Scaling behaviour of optimal simulated annealing schedules

1993 J. Phys. A: Math. Gen. 26 3267

(http://iopscience.iop.org/0305-4470/26/13/028)

J. Phys. A: Math. Gen. 26 (1993) 3267–3277. Printed in the UK

Scaling behaviour of optimal simulated annealing schedules

Michael Christoph† and Karl Heinz Hoffmann

Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 19, D-W-6900 Heidelberg 1, Federal Republic of Germany

Received 25 March 1992, in final form 10 September 1992

Abstract. The success of simulated annealing depends strongly upon the choice of a suitable annealing schedule. For a class of small sample systems the optimal annealing schedules are determined. They show distinct scaling behaviour as a function of the number of Metropolis steps carried out at each temperature of the schedule. This behaviour can be traced back to the influence of dominating barriers during cooling. Knowing the optimal schedule for a few different total annealing steps allows to predict the optimal annealing schedule for intermediate times.

1. Introduction

Simulated annealing is a technique that has attracted significant attention due to its suitability for finding near-optimal solutions to complex optimization problems [1]. Complex optimization problems are characterized by an objective function with many local minima. For such problems, conventional methods like steepest descend [2] fail and simulated annealing (and other techniques, for instance genetic algorithms [3–5]) have shown significant improvements for a variety of problems such as the travelling salesman problem [6, 7], graph partitioning [8], or chip-design [1, 9–11]. Simulated annealing was conceived as an analogy to the annealing of a real physical system: just as careful annealing of a real physical system should bring it into its equilibrium state with the ambient temperature T and thus with T towards zero to its groundstate(s), the proper simulation of this procedure for an optimization problem should result in finding its optimal solution.

The implementation of a simulated annealing algorithm incurs costs, mainly in the form of computer time. So, from the very beginning, the aim has been to lower these costs by optimizing the simulated annealing procedure itself [10, 12–14]. Attempts have been made to improve the performance of simulated annealing by changing the algorithm in such a way as to avoid rejection of new states [15], or by choosing special move classes [16]. There is no doubt, however, that the success of the simulated annealing procedure depends strongly upon the choice of a suitable annealing schedule. Thus optimization of this part of the method seems to be an obvious approach, the central question being as follows. *How can the schedule of an optimal annealing procedure be characterized*?

† Present address: Institut für Geophysik und Meteorologie der Universität zu Köln, Kerpenerstrasse 13, D-W-5000 Köln 41, Federal Republic of Germany.

0305-4470/93/133267+11$07.50

Naturally the optimal schedule will depend on a number of different parameters. In this paper we present our findings about the dependence on one of these parameters, namely the total available annealing time. For applications this is a rather crucial parameter as it translates directly into computer time needed. We find that the optimal schedules of the systems considered here have a distinct scaling behaviour. Knowing the optimal schedule for a few different annealing times this scaling behaviour can be used to predict the optimal schedule for all intermediate times.

2. The optimal annealing schedule

The problem of finding the optimal annealing schedule first requires the definition of a measure of 'optimality' [17, 18]. There are several possible criteria, for instance maximizing the probability of being in the ground state, minimizing the BSF (best-so-far) energy [19] or minimizing the mean final energy. Here we use the mean final energy, which from the viewpoint of statistical mechanics is the most natural choice.

Consider a system of m states with energies $\boldsymbol{E}=(E_1,\ldots,E_m)$. Usually simulated annealing is thought of as an algorithm creating a random path in the state space of the system. When discussing certain statistical properties of simulated annealing an approach along these lines requires very many simulation runs to obtain data with reasonable error margins. Instead we opted to use the appropriate theoretical description in form of a time discrete Markov process (Markov chain). In this description the complete dynamical evolution of the probability $p_\alpha(i)$ of being in state α after Metropolis step i can be given in an analytical form. It is described by the transition probability matrix $\mathbf{G}(x)$ which contains the Arrhenius factors through which the temperature T enters the dynamics. For brevity we introduce a new parameter $x=e^{-1/T}$ which we will also refer to as 'temperature'. To be precise we define an infinite temperature transition matrix $\mathbf{G}(1)=\boldsymbol{\Pi}=(\Pi_{\beta\alpha})$ between states α and β:

