Estimating the Number of Local Maxima for k-SAT Instances

A.A. Albrecht
Science & Technology Research Institute
University of Hertfordshire
Hatfield, Herts AL10 9AB, UK

P.C.R. Lane
School of Computer Science
University of Hertfordshire
Hatfield, Herts AL10 9AB, UK

K. Steinhöfel
Department of Computer Science
King's College London
Strand, London WC2R 2LS, UK

Abstract

We explore the applicability of a sampling method devised by J. Garnier and L. Kallel (SIAM J. Discrete Math., 2002) to approximate the number of local maxima in search spaces induced by k-SAT instances and a simple neighbourhood relation. The objective function is given by the number of satisfied clauses. Although the problem setting for k-SAT instances does not meet all pre-conditions required by the Garnier/Kallel-approach, we obtain approximations of the number of local maxima within the same order of magnitude as the exact values that have been determined by complete search. Since the comparison requires calculation of the complete set of local maxima, only small k-SAT instances have been considered so far. Furthermore, we outline a method for obtaining upper bounds for the average number of local maxima in k-SAT instances, which shows changes in the average number around the phase transition threshold.

1 Introduction

Much attention has been paid in recent years to local search algorithms as one of the basic methods to solve k-SAT problems. A first summary was presented in (Hoos and Stützle 1999) along with an empirical analysis of run-time distributions for various local search-based methods such as WalkSAT (Selman, Kautz, and Cohen 1994). Improvements on run-time estimations for k-SAT problems as well as for CNFs with unconstrained clause lengths are reported in (Schöning 2002; Brueggemann and Kern 2004; Schuler 2005; Dantsin and Wolpert 2005; Paturi et al. 2005), which are partly based on randomised local search methods. Significant progress has been achieved in the analysis of phase transitions since this effect was discovered in (Mitchell, Selman, and Levesque 1992; Selman, Levesque, and Mitchell 1992). Sophisticated methods from statistical mechanics (Martin, Monasson, and Zecchina 2001; Mézard and Zecchina 2002; Montanari, Parisi, and Ricci-Tersenghi 2004) provided quite accurate estimates for the crucial phase transition parameter, which eventually led to a rigorous proof of a tight bound of $2^k \cdot \log 2 - O(k)$ for the phase transition threshold as presented in (Achlioptas and Peres 2004).

In the present paper, we attempt to analyse the number of local maxima in a combinatorial landscape induced by a k-CNF and a simple neighbourhood function, with the objective function being the number of satisfied clauses for a given assignment of binary values. Combinatorial landscape analysis has become a major tool in the design of search-based algorithms, see (Reidys and Stadler 2002).

Reeves and Eremeev (2004) have demonstrated how to incorporate the number of local optima into run-time estimates of local search algorithms. For landscapes that can be partitioned into attraction basins, they proved that with probability α all local optima have been covered by local search with random restart after a waiting time of $\nu \cdot \ln(\nu+\gamma) + z_\alpha \cdot \sqrt{(\nu\cdot\pi)^2/6 + 1 - \nu\cdot\ln(\nu+\gamma)}$, where ν is the number of local optima, γ is the Euler-Mascheroni constant, and z_α is an appropriate confidence coefficient. Thus, estimates for ν provide information, e.g., for the selection of the population size in parallelized versions of local search algorithms.

In a slightly different context (Max-SAT and local minima), Zhang (2005) proposes a landscape-based method that performs especially well on overconstrained random Max-SAT instances. Moreover, Zhang's algorithm finds satisfiable solutions on large k-SAT instances more often than WalkSAT. The paper highlights the importance of how to deal with individual instances rather than with collections of (randomly selected) problem instances.

In our paper, we utilise the approach devised in (Garnier and Kallel 2002) to estimate the number of local maxima for a given problem instance, where sample data are used to approximate a probability distribution associated with the landscape induced by the problem instance. The results are discussed against the information gathered by a complete analysis of the landscape for a limited number of k-SAT problem instances. Given the nature of the problem, we were able to analyse only small scale instances and overall only a limited number of different, randomly generated k-SAT instances. Apart from the experimental analysis based on the Garnier/Kallel-approach, we derive a rough estimate of local maxima in terms of parameters of individual problem instances for the given, simple neighbourhood relation.

2 Basic Notations

We follow mainly the notations from (Achlioptas and Peres 2004): for a set V of n Boolean variables let $C_k(V)$ denote the set of all $\binom{n}{k} \cdot 2^k$ different disjunctive k-clauses on V, i.e. repeated literals and tautologies are excluded. A k-CNF is formed by selecting m different clauses C from $C_k(V)$ and taking their conjunction. We note that the selection does not imply - as in (Achlioptas and Peres 2004) - that the k-CNF strictly depends upon all n variables. The set of all such k-CNF consisting of m clauses is denoted by $\mathcal{F}_k(n,m)$. The set of m clauses forming $F \in \mathcal{F}_k(n,m)$ is denoted by $\mathcal{C}(F)$, and $\mathsf{Z}_F(\tilde{\sigma})$ is the number of satisfied clauses $C \in \mathcal{C}(F)$ on the truth assignment $\tilde{\sigma} = (\sigma_1, ..., \sigma_n)$, i.e. $0 \leq \mathcal{Z}_F(\tilde{\sigma}) \leq m$ and F is satisfiable, if there exists $\tilde{\eta}$ such that $\mathcal{Z}_F(\tilde{\eta}) = m$.

