1

Simulated Annealing
and its Problems to Color Graphs

Andreas Nolte and Rainer Schrader
Universität zu Köln, Weyertal 80, 50931 Köln, Germany

March 1999

Simulated Annealing is a very successful heuristic for various problems in combinatorial optimization. In this paper an application of Simulated Annealing to the 3-coloring problem is considered. In contrast to many good empirical results we will show for a certain class of graphs that the expected first hitting time of a proper coloring, given an arbitrary cooling scheme, is of exponential size.

These results are complementary to those in [13], where the convergence of Simulated Annealing to an optimal solution in exponential time is proved.

1 Introduction

Simulated Annealing has been a very successful general algorithm for the solution of large combinatorial optimization problems. It is a random local search heuristic, that has received much attention, since it was first introduced in [8].

In this paper we consider the Simulated Annealing algorithm applied to the 3-coloring problem which is known to be NP-complete [2]. Let $G = (V, E)$ be an undirected graph with n vertices. A 3-coloring of G is a mapping $f : V \to \{1, 2, 3\}$. f is a proper 3-coloring, if $f(u) \neq f(v)$ for all $(u, v) \in E$. The graph 3-coloring problem is to answer the question whether a given graph G has a proper 3-coloring. According to the experimental results in [7], Simulated Annealing is a very successful heuristic for this problem. In

*Supported by DFG grant (Schr 390/2-1)
Extended Version of [14]

contrast to these practical observations we will show the existence of two classes of graphs, where the expected first hitting time of a proper coloring is of exponential size. For the first class of graphs we assume a fixed cooling temperature for the Simulated Annealing algorithm. We can show for the first time an exponential first hitting time nearly without any limitations on the starting state. In the second class there is one natural starting state given (the empty coloring). Here, we proof an exponential first hitting time of the maximum coloring even for an arbitrary cooling schedule. There are only few negative results concerning the efficiency of Simulated Annealing applied to problems in combinatorial optimization. Observations of similar flavour have been established by Jerrum [6] for the clique problem on random graphs and by Sasaki and Hajek [16] for the matching problem.

This shows that the exponential running time bounds for Simulated Annealing proved by the authors in [13] are essentially tight.

2 A Homogeneous Approach

First, we consider a special case of Simulated Annealing called Metropolis algorithm [10]. It performs the Simulated Annealing strategy with fixed temperature $T \in \mathbb{R}^+$. We describe one natural way to apply the algorithm to the 3-coloring problem.

To start the algorithm, let an arbitrary coloring f be given. $c(f)$ denotes the number of "bad" edges (u, v) with $f(u) = f(v)$. We choose uniformly at random a vertex $v \in V$ and a color $j \in \{1, 2, 3\}$ and color the vertex with j. We move to the resulting coloring f', if $c(f') \leq c(f)$. If $c(f') > c(f)$, we move to f' with probability $\exp((c(f) - c(f'))/T)$ and stay at f with probability $1 - \exp((c(f) - c(f'))/T)$. We call the neighborhood structure defined by all possible transitions N. The sequence of states visited by the algorithm forms a Markov chain $(X_t)_{t \in \mathbb{N}}$ on the state space Ω of all colorings. It is easily seen that this Markov chain is irreducible and aperiodic, hence ergodic. Therefore it exists a unique probability vector π^T with $\pi^T P = \pi^T$, where P is the transition matrix of the chain with
$\forall (f, g) \in \Omega^2$

$$p_{fg} = \begin{cases} \min\{1, e^{\frac{c(f)-c(g)}{T}}\}\frac{1}{3n} & \text{for } f \neq g, (f,g) \in N \\ 0 & \text{for } f \neq g, (f,g) \notin N \\ 1 - \sum_{h \neq f} p_{fh} & \text{for } f = g \end{cases}$$

The important fact is that the distribution of the chain $P(X_t = .)$ tends to π^T as $t \to \infty$. It is easily verified that the Gibbs-distribution

$$\pi_f^T = \frac{e^{\frac{-c(f)}{T}}}{\sum_{h \in \Omega} e^{\frac{-c(h)}{T}}}$$

is the stationary distribution of the chain. By letting $T \to 0$ we can see that π^T converges to the uniform distribution concentrated on the optimal solutions. Therefore we could expect a good asymptotic performance of the chain, if we choose a very "deep" temperature T that is only slightly larger than 0 [9].