$$\Pi_{\beta\alpha}=\begin{cases}0 & \text{if } \beta\notin N(\alpha)\\ \dfrac{1}{|N(\alpha)|} & \text{if } \beta\in N(\alpha)\end{cases} \tag{1}$$

where $N(\alpha)$ is the set of neighbours of α. For the systems considered below it is given by the connectivity of nodes in figure 2 or figure 6. For finite temperatures the transition matrix $\mathbf{G}(x)$ is obtained from $\boldsymbol{\Pi}$ by:

$$G_{\beta\alpha}=\begin{cases}\Pi_{\beta\alpha}\exp(-\Delta E/T) & \text{if } \Delta E>0,\ \alpha\neq\beta\\ \Pi_{\beta\alpha} & \text{if } \Delta E\leqslant 0,\ \alpha\neq\beta\\ 1-\sum\limits_{\xi\neq\alpha}G_{\xi\alpha} & \text{if } \alpha=\beta\end{cases} \tag{2}$$

where $\Delta E=E(\beta)-E(\alpha)$. The transition matrix of the two level system considered below is slightly different (cf section 4).

Let $\boldsymbol{p}(0)$ be the initial probability distribution over the state space. The probability *distribution $\boldsymbol{p}(i)$ after Metropolis step i is calculated from the previous distribution* $\boldsymbol{p}(i-1)$ by $\boldsymbol{p}(i)=\mathbf{G}(x_i)\boldsymbol{p}(i-1)$. The mean energy after step i is given by $\boldsymbol{E}\boldsymbol{p}(i)$. The optimal annealing schedule with respect to minimizing the mean final energy in a finite number N of Metropolis steps is then calculated by

$$\min\bar{E}=\min_{\{x_i\}_{i=1,\ldots,N}}\boldsymbol{E}\mathbf{G}(x_N)\mathbf{G}(x_{N-1})\ldots\mathbf{G}(x_2)\mathbf{G}(x_1)\boldsymbol{p}(0). \tag{3}$$

In the usual application of simulated annealing the temperature is kept fixed for a number of Metropolis steps before it is lowered again. Here we follow the same scheme. We calculate the optimal annealing schedule with respect to the above-defined optimization criterion for $N = j^* \nu$ Metropolis steps. This means in our case that we optimize j temperatue steps and that the system is allowed to spend ν Metropolis steps at each temperature $x_{n=1,\ldots,j}$ before lowering the temperature again. And, knowing that this is certainly true of any optimal schedule [17], we always set the temperature of the final step equal to zero. Thus the final form of the optimization problem is

$$\min \bar{E} = \min_{\{x_n\}_{n=1,\ldots,j}} E\mathbf{G}(0)\mathbf{G}^{\nu}(x_n)\mathbf{G}^{\nu}(x_{n-1}) \ldots \mathbf{G}^{\nu}(x_2)\mathbf{G}^{\nu}(x_1)\boldsymbol{p}(0). \quad (4)$$

This optimization problem is solved numerically by using a conjugated gradient approach based on an algorithm given in *Numerical Recipes* [20]. We would like to point out that we did not use simulated annealing to solve (4) although this could be done in principle. Instead we found fast convergence with the conjugate gradient algorithm using several different starting values.