In (Schuurmans and Southey 2001), various neighbourhood functions are analysed that employ information about $\mathsf{Z}_F(\tilde{\sigma})$ and elements of $\mathcal{C}(F)$ that maximise changes of the objective function in one way or another. For example, flipping values of truth assignments is determined by unsatisfied clauses only, see also (Seitz and Orponen 2003; Seitz, Alava, and Orponen 2005). We consider a simple, unconstrained (i.e., features of clauses w.r.t. $\mathsf{Z}_F(\tilde{\sigma})$ are not taken into account) neighbourhood function where the value of a single variable is flipped, which makes it possible to consider the elements of the unit cube $\{0,1\}^n$ as elements of the configuration space. Thus, the landscape $\mathcal{L}(F)$ for F is induced by $\mathsf{Z}_F(\tilde{\sigma})$, $\tilde{\sigma} \in \{0,1\}^n$, and the neighbourhood relation

$$\mathcal{N}(\tilde{\sigma}) = \big\{\tilde{\sigma}' | d(\tilde{\sigma}, \tilde{\sigma}') = 1\big\}, \tag{1}$$

where $d(\tilde{\sigma}, \tilde{\sigma}')$ is the Hamming distance.

If $\forall \tilde{\sigma}' \big(\tilde{\sigma}' \in \mathcal{N}(\tilde{\sigma}) \rightarrow (\mathsf{Z}_F(\tilde{\sigma}') \leq \mathsf{Z}_F(\tilde{\sigma}))\big)$, then $\tilde{\sigma}$ is called a local maximum (which also covers global maxima).

3 The Garnier/Kallel-Approach

In the present paper, we are solely concerned with the landscape analysis called *inverse problem* in (Garnier and Kallel 2002), i.e. M elements of the landscape are selected at random as initial points of a pre-defined local search procedure. Then, for j initial points, where $1 \leq j \leq M$, the local search procedure is started and executed until a (local) maximum has been detected. The number of *different* (local) maxima is denoted by β_j. The local search procedure is quasi-deterministic and follows the steepest ascent rule: for the intermediate landscape element $\tilde{\sigma}$, all elements of $\mathcal{N}(\tilde{\sigma})$ are examined and one of the neighbours $\tilde{\sigma}'$ with the highest value of $\mathsf{Z}_F(\tilde{\sigma}')$ among all neighbours is chosen as the successor of $\tilde{\sigma}$ in the search procedure. The search terminates, if no improvement of the objective function can be achieved. In (Garnier and Kallel 2002), and the same applies to (Reeves and Eremeev 2004), a single element $\tilde{\sigma}' \in \mathcal{N}(\tilde{\sigma})$ is assumed at each step that maximises $\mathsf{Z}_F(\tilde{\sigma}')$, which implies a partition of $\mathcal{L}$ into attraction basins A_i, where $1 \leq i \leq N$ for a total number of N local and global maxima. The set A_i consists of all elements of $\mathcal{L}$ that lead to the i^{th} local or global maximum by the steepest ascent local search. The assumption affects the normalised size $\alpha_i = |A_i|/|\mathcal{L}|$ of attraction basins and $\sum_{i=1}^{N} |A_i|/|\mathcal{L}| = 1$. Since we employ the Garnier/Kallel-approach in an experimental context, we assume in the following that the impact of random selections among $\tilde{\sigma}'$ that maximise $\mathsf{Z}_F(\tilde{\sigma}')$ within a given neighbourhood is negligible.

Garnier and Kallel (2002) assume that the normalised sizes α_i of attraction basins can be described by a distribution parametrized by some positive number γ as follows: let $(Z_i)_{i=1,..,N}$ be a sequence of independent random variables whose common distribution has density p_γ defined by

$$p_\gamma = \frac{\gamma^\gamma}{\Gamma(z)} \cdot z^{\gamma-1} \cdot e^{-\gamma \cdot z}, \tag{2}$$

where $\Gamma(z) = \int_0^\infty e^{-t} \cdot t^{z-1} dt$. Let H^γ denote the assumption that the $(\alpha_i)_{i=1,...,N}$ can be approximated by $(Z_i/T_N)_{i=1,...,N}$, where $T_N = \sum_{i=1}^N Z_i$ with each Z_i having the density function p_γ. Furthermore, let $\beta_{j,\gamma} = \mathbb{E}_\gamma[\beta_j]$ denote the expected value of β_j, $j = 1, ..., M$. Garnier and Kallel (2002) prove that

$$\begin{aligned}\beta_{j,\gamma} \quad &= \quad N \cdot \binom{M}{j} \cdot \frac{\Gamma(\gamma + j)}{\Gamma(\gamma)} \cdot \frac{\Gamma(N \cdot \gamma)}{\Gamma((N-1)\cdot\gamma)} \times \\ &\times \frac{\Gamma((N-1)\cdot\gamma + M - j)}{\Gamma(N\cdot\gamma + M)}. \end{aligned} \tag{3}$$

We note that for $N = M/r$, a fixed value of M, and appropriate approximations of the Γ-function, the $\beta_{j,\gamma}$ can be approximated according to (3) as functions of (j, γ, r). For fixed r, Garnier and Kallel (2002) propose the χ^2 test to approximate γ for H^γ, which consists of calculating

$$T_\gamma = \sum_{j=1}^{M} \frac{(\beta_j - \beta_{j,\gamma})^2}{\beta_{j,\gamma}}, \tag{4}$$

where the β_j are given from observation and the $\beta_{j,\gamma}$ are approximated according to (3). The goal is then to determine

$$\gamma_0(r) = \operatorname{argmin}\{T_\gamma, \gamma > 0\} \tag{5}$$

by appropriate numerical methods. In our computational experiments, we incorporate the approximation of $\gamma_0(r)$ as a sub-routine in calculations where the parameter r varies (is decremented) until $\gamma_0(r)$ changes only marginally for $r = r_{\text{appr}}$, see Section 4. Thus, for a fixed (but sufficiently large) value of M the number of local maxima is finally estimated by

$$N = \frac{M}{r_{\text{appr}}}. \tag{6}$$

4 Computational Experiments

4.1 Evaluation of random 3-SAT instances

We fixed $k = 3$ and $n = 14, 20$, and randomly generated five instances from $\mathcal{F}_3(n,m)$ for varying ratios m/n: two below the phase transion threshold, two above the threshold, and one instance for $m/n \approx 4.267$.