In the following we are interested in the performance of the chain depending on the problem size., i.e. convergence bounds of the chain after a finite number of steps. We will show in this section, that at least for general graphs good performance bounds do not exist by defining a class of graphs, where the expected first hitting time of a proper coloring will be of exponential size, even when we assume, that the starting state could be chosen uniformly at random from all possible starting states. Our example gives a stronger evidence of the sometimes bad performance of Simulated Annealing than those in [6] and [16], where only the existence of one starting state with exponential first hitting time of the optimum is shown.

Let $G_n = (V_n, E_n)$ with $V_n = \{x, z, y_1, \ldots, y_n\}$ and $E_n = \{(x, y_i), (y_i, z), i = 1, \ldots, n\} \cup \{(x, z)\}$ (Fig. 1). Obviously the graph G_n has only one proper

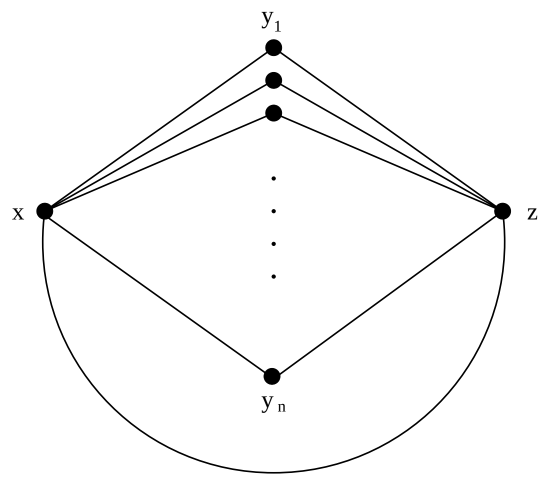

Figure 1: Graph G_n

partition in 3 parts and therefore only 6 proper colorings.

Theorem 1 *Almost all colorings f of G_n have an expected first hitting time H_f of a proper 3-coloring of G_n of exponential size, when f is selected as a starting state.*

Before we proof the theorem, we want to present the idea of the proof and show some helpful lemmata. In the following we will define a mapping

$g : \Omega \to \mathbb{N}$ from the set of all possible colorings to the natural numbers. A proper coloring will have the value $g(f) = 0$ and nearly all others will have positive values. A transition of the Markov chain will change the value of g by at most 1. We will consider the process $g(X_t)$, that is not necessarily Markovian any more and show under the hypothesis $X_t = f$ with certain colorings f that the expected value after one transition step is greater than $g(X_t) + \epsilon, \epsilon > 0$. Then we are able to prove with the help of a result in [4] about hitting time bounds implied by drift analysis, that the expected first hitting time of a coloring $X_t = f$ with $g(f) = 0$ is of exponential size, given a certain starting state. By specializing Theorem 2.3 of [4] on the discrete state space we get:

Theorem 2 [HAJEK] *Let $(Y_j)_{j\in\mathbb{N}}$ a sequence of random variables with values in $\mathbb{N}$ and $a, b \in \mathbb{R}, b < a$. The conditions*

C1 $\quad E(|Y_{j+1} - Y_j|\ |Y_j = i_j, \ldots, Y_0 = i_0) \leq s$

C2 $\quad E(Y_{j+1} - Y_j - \epsilon, Y_j < a, H_b > j|Y_j = i_j, \ldots, Y_0 = i_0) \geq 0$

with $\epsilon > 0, H_b = \min\{j|Y_j = b\}$ (first hitting time of b)

imply

$$\exists D, \eta \in \mathbb{R} \quad \forall i_0 \geq a, t \in \mathbb{N} \qquad P(H_b = t|Y_0 = i_o) \leq De^{\eta(b-a)}$$

(D, η depend only on s and ϵ, not on a or b).

□

Suppose f is a coloring of G_n with 3 colors. Then the set $\{y_1, \ldots y_n\}$ of nodes can be divided in at most 3 parts according to the coloring. Let $g(f) : \Omega \to \mathbb{N}$ be the size of the second largest part.

Lemma 3 *Let $c = \frac{1}{4}\frac{e^{-2/T}}{2+e^{-2/T}}, \epsilon = \frac{e^{-2/T}}{12}$ and f a coloring of G_n with $0 < g(f) \leq cn$. Let $X_t = f$. Then*

$$E(g(X_{t+1}) - g(X_t)) \geq \epsilon$$

follows.