$\bar{E}$ is minimized by the technique of Polak and Ribiere (cf *Numerical Recipes* [20]). This algorithm requires the calculation of partial derivatives of $\bar{E}$ with respect to the variables x_n. Problems arose because of the constraints that had to be met (temperature $x \in [0, 1]$). The inequality constraints can be treated by performing the line minimization exclusively within a j-dimensional cube whose corners R_s have the following coordinates:

$$R_s = (x_1, x_2, \ldots, x_j),\ s = 1, \ldots, 2^j \text{ where } x_n = 0, 1 \quad n = 1, \ldots, j. \quad (5)$$

The algorithm proceeds as follows. A starting point $X = (x_1, x_2, \ldots, x_j)$ in the j-dimensional space is selected. Then the value and the derivatives of the objective function are calculated from their analytical formulae. From there, one proceeds in a vector direction determined by the conjugate gradient method. Next, using the golden-section method, the function is minimized along the line beginning at point X and ending on the surface of the j-dimensional cube. The minimum along this line becomes the starting point for the next iteration. Convergence is achieved when $|\nabla \bar{E}|^2$ is less than an input tolerance value.

3. Description of systems

We study the optimal annealing schedule for three different systems: two with a tree structure and then a two-level system. The use of tree structures is motivated by the following consideration. Generally speaking, optimization problems are characterized by having a complex state-space structure often due to various constraints which have to be met. These constraints introduce a large number of local minima, often on the order of e^N, where N is the number of degrees of freedom. These minima are separated by barriers of different height and the necessity to overcome those barriers to reach one local minimum from another constitutes the complexity of the problem. This essential state space structure is preserved under a lumping technique recently introduced by Hoffmann and Sibani [21] which allows to reduce the enormous number of states. A new coarse-grained state space shaped like a 'tree' is generated from an energy landscape; see figure 1. The nodes of the tree consist of subsets of 'microscopic' states lumped together. Then for low temperatures and rough energy surface the original dynamics according to the METROPOLIS algorithm can be mapped approximately into a random walk on the tree structure. Essentially, the simulated annealing procedure

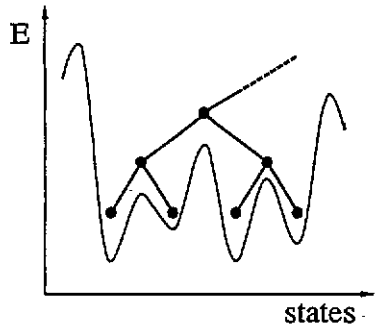

Figure 1. Tree construction from an energy landscape.

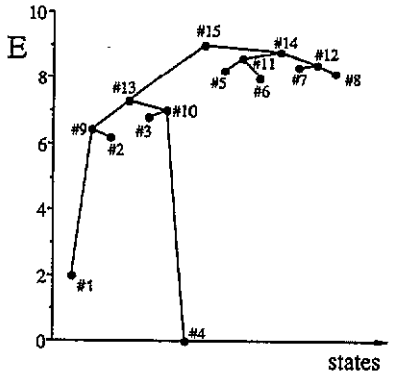

Figure 2. System 1: the connectivity between the nodes is given, while the energy differences are assigned randomly.

can be described as 'hopping over barriers', so the tree model can be taken to be a good approximation of the original state space.

The first model, System 1, has a simple state space comprised of 15 states. The size of the model is a compromise between computational limits and the number of barriers desired. The states of the system are organized in a tree-like way. The move class, i.e. the sets of neighbours, of the system is determined by the bonds in figure 2. Here the energy values were chosen randomly.

The second model, System 2, is designed as shown in figure 6. Again it is a system consisting of 15 states with the same move class as System 1. The main feature of this tree is that for each 'parent' node the ratio of the barrier heights viewed from the two 'daughter' nodes is always two. This choice is motivated by expected self-similarities in certain optimization problems [22].

Finally we consider a system of two levels separated by a barrier. The system has activation energies D^* and D as viewed from the ground state and the high state, respectively. This system is studied as it allows an analytical treatment which provides some insight into the behaviour of Systems 1 and 2.