For each of the five k-CNF we executed a complete search for local/global maxima in $\{0,1\}^{14}$ and $\{0,1\}^{20}$. The corresponding values of the number N of maxima are shown in the second column of Table 1 and Table 2, respectively.

We then selected three values for M, the number of random points chosen in $\{0,1\}^{14}$ and $\{0,1\}^{20}$ as initial elements for a deteministic steepest ascent search for local maxima. For each of the values M_i, $i = 1,2,3$, and some order of the M_i points we counted by β^i_j, $j = 1, ..., M_i$, the number of different maxima detected by the first j starting points for steepest ascent search. The values of β^i_M are recorded in the fourth column of Table 1 and Table 2.

An example of one of the five 3-CNF for $n = 14$ is given below:

$$x_6 \vee x_5 \vee x_7 \wedge \neg x_7 \vee x_5 \vee x_2 \wedge x_0 \vee \neg x_7 \vee x_4 \wedge \neg x_8 \vee x_{13} \vee x_1 \wedge x_7 \vee \neg x_{11} \vee \neg x_{12} \wedge x_{10} \vee x_3 \vee \neg x_9 \wedge x_8 \vee x_4 \vee x_9 \wedge x_3 \vee x_{12} \vee x_{11} \wedge x_{12} \vee x_{10} \vee x_9 \wedge \neg x_3 \vee x_2 \vee x_{11} \wedge x_5 \vee x_4 \vee x_6 \wedge x_7 \vee x_5 \vee \neg x_2 \wedge \neg x_2 \vee \neg x_7 \vee x_4 \wedge \neg x_5 \vee x_6 \vee x_3 \wedge x_2 \vee x_0 \vee x_3 \wedge \neg x_0 \vee x_3 \vee x_1 \wedge \neg x_0 \vee \neg x_{10} \vee x_4 \wedge \neg x_4 \vee \neg x_{11} \vee x_0 \wedge \neg x_1 \vee \neg x_{11} \vee \neg x_9 \wedge x_2 \vee x_6 \vee x_4 \wedge \neg x_{12} \vee x_9 \vee \neg x_1 \wedge x_6 \vee x_0 \vee \neg x_2 \wedge \neg x_2 \vee x_0 \vee \neg x_1 \wedge x_2 \vee \neg x_{11} \vee \neg x_3 \wedge \neg x_5 \vee \neg x_8 \vee \neg x_0 \wedge x_8 \vee \neg x_6 \vee \neg x_3 \wedge x_7 \vee x_{13} \vee x_3 \wedge \neg x_{12} \vee \neg x_{11} \vee x_3 \wedge \neg x_0 \vee x_7 \vee x_{13} \wedge x_6 \vee \neg x_3 \vee x_9 \wedge \neg x_7 \vee x_{10} \vee x_{13} \wedge \neg x_3 \vee x_7 \vee \neg x_2 \wedge \neg x_7 \vee x_1 \vee x_5 \wedge x_4 \vee x_2 \vee x_{13} \wedge \neg x_{11} \vee x_{12} \vee \neg x_{10} \wedge \neg x_8 \vee x_{10} \vee \neg x_6 \wedge x_{11} \vee \neg x_6 \vee \neg x_2 \wedge x_6 \vee \neg x_{10} \vee x_{13} \wedge x_3 \vee x_1 \vee \neg x_8 \wedge x_0 \vee \neg x_{10} \vee \neg x_4 \wedge x_{12} \vee x_0 \vee \neg x_2 \wedge x_3 \vee x_{12} \vee \neg x_{13} \wedge \neg x_2 \vee \neg x_8 \vee x_{12} \wedge x_{12} \vee x_8 \vee \neg x_4 \wedge \neg x_7 \vee x_3 \vee x_0 \wedge x_6 \vee \neg x_4 \vee x_8 \wedge x_9 \vee x_2 \vee \neg x_4 \wedge \neg x_5 \vee x_8 \vee \neg x_3 \wedge x_4 \vee \neg x_5 \vee x_8 \wedge x_{12} \vee x_9 \vee x_0 \wedge \neg x_5 \vee x_8 \vee \neg x_9.$$

4.2 Approximation of H^γ

For the calculation of $\beta_{j,\gamma}$ according to (3) we implemented the following procedure, which actually approximates $\beta_{j,\gamma}$ since we employ an approximation of the Γ-function. We recall that in (3) the (unknown) N is substituted by M/r, where M is selected as described in Section 4.1 and r is a variable in our calculations.

At first, we represent Eqn. 3 by

$$\beta_{j,\gamma} = \frac{M}{r} \cdot \binom{M}{j} \cdot \frac{A_1}{A_2} \cdot \frac{B_1}{B_2} \cdot \frac{C_1}{C_2}, \quad \text{where} \tag{7}$$

$$A_1 = \Gamma(a_1) \ \text{ for } \ a_1 = \gamma + j; \tag{8}$$

$$A_2 = \Gamma(a_2) \ \text{ for } \ a_2 = \gamma; \tag{9}$$

$$B_1 = \Gamma(b_1) \ \text{ for } \ b_1 = \gamma \cdot \frac{M}{r}; \tag{10}$$

$$B_2 = \Gamma(b_2) \ \text{ for } \ b_2 = \gamma \cdot \Big(\frac{M}{r} - 1\Big); \tag{11}$$

$$C_1 = \Gamma(c_1) \ \text{ for } \ c_1 = M - j + \gamma \cdot \Big(\frac{M}{r} - 1\Big); \tag{12}$$

$$C_2 = \Gamma(c_2) \ \text{ for } \ c_2 = M + \gamma \cdot \frac{M}{r}. \tag{13}$$