Proof: Let $g(f) = k$. By considering the transition probabilities of the graph we get:

$$P(g(X_{t+1}) = k + 1) \geq \frac{e^{-2/T}}{3(n + 2)}(n - k)$$

and

$$P(g(X_{t+1}) = k - 1) \leq \frac{2}{3(n+2)}k.$$

Therefore we can conclude

$$\begin{aligned} & E(g(X_{t+1}) - g(X_t)) \\ \geq & \left(\frac{e^{-2/T}}{3(n+2)}(n-k) - \frac{2}{3(n+2)}k \right) \\ \geq & \frac{e^{-2/T}}{12} \\ = & \epsilon \end{aligned}$$

as required. □

To count those colorings with $g(f) > cn$ the following version of the Chernoff bound [12] is helpful. Let $B(n,p)$ denote the binomial distribution. By definition, if $x \in B(n,p)$, then $P(x = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Proposition 4 *If $x \in B(n,p)$, then for all $\alpha, 0 < \alpha < 1, P(x \leq (1-\alpha)np) < e^{-\alpha^2 np/2}$ and $P(x \geq (1+\alpha)np) < e^{-\alpha^2 np/3}$.*

Lemma 5 *Let c be as in Lemma 3. All but an exponential small fraction of the colorings of G_n satisfy $g(f) > cn$.*

Proof: Let z and y be colored arbitrarily and Y_r be the set of vertices y_i, that are colored red. By assuming, that every coloring of $\{y_1, \ldots, y_n\}$ has the same probability $1/3^n$, we get, that the color red is $B(n, 1/3)$ distributed. Therefore it follows with the last proposition:

$$P(|Y_r| < cn) < e^{-(1-3c)^2 n/6} < e^{-n/96}.$$

Because we could argue symmetrically for the other colors, we get:

$$P(g(f) < cn) \leq 3e^{-n/96}$$

as required. □

Now we are able to proof the main theorem:

Proof: Let $X_0 = f_0$ with $g(f_0) > cn$ be the starting state of the Metropolis-process. We consider only a subprocess of the actual Markov chain. Let Z_j be the points of time, when the value g of the actual coloring changes:

$Z_0 = 0$ and for all $j \in \mathbb{N} \quad Z_j = \min\{j > Z_{j-1} | g(X_j) \neq g(X_{Z_{j-1}})\}$. Suppose f is a coloring with $0 < g(F) < cn$. We get with Lemma 3 for all $j \in \mathbb{N}$:

$$E(g(X_{Z_{j+1}}) - g(X_{Z_j}) | X_{Z_{j+1}-1} = f, X_{Z_{j+1}-2}, \ldots, X_0) \geq \epsilon.$$

By the summation of the disjoint, given events we get:

$$E(g(X_{Z_{j+1}}) - g(X_{Z_j}) - \epsilon | 0 < g(X_{Z_j}) < cn, X_{Z_j}, \ldots, X_{Z_0}) \geq 0$$

and it follows

$$E(g(X_{Z_{j+1}}) - g(X_{Z_j}) - \epsilon | 0 < g(X_{Z_j}) < cn, g(X_{Z_j}) = i_j, \ldots, g(X_{z_0}) = i_0) \geq 0.$$

With $Y_j = g(X_{Z_j}), a = cn$ and $H_0 = \min\{t | Y_t = 0\}$ we obtain

$$E(Y_{j+1} - Y_j - \epsilon, Y_j < a, H_0 > j | Y_j = i_j, \ldots, Y_0 = i_0) \geq 0$$

as required for condition C1 of Theorem 2. Due to $|Y_{j+1} - Y_j| \leq 1$ condition C2 is also fulfilled and it follows:

$$\exists D, \eta \in \mathbb{R} \quad \forall i_0 \geq a, k \in \mathbb{N} \qquad P(H_0 = k | Y_0 = i_o) \leq De^{\eta(-a)} =: \delta$$

Using the fact that $X_0 = f_0$ we get

$$P(H_0 = t | Y_0 = g(f_0)) = P(H_0 = t | X_0 = f_0)$$

and

$$P(H_0 > t | X_0 = f_0) = 1 - \sum_{k=1}^{t} P(H_0 = k | X_0 = f_0) \geq \max\{0, 1 - t\delta\}.$$

Combining the last results yields

$$E(H_0 | X_0 = f_0) = \sum_{t=0}^{\infty} P(H_0 > t | X_0 = f_0) \geq \sum_{t=0}^{\infty} \max\{0, 1 - t\delta\} \geq \frac{1}{2\delta}.$$

Because $X_t =$ (proper 3-coloring) implies $g(X_t) = 0$

$$E(H_{f_0} | X_0 = f_0) \geq \frac{1}{2\delta} = \frac{De^{\eta cn}}{2} = e^{\Omega(n)}$$

follows. □

Remark: We suspect, that one can establish exponential first hitting time even for an arbitrary time dependent sequence T_t of temperatures. Unfortunately we have not been able to prove this inhomogeneous case for X_t by now. But in the next section we establish this for another application of the Simulated Annealing algorithm to the 3-coloring problem.