4. Results

System 1

Solving (4) yields optimal annealing schedules shown in figure 3. We optimize $j = 30$ temperature steps and plot the temperature x versus the step number n while varying the number of equal-temperature steps ν between 50 and 600. Here and in all the following computations the initial probability distribution $p(0)$ is taken to be uniform. Optimal annealing schedules are thus obtained for a total number N of Metropolis steps ranging between 1500 and 18 000. A decrease in both the starting temperature and the plateau of the annealing schedule is observed for an increasing number of equal-temperature steps ν.

The outstanding feature contained in the observed data is the *scaling behaviour* shown by the annealing schedules of System 1. This is demonstrated in figure 4. At step $n = 15$, for example, we take the temperature $x_{n=15}$ of each optimal schedule in

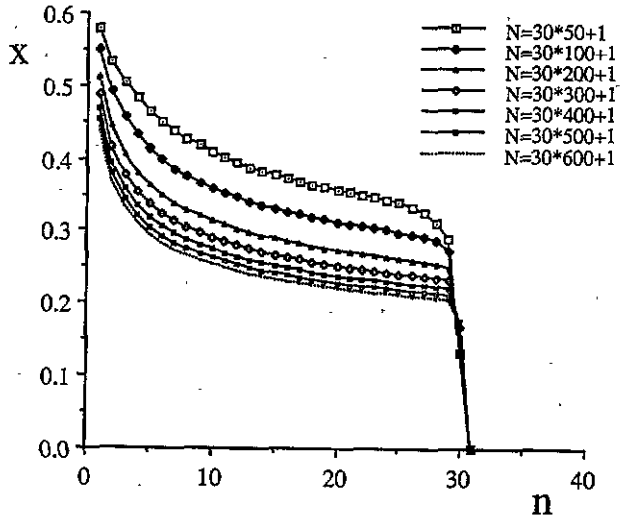

Figure 3. Optimal annealing schedules of System 1. Temperature x plotted versus temperature step n while varying ν between 50 and 600.

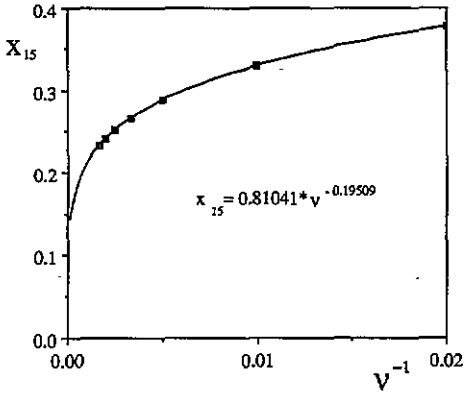

Figure 4. Scaling behaviour of optimal annealing schedules (System 1) shown at step $n = 15$.

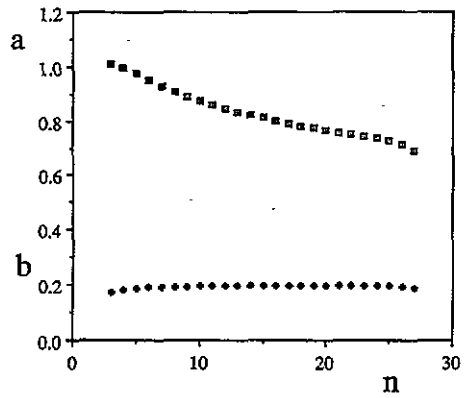

Figure 5. Scaling coefficients a and b as a function of step number n.

figure 3, plot them versus ν^{-1}, and fit the data with a curve whose equation is given by

$$x_n = a_n \nu^{-b_n} \tag{6}$$

with $a_{n=15} = 0.810\,41$ and $b_{n=15} = 0.195\,09$. Such scaling behaviour is also clearly to be seen at each temperature step $n = 3, \ldots, 27$. The data of the first two temperature steps and steps $n = 28, 29$ cannot be fitted exactly, and for $n = 30$ a fit with (6) fails completely.