Since in our case some of the values are very large, we use intermediately a representation by the natural logarithm, i.e. in the second step we calculate

$$Z = \ln\Bigg(\binom{M}{j} \cdot \frac{A_1}{A_2} \cdot \frac{B_1}{B_2} \cdot \frac{C_1}{C_2}\Bigg) \tag{14}$$

$$= \ln\binom{M}{j} + \ln A_1 + \ln B_1 + \ln C_1 - \tag{15}$$

$$- \ln A_2 - \ln B_2 - \ln C_2. \tag{16}$$

For each of the $\ln\Gamma(x)$ we employ the following approximation (due to Robert H. Windschitl, 2002):

$$\ln\Gamma(x) \approx \frac{1}{2} \cdot \Big(\ln(2 \cdot \pi) - \ln x\Big) + \tag{17}$$

$$+ x \cdot \Bigg(-1 + \ln\Bigg(x + \frac{1}{12 \cdot x - \frac{1}{10 \cdot x}}\Bigg)\Bigg), \tag{18}$$

i.e. $x = a_1, ..., c_2$. For the binomial coefficient we use the formula

$$\ln\binom{M}{j} = \sum_{s=1}^{M} \ln s - \sum_{t=1}^{j} \ln t - \sum_{u=1}^{M-j} \ln u. \tag{19}$$

Finally, we set

$$\beta_{j,\gamma} = \frac{M}{r} \cdot e^Z. \tag{20}$$

Eqn. 20 was then used as a sub-routine in the search for optimum settings of (r, γ):

1. For a fixed $r \geq r_0$ we searched for γ such that T_γ from (4) is minimised, i.e. Eqn. 4 and Eqn. 20 were repeatedly calculated for $\gamma \geq \gamma_0 = 0.1$ and $\gamma = \gamma + \delta$, until T_γ changed only marginally or increased above the minimum value obtained so far.
2. For r_0 and $r = r + \Delta \leq r_{\text{max}}$, the triplets (r, γ, T_γ) were recorded and finally r_{appr} with the minimum value of T_γ was selected.
3. The output was then determined by $N_{\text{appr}} = M/r_{\text{appr}}$.

The results are summarised in Table 1 and Table 2. Since both the instances as well as the number of instances are small, the values of N do not provide any statistical information. Our main goal here is to demonstrate that the implementation of the Garnier/Kallel-approach as described in the present section provides approximations N_{appr} in the region of the exact values N.

Since deterministic search is easy to implement and fast, if neighbours can be identified in an easy way, the procedure can be executed for large numbers of M, which has been done in the present study, i.e. the M_i are relatively large compared to 2^{14} and 2^{20}, respectively. As a result, we obtained values β_M that are close or even equal to the corresponding N for $n = 14$. We note that for $M = 512$ the value of N_{appr} is in four out of the five examples equal or close to $N/2$.

m	N	M	β_M	$\gamma(r_{\text{appr}})$	r_{appr}	$T_{\gamma(r_{\text{appr}})}$	N_{appr}
51	6	128	5	2.2	0.1	1.37	2
51	6	256	5	2.2	0.1	1.37	2
51	6	512	6	2.0	0.1	1.54	3
55	13	128	8	2.6	0.4	81.2	3
55	13	256	8	2.5	0.4	90.7	3
55	13	512	10	2.3	0.2	308.6	4
59	11	128	6	2.8	0.1	1.35	2
59	11	256	8	2.4	0.3	156.0	3
59	11	512	8	2.1	0.2	12.8	4
63	32	128	20	2.4	0.1	4859	8
63	32	256	24	2.3	0.1	9185	10
63	32	512	27	2.2	0.1	17574	12
68	10	128	6	2.7	0.4	14.3	2
68	10	256	10	2	0.1	2.1	5
68	10	512	9	1.7	0.2	23.6	5

Table 1: Results for n = 14, k = 3

For $n = 20$ we see in all five sample functions a clear improvement of the approximation with increasing M. The the value of N_{appr} is again in four out of the five examples close to $N/2$ for the largest $M = 2^{14}$.

Future research will focus on a wider range of parameter settings, a larger size of k-CNF instances and significantly larger sets of randomly selected k-CNF instances.

5 Local Maxima and k-CNF

For an arbitrary $\tilde{\sigma} \in \{0,1\}^n$ and $F \in \mathcal{F}_k(n,m)$, we set $\mathcal{C}_{\mathbf{0}}(F, \tilde{\sigma}) = \{C | C \in \mathcal{C}(F) \wedge C(\tilde{\sigma}) = 0\}$ and $\mathcal{C}_{\mathbf{1}}(F, \tilde{\sigma}) = \{C | C \in \mathcal{C}(F) \wedge$

m	N	M	β_M	$\gamma(r_{\text{appr}})$	r_{appr}	$T_{\gamma(r_{\text{appr}})}$	N_{appr}
78	213	2^{10}	66	0.1	4.23	1×10^{6}	16
78	213	2^{12}	112	0.1	2.90	2×10^{9}	39
78	213	2^{14}	157	0.1	2.36	1×10^{14}	67
82	35	2^{10}	19	0.1	2.64	321.39	7
82	35	2^{12}	32	0.1	2.09	5.06	15
82	35	2^{14}	33	0.1	2.06	2139.39	16
86	142	2^{10}	72	0.1	2.97	6×10^{6}	24
86	142	2^{12}	117	0.1	2.21	2×10^{11}	53
86	142	2^{14}	136	0.1	2.04	6×10^{13}	67
90	33	2^{10}	21	0.1	2.57	1012.02	8
90	33	2^{12}	31	0.1	2.06	134.94	15
90	33	2^{14}	33	0.1	2.00	2723	17
94	15	2^{10}	5	5.1	3.10	5.52	2
94	15	2^{12}	13	0.1	2.16	27.2	6
94	15	2^{14}	15	0.1	2.00	109.57	8