3 A different Approach

In this section we consider another approach for solving the 3-coloring problem. According to some positive experimental results [7] we try to find a maximum 3-colorable induced subgraph of a given graph $G = (V, E)$ (of maximal cardinality). Obviously the problem is also NP-hard.

Therefore the state space Ω of the corresponding Markov chain $(X_t)_{t\in\mathbb{N}}$ consists of all proper 3-colorings of arbitrary subgraphs of G. Let $(T_t)_{t\in\mathbb{N}}$ be an arbitrary sequence of temperatures, so that we are not concerned with the homogeneous case any more. Given a state $f \in \Omega$ at time t, we choose uniformly at random a vertex $v \in V$. If v is already colored, we remove v from the coloring with probability $\exp(1/T_t)$ and stay at f with probability $1-\exp(1/T_t)$. If v is not yet colored, we choose one of the 3 colors uniformly at random and try to color v. If we get a proper coloring, we move to this coloring. Otherwise we stay at f.

3.1 Convergence to the optimal solution

In this subsection we briefly describe the known theoretical bounds of the convergence of the distribution of the Markov chain to a distribution, that is concentrated on the optimal states (i.e. maximal 3-colorable subgraphs). We start with a description of the underlying inhomogeneous Markov chain.

According to the above defined strongly connected neighborhood structure $N \subset \Omega^2$ of the 3-colorings of subgraphs of G we get the following transition probabilities. For technical reasons we also assume, that we stay at each state with probability at least $1/2$, so that we obtain:

$\forall (f,g) \in \Omega^2$

$$p_{fg}^{(t)} = \begin{cases} \min\{1, e^{\frac{|g|-|f|}{T_t}}\}\frac{1}{2n} & \text{for } f \neq g, (f,g) \in N \\ 0 & \text{for } f \neq g, (f,g) \notin N \\ 1 - \sum_{f' \neq f} p_{ff'} & \text{for } f = g \end{cases}$$

where $|f| = \sharp\{v|v \text{ is colored by f}\}$. Clearly $p_{ff}^{(t)} \geq 1/2$, so that the chain is aperiodic. Assuming fixed transition probabilities p_{fg}^t (fixed temperature T_t) the resulting homogeneous chain is also ergodic and

$$\pi(t)_f = \frac{e^{\frac{|f|}{T_t}}}{\sum_g e^{\frac{|g|}{T_t}}} = \frac{t^{\frac{|f|}{\gamma}}}{\sum_g t^{\frac{|g|}{\gamma}}}$$

is the stationary distribution. Let $c_{\max} = \max\{|f| \mid f \in \Omega\}$. Obviously

$$\pi^* = \lim_{t\to\infty} \pi(t)_f = \lim_{t\to\infty} \frac{1}{\sum_g t^{\frac{|g|-|f|}{\gamma}}} = \begin{cases} 0, & \text{if } |f| < c_{\max} \\ \frac{1}{\sharp\{f \mid |f| = c_{\max}\}}, & \text{if } |f| = c_{\max} \end{cases}$$

The limit distribution is therefore a constant probability vector, which probability charges are concentrated on the global optima of the solution space.

Almost all results in the inhomogeneous case assume a logarithmic cooling schedule $T_t = \gamma/\ln(t)$, where γ is a problem-dependent parameter. Asymptotic results have been published by Hajek [5] and Tsitsiklis [17], who were even been able to find necessary and sufficient conditions for the convergence of Simulated Annealing. However, since we are concerned with finite time bounds, these asymptotic results are not helpful for us.

Anily and Federgruen [1], Gidas [3] and Mitra et al. [11] have obtained independently similar deterministic upper bounds for the proximity of the probability distribution of the state space after the generation of t transitions to the uniform distribution on the set of optimal states π^*. Applied to the 3-coloring problem the bound of Mitra [11] for example would yield:

$$t \geq c_1 \frac{1}{\epsilon^{n(3n)^n}} \quad \Rightarrow \quad \sum_f |P(X(t) = f) - \pi_f^*| \leq \epsilon$$

for a $c_1 \in \mathbb{N}$. In [13] the authors could improve these running time bounds by an application of the theory of rapidly mixing Markov chains. They could guarantee the same inequality with a considerably smaller bound $t \geq \epsilon^{-n} \exp(n^{c_2})$ with $c_2 \in \mathbb{N}$.

