokoo, M.: Why adding more constraints makes a problem easier for hill-climbing algorithms: Analyzing landscapes of CSPs. In: Smolka, G. (ed.) CP 1997. LNCS, vol. 1330, pp. 356–370. Springer, Heidelberg (1997)
7. Stadler, P.F.: Toward a theory of landscapes. In: Lopéz-Peña, R., Capovilla, R., García-Pelayo, R., Waelbroeck, H., Zertruche, F. (eds.) Complex Systems and Binary Networks, pp. 77–163. Springer, Heidelberg (1995)
8. Grover, L.K.: Local search and the local structure of NP-complete problems. Operations Research Letters 12, 235–243 (1992)
9. Stadler, P.F.: Landscapes and their correlation functions. J. of Mathematical Chemistry 20, 1–45 (1996)
10. Rockmore, D., Kostelec, P., Hordijk, W., Stadler, P.F.: Fast Fourier transform for fitness landscapes. Applied and Computational Harmonic Analysis 12, 57–76 (2002)
11. Whitley, L.D., Sutton, A.M., Howe, A.E.: Understanding elementary landscapes. In: Proc. of GECCO, Atlanta, GA (July 2008)
12. Rana, S., Heckendorn, R.B., Whitley, L.D.: A tractable Walsh analysis of SAT and its implications for genetic algorithms. In: Proc. of AAAI 1998, pp. 392–397 (1998)