Table 2: Results for n = 20, k = 3

$C(\tilde{\sigma}) = 1\}$. Thus, clauses from $\mathcal{C}_{\mathbf{1}}(F,\tilde{\sigma})$ have at least one literal among the k literals that returns 1 on $\tilde{\sigma}$. Since in (1) we have $d(\tilde{\sigma},\tilde{\sigma}') = 1$, clauses with at least two literals returning 1 on $\tilde{\sigma}$ do not affect the re-calculation of Z_F in neighbourhood transitions out of $\tilde{\sigma}$. We therefore partition $\mathcal{C}_{\mathbf{1}}(F,\tilde{\sigma})$ into $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ and $\mathcal{C}_{\mathbf{1}}^{(\geq 2)}(F,\tilde{\sigma})$, i.e. $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ contains all $C \in \mathcal{C}(F)$ with exactly one literal that returns 1 on $\tilde{\sigma}$. We note the following simple observation:

Lemma 1 *The truth assignment $\tilde{\sigma}$ is a local maximum in $\mathcal{L}(F)$ iff for all $\tilde{\sigma}' \in \mathcal{N}(\tilde{\sigma})$:*

$$\begin{aligned} & |\{C|C(\tilde{\sigma}')=1 \wedge C\in\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})\}| && (21)\\ \leq\; & |\{C|C(\tilde{\sigma}')=0 \wedge C\in\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})\}|. && (22)\end{aligned}$$

Here, we do not exclude $\mathsf{Z}_F(\tilde{\sigma}) = m$.

For a literal x^{η} we use $x^{\eta} \in C$ to express that x^{η} is part of the disjunctive term C. Let $X_{\mathbf{0}}(\tilde{\sigma}) = \{x|\exists C \in \mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma}) \wedge x^{\overline{\sigma}} \in C\}|$ and $p = |X_{\mathbf{0}}(\tilde{\sigma})|$ be the number of variables that occur in clauses of $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$, where we employ $\sigma^{\overline{\sigma}} \equiv 0$. Furthermore, we set $q = |\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})|$, $r = |\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})|$, and $s = |\mathcal{C}_{\mathbf{1}}^{(\geq 2)}(F,\tilde{\sigma})|$. Thus, we have for $F \in \mathcal{F}_k(n,m)$

$$m \;=\; q + r + s. \tag{23}$$

For $X_{\mathbf{1}} = \{x|\exists C \in \mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma}) \wedge x^{\sigma} \in C\}$, $t = |X_{\mathbf{1}}|$, and $h_u = |\{C|C \in \mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma}) \wedge x_{i_u}^{\sigma_{i_u}} \in C\}|$ we have

$$\sum_{u=1}^{t} h_u = r, \tag{24}$$

since the corresponding subsets of clauses have to be disjoint (otherwise, a clause from the intersection would belong to $\mathcal{C}_{\mathbf{1}}^{(\geq 2)}(F,\tilde{\sigma})$).

Lemma 2 *If $x_{i_u} \in X_{\mathbf{1}} \backslash X_{\mathbf{0}} \neq \emptyset$, then the neighbourhood transition that involves x_{i_u} diminishes $\mathsf{Z}_F(\tilde{\sigma})$ by h_u.*

This follows from the definitions of $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$ and $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$. For $f_u = |\{C|C \in \mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma}) \wedge x_{i_u}^{\overline{\sigma_{i_u}}} \in C\}|$, Lemma 1 can now be rewritten as

Lemma 3 *The truth assignment $\tilde{\sigma}$ is a local maximum in $\mathcal{L}(F)$ iff $X_{\mathbf{0}} \subseteq X_{\mathbf{1}}$ and for $x_{i_u} \in X_{\mathbf{0}}$:*

$$f_u \leq h_u. \tag{25}$$

We note that by definition

$$\sum_{u=1}^{p} f_u = q \cdot k, \tag{26}$$

and (24) and (25) imply for a local maximum

$$q \cdot k \leq r. \tag{27}$$

Let $\mathcal{M}_k^{\tilde{\sigma}}(n,m) \subseteq \mathcal{F}_k(n,m)$ denote the set of k-CNF that have $\tilde{\sigma}$ as a local maximum for the neighbourhood defined by $\mathcal{N}(\tilde{\sigma})$ and the objective function defined by Z_F, where we require $\mathsf{Z}_F(\tilde{\sigma}) < m$, i.e. $q \geq 1$ and $\tilde{\sigma}$ is not a satisfying assignment.

We are now going to derive a (rough) upper bound for $M_{\tilde{\sigma}} = |\mathcal{M}_k^{\tilde{\sigma}}(n,m)|$. As will be seen later, the ratio $2^n \cdot M_{\tilde{\sigma}}/|\mathcal{F}_k(n,m)|$, when approximated by using an upper bound of $M_{\tilde{\sigma}}$, then provides some information about typical values for the number of local maxima for k-CNF in terms of parameters (k,n,m).

For fixed (p,q,r,s), we consider the number of potential sets $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$, $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$, and $\mathcal{C}_{\mathbf{1}}^{(2)}(F,\tilde{\sigma})$ under the assumption that the fixed truth assignment $\tilde{\sigma}$ is a local maximum. Here, it is useful to consider bipartite graphs where one set of nodes represents the clauses of $\mathcal{C}$ with fixed degree k, and the other set of nodes is formed by the elements of $\{\sigma_i; i = 1,2,...,n\}$.