In the following section we will show that there is no hope to further improve this bound, i.e. we will prove that the exponential bound in [13] is essentially tight.

3.2 Negative Results

We will show in this section, that there will not exist better than exponential performance bounds of Simulated Annealing in general for the 3-coloring problem. In the following we give an example of a class of graphs, where the expected first hitting time of a maximal 3-colorable subgraph is of exponential size under the hypothesis $X_0 = \emptyset$, i.e. we start with the empty coloring, and an arbitrary cooling schedule.

Consider the following family of graphs:

Let $G_r = (R_1 \cup R_2, E_r)$ with $R_1 = \{r_i, 1 \le i \le n\}, R_2 = \{r_{jk}, 1 \le j,k \le n, j \neq k\}$ and $E_r = \{(r_{jk}, r_k), k \in \{1,\dots,n\} \setminus \{j\}\} \cup \{(r_{jk}, r_{lm}), k,m,j,l \in \{1,\dots,n\}, j \neq l, k \neq l\} \cup \{(r_{jk}, r_j), k \in \{1,\dots,n \setminus \{j\}\}\}$. Define now G_b and G_g as copies of G_r. G_n is formed by considering G_r, G_b, G_g as subgraphs and adding all edges between vertices, that are in different graphs G_*. Despite the long description G_n has a quite simple structure: $(R_1 \cup G_1 \cup B_1)$ is a complete 3-partite graph with $3n$ vertices. Every part R_1, G_1, B_1 has an additional set R_2, G_2, B_2 of vertices with the same structure.

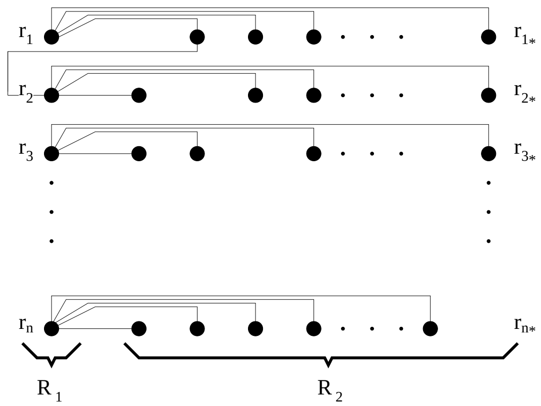

Figure 2: Graph G_r

Lemma 6 *The subgraph induced by $R_1 \cup B_1 \cup G_1$ is the maximal 3-colorable subgraph of G_n.*

Proof: $G(R_1 \cup B_1 \cup g_1)$ is obviously 3-colorable. Because every set of vertices colored with the same color has to be independent, we look for independent sets of size at least n. As every edege between G_r, G_b and G_g is realized, every independent set is completely in one of these subgraphs. Due to symmetry we only have to look at G_r.

Let U be an independent set in $G(R_1 \cup R_2)$. If $R_2 \cap U = \emptyset$ then $U \subset R_1$ and $|U| = n \Leftrightarrow R_1 = U$. If $R_2 \cap U \neq \emptyset$, then there exists exactly one j with $\{r_{jk}, k \in \{1,\dots,n\} \setminus \{j\}\} \cap U \neq \emptyset$ and $r_j \notin U$, because $G(R_2)$ is a complete n-partite graph.

Suppose now $\{r_{i_1}, \dots, r_{i_k}\} = U \cap R_1$. Then we obtain with $A = \{1,\dots,n\} \setminus \{i_1,\dots,i_k\} \setminus \{j\}$: $U \subset \{r_{i_1}, \dots, r_{i_k}, r_{jl}, l \in A\}$. Hence $|U| \le n-1$.

This implies, that the only independent sets of size of at least n are R_1, G_1 and B_1 and the Lemma follows. $\square$

Similar to the first section we will define a mapping $\omega : \Omega \to \mathbb{N}$ with $\omega(\emptyset) = n$ and $\omega(\text{maximal coloring}) = 0$. We will consider the process $\omega(X_t)$, that is not necessarily Markovian any more and show under the hypothesis $X_t = f$ with $\omega(X_t) < a$, that the expected value after one transition step is greater than $\omega(X_t) + \epsilon, \epsilon > 0$. Moreover the difference $|\omega(X_t) - \omega(X_{t-1})|$ is bounded. Then we are able to prove with the help of Theorem 2, that the expected first hitting time of a coloring $X_t = f$ with $\omega(f) = 0$ is of exponential size, given a starting state h with $\omega(h) > a$.