In figure 5 the scaling coefficients a and b are plotted versus the temperature step $n = 3, \ldots, 27$. One sees that coefficient a decreases with increasing temperature step number n while coefficient b remains approximately constant. The data of coefficient a are fitted well by

$$a_n = 1.3122 n^{-0.18007} \tag{7}$$

and for coefficient b a linear curve fit yields

$$b_n = 0.18865 + 0.0003 n. \tag{8}$$

Having observed scaling behaviour in System 1 the question arises whether this behaviour can also be found in systems with a certain self-similarity.

System 2

The optimal annealing schedules of System 2 computed with (4) are shown in figure 7. Again we optimize $j = 30$ temperature steps and plot temperature x versus step number n while varying the number of same-temperature steps ν between 80 and 300. Note that this time the optimal schedules consist of three different plateaux instead of a single one as in System 1. The first jump in temperature x takes place at step $n \approx 21$ and the second jump at $n \approx 28$.

Again we find significant scaling behaviour in System 2 and the coefficients a and b as a function of step number n are shown in figure 8. The evolution of coefficients a and b with increasing temperature step n, i.e. with time, reveals the following

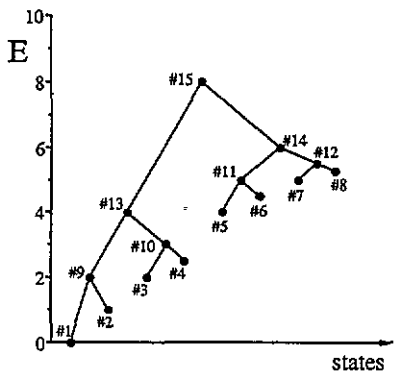

Figure 6. System 2: the main feature of this system is that for each parent node the ratio of the barrier heights viewed from the two daughter nodes is always two.

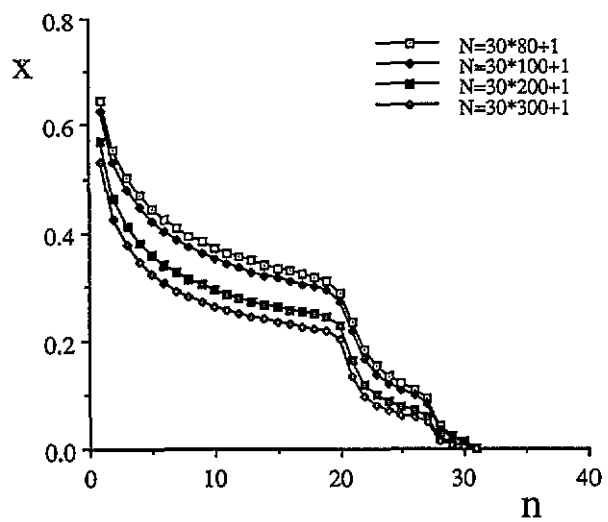

Figure 7. Optimal annealing schedules of System 2. Temperature x plotted versus temperature step n while varying ν between 80 and 300.

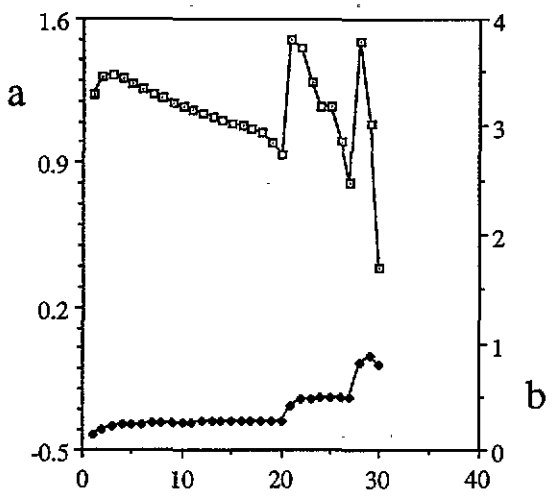

Figure 8. *Scaling coefficient **a** and **b** as a function of step number n.*

remarkable characteristics. Coefficient ***a*** decreases with increasing n, and at temperature step $n=21$ coefficient ***a*** jumps to a significantly higher value before continuing its decrease. At step $n=28$ the coefficient jumps a second time and then decreases faster than in the previous two instances. Coefficient ***b*** behaves similarly: it remains approximately constant up to step $n=20$, then jumps to a higher value and remains constant until step $n=28$ is reached, at which point the second jump to a even higher value takes place.