For $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$ we have to ensure that each of the p elements of $X_{\mathbf{0}}(\tilde{\sigma})$ is present in at least one of the clauses from $\mathcal{C}_{\mathbf{0}}$ and we therefore need

$$q \cdot k \geq p \geq k \;\text{ and }\; \binom{p}{k} \geq q. \tag{28}$$

Let $A(p,q,k)$ denote the number of sets $\mathcal{H}$ of size q consisting of k-selections $S = \{x_{i_1},...,x_{i_k}\}$ out of p variables of $X_{\mathbf{0}}(\tilde{\sigma})$ such that $\forall x\Big(x \in X_{\mathbf{0}} \rightarrow \exists S(S \in \mathcal{H} \wedge x \in S)\Big)$. Since in the given context the elements of $X_{\mathbf{0}}(\tilde{\sigma})$ are independent of each other, we have

$$\begin{aligned} A(p,q,k) \;&=\; \binom{\binom{p}{k}}{q} - p \cdot A(p-1,q,k) - \\ &\quad - \binom{p}{2} \cdot A(p-2,q,k) - \ldots - \\ &\quad - \binom{p}{s_{\text{fin}}} \cdot A(p-s_{\text{fin}},q,k), \end{aligned} \tag{29}$$

where s_{fin} is defined by $\binom{p-(s_{\text{fin}}+1)}{k} < q$. By substituting the $A(p-i,q,k)$ recurrently, we obtain the inclusion/exclusion-type equation

$$A(p,q,k) = \sum_{i=0}^{s_{\text{fin}}} \Big(-1\Big)^i \cdot \binom{p}{i} \cdot \binom{\binom{p-i}{k}}{q}, \tag{30}$$

which represents the number of sets $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$. Taking $\binom{p}{2\cdot s-1} \cdot \binom{\binom{p-2\cdot s-1}{k}}{q} - \binom{p}{2\cdot s} \cdot \binom{\binom{p-2\cdot s}{k}}{q}$ together and applying Stirling's formula, one can see that the impact of $\sum_{i=1}^{s_{\text{fin}}} \Big(-1\Big)^i \cdot \binom{p}{i} \cdot \binom{\binom{p-i}{k}}{q}$ is only marginal and we therefore employ $\binom{\binom{p}{k}}{q}$ to upper bound the number of sets $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$.

For $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ we consider the set $X_{\mathbf{1}}$: for f_u clauses from $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$ with $x_{i_u}^{\overline{\sigma_{i_u}}}$ we have $h_u \geq f_u$ clauses from $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ with $x_{i_u}^{\sigma_{i_u}}$, if $\tilde{\sigma}$ is a local maximum. In each of the h_u clauses, the literals different from $x_{i_u}^{\sigma_i}$ are of the type $x_j^{\overline{\sigma_j}}$ due to the definition of $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$. Thus, the number of different sets $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ can be upper bounded by

$$\sum_{t=p}^{n}\binom{n-p}{t-p}\cdot\sum_{\substack{(h_1,\ldots,h_t)\\ h_1\geq f_1,\ldots,h_p\geq f_p}}\prod_{u=1}^{t}\binom{\binom{n-1}{k-1}}{h_u}. \tag{31}$$

We recall that $t \geq p$ is required by Lemma 1.

For $\mathcal{C}_{\mathbf{1}}^{(\geq 2)}(F,\tilde{\sigma})$ we consider the set of all $\binom{n}{k}\cdot 2^k$ clauses: since $\tilde{\sigma}$ is fixed, among the set of all clauses there are $\binom{n}{k}$ clauses that return 0 on $\tilde{\sigma}$ (the clauses of $\mathcal{C}_{\mathbf{0}}(F,\tilde{\sigma})$ are drawn from this subset); there are $\binom{n}{k}\cdot k$ clauses with excatly one literal of type $x_{i_u}^{\sigma_i}$ (the clauses of $\mathcal{C}_{\mathbf{1}}^{(1)}(F,\tilde{\sigma})$ are drawn from this subset). Thus, the number of different sets $\mathcal{C}_{\mathbf{1}}^{(\geq 2)}(F,\tilde{\sigma})$ can be upper bounded by

$$\binom{\binom{n}{k}\cdot(2^k-k-1)}{s}. \tag{32}$$

We assume at first $t = n$ (implicitly also $q \geq n/k$) and set $r = q\cdot k + \Delta$ for $\Delta \geq 0$, cf. (27). Based on $\binom{K}{a}\cdot\binom{K}{b} \geq \binom{K}{a+b}$ (and the remark after (30)), we summarize (30), (31) and (32) to

$$M_{\tilde{\sigma}} < \sum_{p=k}^{n}\binom{n}{p}\cdot\Bigg\{\sum_{\substack{q+r+s=m\\ r\geq k\cdot q,\, q\geq 1}}\binom{\binom{p}{k}}{q}\cdot\left(\sum_{t=p}^{n}\binom{n-p}{t-p}\times\right.$$
$$\left.\times\binom{\binom{n-1}{k-1}}{\frac{r}{n}}^{n}\right)\cdot\binom{\binom{n}{k}\cdot(2^k-k-1)}{s}\Bigg\}. \tag{33}$$

(Note: we use $\binom{A}{B}^n$ for $\{\binom{A}{B}\}^n$.

We are now going to identify (q,r,s) such that the product on the RHS of (33) is maximised for fixed (p,k,n,m). At first, we consider for variable r and s the product

$$P_q(r;s) = \binom{\binom{n-1}{k-1}}{\frac{r}{n}}^{n}\cdot\binom{\binom{n}{k}\cdot(2^k-k-1)}{s}, \tag{34}$$

where $r{+}s = m{-}q$. We analyse $P_q(r;s) \leq P_q(r{-}1;s{+}1)$, which for $r = a\cdot n + b$, $1 \leq b < n$, turns to

$$\binom{\binom{n-1}{k-1}}{a}^{n-b}\cdot\binom{\binom{n-1}{k-1}}{a+1}^{b}\cdot\binom{\binom{n}{k}\cdot(2^k-k-1)}{s}$$
$$\leq \binom{\binom{n-1}{k-1}}{a}^{n-b+1}\cdot\binom{\binom{n-1}{k-1}}{a+1}^{b-1}\cdot\binom{\binom{n}{k}\cdot(2^k-k-1)}{s+1}. \tag{35}$$