Suppose f is a proper 3-coloring of a subgraph of G_n and F is the set of the colored vertices. Let

$$\omega_r(F) = 3\sharp F \cap R_2 + \max\{\sharp\{\text{uncolored vertices in } R_1\} - 1, 0\}$$

and define $\omega_b(F)$ and $\omega_g(F)$ analogously. Then we can define

$$\omega(F) = \max\{\omega_r(F), \omega_g(F), \omega_b(F)\}.$$

It is easily seen, that $\omega(\emptyset) = n-1$ and $\omega(R_1 \cup G_1 \cup B_1) = 0$. Let $\omega(F) < n/2$. Then we can derive from the structure of G_n, that all vertices, that are colored with the same color, must be contained in one of the sets $R_1 \cup R_2, G_1 \cup G_2, B_1 \cup B_2$. We consider now those transitions, that color an additional vertex. Let

$$\begin{array}{rcl} M_\cup(F) & = & \{v \in V_n | \omega(F \cup \{v\}) \neq \omega(F)\} \\ M_\cup^+(F) & = & \{v \in V_n | \omega(F \cup \{v\}) > \omega(F)\} \\ M_\cup^-(F) & = & \{v \in V_n | \omega(F \cup \{v\}) < \omega(F)\}. \end{array}$$

($F \cup \{v\}$ denotes the set of colored vertices obtained by coloring the additional vertex v.)

Lemma 7 *Suppose* $0 < \omega(F) < n/2$. *Then* $\sharp M_\cup^-(F) \leq \sharp M_\cup^+(F)$.

Proof: First we suppose $\omega(F) = \omega_r(F) \wedge \omega(F) > \omega_b(F) \wedge \omega(F) > \omega_g(F)$ We get $M_\cup^-(F) \subset R_1$ from the construction of G_n. Let $\{r_{i_1}, \ldots, r_{i_k}\} = M_\cup^-(F)$. Because $\omega(F) > 0$, we obtain $|M_\cup^-| \geq 2$. Let $r_{i_1}, r_{i_2} \in M_\cup^-$. If $F \cap R_2 = \emptyset$, $\{r_{i_1 i_2}, r_{j i_1}, j \in \{i_2, \ldots, i_k\}\} \subset M_\cup^+(F)$. If $F \cap R_2 \neq \emptyset$, then there exists exactly one $j \notin \{i_1, \ldots, i_k\}$ with $F \cap R_2 \subset \{r_{ji}, i \in \{1, \ldots, n\} \setminus \{j\}\}$ and $\{r_{ji}, i \in \{i_1, \ldots, i_k\}\} \subset M_\cup^+(F)$. Thus the claim follows.

If the maximum in ω is reached in more than one value $\omega(F) = \omega_r(F) = \omega_b(F)$ then $M_\cup^- = \emptyset$ and the claim follows trivially. The other cases follow by symmetry. $\square$

Now we consider the transitions, that remove a vertex from a coloring. Let

$$
\begin{array}{rcl}
M_{\smallsetminus}(F) & = & \{v \in V_n | \omega(F \smallsetminus \{v\}) \neq \omega(F)\} \\
M_{\smallsetminus}^{+}(F) & = & \{v \in V_n | \omega(F \smallsetminus \{v\}) > \omega(F)\} \\
M_{\smallsetminus}^{-}(F) & = & \{v \in V_n | \omega(F \smallsetminus \{v\}) < \omega(F)\}.
\end{array}
$$

($F \smallsetminus \{v\}$ denotes the set of colored vertices obtained by uncoloring v in the given coloring f.)

Lemma 8 *Suppose $\omega(F) < n/12$. If $M_{\smallsetminus}^{-} \neq \emptyset$ is fulfilled, then $\sharp M_{\smallsetminus}^{-} \leq n/36$ and $\sharp M_{\smallsetminus}^{+} \geq 11/12n - 1$ follows.*

Proof: First we suppose $\omega(F) = \omega_r(F) \wedge \omega(F) > \omega_b(F) \wedge \omega(F) > \omega_g(F)$. Then $\sharp M_{\smallsetminus}^{-} = |F \cap R_2| \leq n/36$. Moreover we obtain $M_{\smallsetminus}^{+} \supset F \cap R_1$ from the construction of G_n and $\omega(F) < n/12$ yields $\sharp\{\text{uncolored vertices in } R_1\} \leq n/12 + 1$. Then follows $\sharp M_{\smallsetminus}^{+} \geq \sharp F \cap R_1 \geq 11/12n - 1$ as required. If $\omega(F) = \omega_r(F) = \omega_b(F)$, then $M_{\smallsetminus}^{-} = \emptyset$ and the Lemma follows trivially. The remaining cases follow by symmetry. $\square$

Combining these two Lemmata we will now show, that the expected value of the difference $\omega(X_{t+1} - \omega(X_t)$, given either a $\sharp F$-increasing or -decreasing move, is separated from 0.