For $b = 0$ we take a–1 and $b' = n$. By straightforward calculations one obtains that (35) is valid if $r \geq \widehat{r}$ for

$$\widehat{r} = \frac{(m-q)\cdot(k+\varepsilon_1)-(n-b)\cdot(2^k-k-1)+k+\varepsilon_2}{2^k-1+\varepsilon_3}, \tag{36}$$

where $\varepsilon_1 = n/\binom{n}{k}$, $\varepsilon_2 = b/\binom{n}{k}$, and $\varepsilon_3 = (n+1)/\binom{n}{k}$. Here, we assume that m is sufficiently large in relation to n and 2^k, which will be discussed further below in more detail.

Thus, if we assume $\widehat{r} > k\cdot q$ (cf. (27)), then (34) increases for increasing s from $0 \leq s \leq \widehat{s}$ and (34) decreases for increasing s from $\widehat{s} < s \leq s_{\max}$, where $\widehat{s} = m{-}q{-}\widehat{r}$ and $s_{\max} = m{-}q{-}k{\cdot}q$.

The condition $\widehat{r} \geq k\cdot q$ results in an upper bound for q:

$$k\cdot q \;\leq\; \frac{(m-q)\cdot(k+\varepsilon_1)-(n-b)\cdot(2^k-k-1)+k+\varepsilon_2}{2^k-1+\varepsilon_3}$$
$$q \;\leq\; q_1 = \frac{m\cdot(k+\varepsilon_1)-(n-b)\cdot(2^k-k-1)+k+\varepsilon_2}{k\cdot(2^k+\varepsilon_4)}, \tag{37}$$

where $\varepsilon_4 = (n+n/k+1)/\binom{n}{k}$. Furthermore, we need $k\cdot q/n \leq r/n < \binom{n-1}{k-1}$ and $k\cdot q \leq r = m-q-s \leq m-q$, which leads to

$$q \;<\; q_2 \;=\; \binom{n}{k}, \tag{38}$$

$$q \;\leq\; q_3 \;=\; \frac{m}{k+1}. \tag{39}$$

Summarizing these observations, we obtain that $P_q(r;s)$ from (34) is maximized (ignoring integer representations) at $P_q(\widetilde{r};\widetilde{s})$ for $\widetilde{s} = m-q-\widetilde{r}$ with $\widetilde{r} = \max\{k\cdot q,\widehat{r}\}$, where q is fixed but must obey the minimum upper bound defined by (37) until (39), depending upon the value of $\widetilde{r}$.

So far, we kept the parameter q fixed. Now we take into account $\binom{\binom{p}{k}}{q}$ from (33) for $p = n$ and try to maximize (and to compare)

$$\binom{\binom{n}{k}}{q}\cdot P_q(k\cdot q; m-q-k\cdot q); \tag{40}$$

$$\binom{\binom{n}{k}}{q}\cdot P_q(\widehat{r}; m-q-\widehat{r}). \tag{41}$$

By using a representation similar to the one of (35) and setting $A = \binom{n}{k}\cdot(2^k{-}k{-}1)$, $B = \binom{n}{k}$, the assumption about an increasing value of (40) for increasing q leads to

$$\left(\frac{A-m+(k+1)\cdot q+1}{m-(k+1)\cdot q-k}\right)^{k+1} < \left(\frac{B-q+a/k}{q+(n-a)/k}\right)^{k}\cdot\frac{B-q}{q+1}. \tag{42}$$

Depending on the value of a, one has to consider two cases of the type

$$\frac{A-m+(k+1)\cdot q+1}{m-(k+1)\cdot q-k} < \frac{B-q}{q+1}. \tag{43}$$

The resulting upper bound

$$q < \frac{m-2^k+1+(m-1)/\binom{n}{k}}{2^k+2/\binom{n}{k}} \tag{44}$$

is similar to the upper bound (37). Based on these observations we conjecture that (40) and (41) are maximized for $q^* \sim m/2^k$ and $\widehat{r} \approx k\cdot q^*$. A detailed analysis of all sub-cases will be subject of further research.

We note that $q^* \sim m/2^k$ and $\widehat{r} \approx k{\cdot}q^*$ actually ignore the relation of m to n and k. A more detailed analysis of (37) shows that $q_1 \sim m/2^k - (n{-}b)/k$, where $1 \leq b < n$. If $m \lesssim (n{-}b)\cdot\frac{2^k}{k}$ and $1 << b < n$, i.e. m is in the region of the phase transition threshold (Achlioptas and Peres 2004), then $q_1 < 0$ and, moreover, for q from (44) the condition $q > \lceil n/k \rceil$ might no longer be valid, which is required by (26) for $p = n$. If this is the case, the maximum value of (40) changes significantly. Therefore, we conjecture that for m in the region of the phase transition threshold (and below) the value of $M_{\tilde{\sigma}}$ (for a large fraction of $\tilde{\sigma}$) is significantly smaller than for $m >> O(n\cdot 2^k)$.