Lemma 9 *Let $0 < \omega(F) < n/12$. Then*

1. *$\sum_{\cup}(F) = |M_{\cup}(F)|^{-1} \sum_{v \in M_{\cup}(F)}(\omega(F \cup \{v\}) - \omega(F)) \geq 1$, if $M_{\cup}(F) \neq \emptyset$.*
2. *$\sum_{\smallsetminus}(F) = |M_{\smallsetminus}(F)|^{-1} \sum_{v \in M_{\smallsetminus}(F)}(\omega(F \smallsetminus \{v\}) - \omega(F)) \geq 3/4$, if $M_{\smallsetminus}(F) \neq \emptyset$.*

Proof: Straightforward calculation with the help of Lemma 7 and 8:

1.

$$
\begin{array}{rcl}
\sum\limits_{\cup}(F) & = & |M_{\cup}(F)|^{-1} \left(\sum\limits_{v \in M_{\cup}^{-}(F)} (\omega(F \cup \{v\})) - \omega(F) \right. \\
& & \left. \sum\limits_{v \in M_{\cup}^{+}(F)} (\omega(F \cup \{v\}) - \omega(F)) \right) \\
& \geq & |M_{\cup}(F)|^{-1} \left((-1)|M_{\cup}^{-}(F)| + 3|M_{\cup}^{+}(F)|\right) \\
& \geq & 1
\end{array}
$$

2. analogously

$\square$

Finally we can prove the main theorem.

Theorem 10 *Let* $H_m = \min\{t \in \mathbb{N} | X_t = \textit{maximal 3-coloring}\}$. *Then* $E(H_m | X_0 = \emptyset) = \exp(\Omega(n))$.

Proof: The argumentation is analogous to that in the proof of Theorem 1. First we define the points of time Z_j, when the value ω of the actual coloring changes: $Z_0 = 0$ and for all $j \in \mathbb{N} \quad Z_j = \min\{j > Z_{j-1} | \omega(X_j) \neq \omega(X_{Z_{j-1}})\}$. Suppose f is a 3-coloring with $0 < \omega(F) < n/12$. Let $M_t = +$, if the tth transition colors an uncolored vertex, $M_t = -$, if a vertex is removed from the coloring and $M_t = 0$, if the chain stays in the same state. We get with Lemma 9 for all $j \in \mathbb{N}, f \in \Omega$ with $0 < \omega(F) < n/12$:

$$E(\omega(X_{Z_{j+1}}) - \omega(X_{Z_j}) | X_{Z_{j+1}-1} = f, M_{Z_{j+1}-1} = +, X_{Z_{j+1}-2}, \ldots, X_0) \geq 1$$

and analogously

$$E(\omega(X_{Z_{j+1}}) - \omega(X_{Z_j}) | X_{Z_{j+1}-1} = f, M_{Z_{j+1}-1} = -, X_{Z_{j+1}-2}, \ldots, X_0) \geq 3/4$$

By the summation of the disjoint, given events we get:

$$E(\omega(X_{Z_{j+1}}) - \omega(X_{Z_j}) - 3/4 | 0 < \omega(X_{Z_j}) < n/12, X_{Z_j}, \ldots, X_{Z_0}) \geq 0$$

due to

$$E(f | \dot{\cup} A_i) = \frac{\sum_i P(A_i) E(f | A_i)}{P(\dot{\cup} A_i)}$$

and it follows that

$$E(\omega(X_{Z_{j+1}}) - \omega(X_{Z_j}) - 3/4 | 0 < \omega(X_{Z_j}) < n/12, \omega(X_{Z_j}) = i_j, \ldots, \omega(X_{z_0}) = i_0)$$

is not negative. With $Y_j = \omega(X_{Z_j}), a = n/12, \epsilon = 3/4$ and $H_0 = \min\{t | y_t = 0\}$ we obtain

$$E(Y_{j+1} - Y_j - \epsilon, Y_j < a, H_0 > j | Y_j = i_j, \ldots, Y_0 = i_0) \geq 0$$

as required for condition C1 of Theorem 2. Due to $|Y_{j+1} - Y_j| \leq 3$ condition C2 is also fulfilled and it follows

$$\exists D, \eta \in \mathbb{R} \quad \forall i_0 \geq a, t \in \mathbb{N} \qquad P(H_0 = t | Y_0 = i_o) \leq D e^{\eta(-a)} =: \delta$$