In Lemma 3, (30), (31), and (32) we exploit only information about $x_i^{\overline{\sigma_i}}$ vs. $x_i^{\sigma_i}$, i.e. information about the actual values of σ_i has no impact on $M_{\tilde{\sigma}}$ at all. Thus, $M_{\tilde{\sigma}}$ depends on structural parameters (n,k,m) only:

Lemma 4 *If $\tilde{\sigma}, \tilde{\eta} \in \{0,1\}^n$, then $M_{\tilde{\sigma}} = M_{\tilde{\eta}}$ for a given class $\mathcal{F}_k(n,m)$.*

Given (n,k,m), we denote by $R(n,k,m,q^*)$ the maximum value of $\binom{\binom{n}{k}}{q} \cdot P_q(r; m-q-r)$ as presented in (40) and (41). In (33), the value of $\sum_{p=k}^{n} \binom{n}{p} \sum_{t=p}^{n} \binom{n-p}{t-p}$ can be upper bounded by $(n-k)^2 \cdot 3^{(1+\varepsilon)n/3}$. We now have for $m \leq \binom{n}{k} \cdot (2^k - k - 1)$ the upper bound

$$M_{\tilde{\sigma}} \;<\; (n-k)^2 \cdot \frac{m^2}{2\cdot(k+1)} \cdot 3^{\frac{(1+\varepsilon)\cdot n}{3}} \cdot R(n,k,m,q^*), \quad (45)$$

where $R(n,k,m,q^*)$ depends on the relation of m to n and k. For an upper bound of the average number of local maxima per k-CNF, one has to multiply the RHS of (45) by 2^n (cf. Lemma 4) and to divide the expression by the number of k-CNF of length m. To obtain asymptotic expressions for different intervals of m requires, however, a detailed analysis of $R(n,k,m,q^*)$.

6 Concluding Remarks

The Garnier/Kallel-approach requires a partition of the search space into attraction basins, i.e. within each neighbourhood a single element with the maximum value of the objective function is assumed. This assumption does not apply to the neighbourhood in our study. Nevertheless, our computational experiments provide evidence that the sampling-based method for the approximation of the number of local maxima seems to work in the context of k-SAT instances. Furthermore, the outline of our method for obtaining upper bounds for the average number of local maxima per k-SAT instance suggests that the fraction of local maxima relative to the total number of truth assignments changes around the phase transition threshold. Future research will be directed towards a more comprehensive analysis of the Garnier/Kallel-approach for larger k-SAT instances and larger sets of parameter settings. Furthermore, we intend to derive asymptotic formulas for the average number of local maxima in terms of the basic parameters n, k, and m.

Acknowledgements

The author would like to thank the anonymous referees for their careful reading of the manuscript and helpful suggestions, in particular, one referee for pointing us to the paper by Weixiong Zhang (2004).

The research has been partially supported by EPSRC Grant EP/D062012/1.

References

Achlioptas, D., and Peres, Y. 2004. The threshold for random k-SAT is $2^k \cdot \log 2 - O(k)$. *J. Amer. Math. Soc.* 17(4): 947–973.

Brueggemann, T., and Kern, W. 2004. An improved local search algorithm for 3-SAT. *Theor. Comput. Sci.* 329(1-3): 303–313.

Dantsin, E., and Wolpert, A. 2005. An improved upper bound for SAT. In *Proc. SAT 2005*, LNCS 3569, 400407. Heidelberg: Springer-Verlag.

Garnier, J., and Kallel, L. 2002. Efficiency of local search with multiple local optima. *SIAM J. Discrete Math.* 15(1): 122–141.

Hoos, H.H., and Stützle, Th. 1999. Towards a characterisation of the behaviour of stochastic local search algorithms for SAT. *Artif. Intell.* 112: 213–232.

Martin, O.C.; Monasson, R.; and Zecchina, R. 2001. Statistical mechanics methods and phase transitions in optimization problems. *Theor. Comput. Sci.* 265 3–67.

Mézard, M., and Zecchina, R. 2002. Random K-satisfiability problem: from an analytic solution to an efficient algorithm. *Phys. Rev* 66:056126-1–27.

Mitchell, D.; Selman, B.; and Levesque, H. 1992. Hard and easy distributions of SAT problems. In *Proc. 10*$^{\text{th}}$ *Natl. Conf. on Artificial Intelligence*, 459–465. Menlo Park, CA: AAAI Press.

Montanari, A.; Parisi, G.; and Ricci-Tersenghi, F. 2004. Instability of one-step replica-symmetry-broken phase in satisfiability problems. *J. Phys. A: Math. Gen.* 37: 2073–2091.

Paturi, R.; Pudlák, P.; Saks, M.E.; and Zane, F. 2005. An improved exponential-time algorithm for k-SAT. *J. ACM* 52(3):337–364.

Reeves, C.R., and Eremeev, A.V. 2004. Statistical analysis of local search landscapes. *J. Oper. Res. Soc.* 55: 687–693.

Reidys, Ch.M., and Stadler, P.F. 2002. Combinatorial landscapes. *SIAM Rev.* 44(1):3–54.

Schöning, U. 2002. A probabilistic algorithm for k-SAT based on limited local search and restart. *Algorithmica* 32(4): 615–623, 2002.

Schuler, R. 2005. An algorithm for the satisfiability problem of formulas in conjunctive normal form. *J. Algorithms* 54(1): 40–44.

Schuurmans, D., and Southey, F. 2001. Local search characteristics of incomplete SAT procedures. *Artif. Intell.* 132: 121–150.

Seitz, S.; Alava, M.; and Orponen, P. 2005. Threshold behaviour of WalkSAT and Focused Metropolis search on random 3-satisfiability. In *Proc. SAT 2005*, LNCS 3569, 475–481. Heidelberg: Springer-Verlag.

Seitz, S., and Orponen, P. 2003. An efficient local search method for random 3-satisfiability. *Electr. Notes Discrete Math.* 16: 71–79.

Selman, B.; Kautz, H.A.; and Cohen, B. 1994. Noise strategies for improving local search. In *Proc. AAAI-94*, 337-343. Cambridge, MA: MIT Press.

Selman, B.; Levesque, H.; and Mitchell, D. 1992. A new method for solving hard satisfiability problems. In *Proc. AAAI-92*, 440–446. Cambridge, MA: MIT Press.

Zhang, W. 2004. Configuration landscape analysis and backbone guided local search. Part I: Satisfiability and maximum satisfiability. *Artif. Intell.* 158: 1–26.