Using the fact that $\omega(\emptyset) = n - 1$ and $X_0 = \emptyset$ we get

$$P(H_0 = t | Y_0 = n - 1) = P(H_0 = t | X_0 = \emptyset)$$

and

$$P(H_0 > t|X_0 = \emptyset) = 1 - \sum_{k=1}^{t} P(H_0 = k|X_0 = \emptyset) \geq \max\{0, 1 - t\delta\}$$

(We could assume $P(H_0 < \infty|X_0 = \emptyset) = 1$, because otherways the proof of the claim would be trivial). Combining the last results yields

$$E(H_0|X_0 = \emptyset) = \sum_{t=0}^{\infty} P(H_0 > t|X_0 = \emptyset) \geq \sum_{t=0}^{\infty} \max\{0, 1 - t\delta\} \geq \frac{1}{2\delta}.$$

Because $X_t =$ (maximal 3-coloring) implies $\omega(X_t) = 0$, it follows

$$E(H_m|X_0 = \emptyset) \geq \frac{1}{2\delta} = De^{\eta m}$$

as required. □

4 Closing Remarks

We have showed showed that there is no hope to further improve the exponential running time bounds for Simulated Annealing.

As already pointed out, we suspect that an exponential first hitting time can be established in the case of the first approach even for arbitrary cooling schedules and not only for the Metropolis process with constant temperature. A similar definition of ω with constant values on the energy levels as in the second approach does not seem to be applicable, since the difference $c(f) - c(f')$ of two neighboring colorings is not bounded by a constant.

Despite the negative results in this paper we believe, that the performance of Simulated Annealing on certain random graphs is much better than on the graphs considered in this paper. The reason for this is the smooth structure of these graphs, the ratio of optimal to near-optimal solutions is not so extremely bad as here. The authors have succeeded to show convergence of the Metropolis process on certain random graphs with high probability [15], but the proof, that Annealing is beneficial for some combinatorial optimization problem, is still missing.

References

[1] Anily, S. and Federgruen, A.:*Ergodicity in Parametric Nonstationary Markov Chains: An Application to Simulated Annealing Methods*, Oper. Res. 35, 1987

[2] Garey, M.R.; Johnson, S.J.: *Computers and Intractability*, W.H. Freeman and Company, 1979

[3] Gidas, B.: *Nonstationary Markov Chains and Convergence of the Annealing Algorithm*, J. Statis. Phys. 39, 1985

[4] Hajek, B.: *Hitting Time and Occupation -Time Bounds Implied By Drift Analysis With Applications*, Advanced Applied Probability 14, 1982

[5] Hajek, B.: *Cooling Schedules for Optimal Annealing*, Mathematics of Operations Research 13, 1988

[6] Jerrum, M.: *Large Cliques elude the Metropolis Process*, Random Structures and Algorithms 3, 1992

[7] Johnson, D.S. et al.: *Optimization by Simulated Annealing: An Experimental Evaluation*, Operations Research 39, 1991

[8] Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P.:*Optimization by Simulated Annealing*, Science 220, 1983.

[9] Laarhoven, P.J.M.; Aarts, E.H.L.: *Simulated Annealing: Theory and Applications*, Kluwer Academic Publishers, 1989.

[10] Metropolis, N. et al.: *Equations of state calculations by fast computer machines*, Journal of Chemical Physics 21, 1953

[11] Mitra, D.; Romeo, F.; Sangiovanni-Vincentelli, A.L.:*Convergence and Finite Time Behaviour of Simulated Annealing*, Adv. Appl. Prob. 18, 1986

[12] Motwani, R.; Raghavan, P.:*Randomized Algorithms*, Cambridge University Press, 1995

[13] Nolte, A.; Schrader, R.: *A Note on the Finite Time Behaviour of Simulated Annealing*, Mathematics of Operations Research, to appear

[14] Nolte, A.; Schrader, R.: *Simulated Annealing and its Problems to Color Graphs*, Proceedings of the ESA 1996, Springer Lecture Notes in Computer Science 1136

[15] Nolte, A.; Schrader, R.: *Coloring in Sublinear Time*, Proceedings of the ESA 1997, Springer Lecture Notes in Computer Science 1284

[16] Sasaki, G.H.; Hajek, B.: *The Time Complexity of Maximum Matching by Simulated Annealing*, Journal of the Association for Computing Machinery 35, 1988

[17] Tsitsiklis, J.: *Markov Chains with Rare Transitions and Simulated Annealing*, Mathematics of Operations Research 14, 1989