Due to these characteristics, the data of coefficient ***b*** can be fitted linearly for each section of the graph. Note that the n-dependence for ***b*** is very weak.

$$b_{1,n}=0.2063+0.004\,13n \qquad n=1,\ldots,20 \tag{9a}$$

$$b_{2,n}=0.4404+0.010\,47n \qquad n=21,\ldots,27 \tag{9b}$$

$$b_{3,n}=0.8591-0.013\,21n \qquad n=28,29,30. \tag{9c}$$

Two-level system

Here we optimize only a single temperature step and vary the number of equal-temperature steps ν between 1000 and 15 000. The results plotted in figure 9 are based on the following system parameters. The energy of the ground state and the high state are taken to be zero and one, respectively. Activation energies D^* and D are equal to three and to two, respectively. In this case the optimal annealing schedule is only a function of ν and is represented by a single value of temperature x. Again scaling behaviour is observed:

$$x=1.3061\nu^{-0.43793}. \tag{10}$$

The optimal annealing schedule of this system can also be determined analytically within certain approximations. Schedules with more than one temperature can be obtained iteratively from the following single temperature step optimization. The probability distribution $p(n)$ after ν equal-temperature steps with temperature x_n is

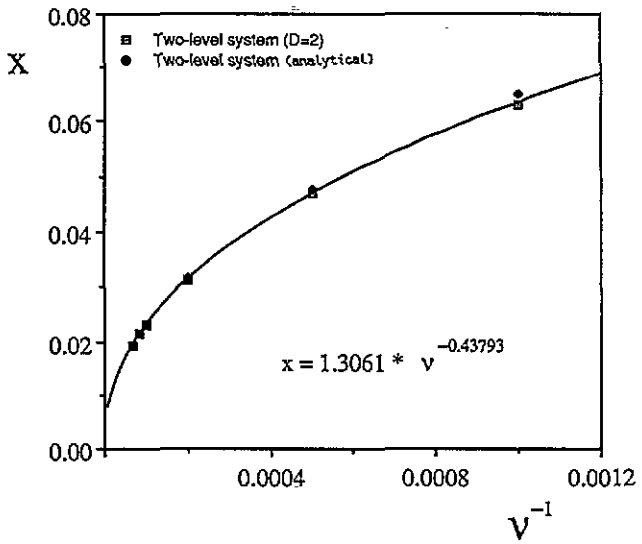

Figure 9. Optimal annealing schedules and scaling behaviour of the two level system (with $n=j=1$ and $\nu=1000$, 2000, 5000, 10 000, 12 000, 15 000) and corresponding analytical data predicted from (13).

calculated from distribution $p(n-1)$ at the previous temperature by $p(n)=\tilde{\mathbf{G}}(x_n)^{\nu}p(n-1)$, with

$$\tilde{\mathbf{G}}(x_n)=\begin{pmatrix}1-x_n^{D^*} & x_n^{D}\\ x_n^{D^*} & 1-x_n^{D}\end{pmatrix}. \tag{11}$$

Let $p_2(n)$ be the probability of being in the high state. Going into the eigenbasis of $\tilde{\mathbf{G}}(x_n)$, it can be written as:

$$p_2(n)=\lambda^{\nu}p_2(n-1)+(1-\lambda^{\nu})\frac{x_n}{x_n+1} \tag{12}$$

with the second largest eigenvalue $\lambda(x_n)=1-x_n^{D}-x_n^{D^*}$. Here D^* and D represent the activation energy as viewed from the ground state and the activation energy as viewed from the high state, respectively.

According to the optimization criterion (4) one is interested in finding the optimal temperature $x_n(\nu)$ which minimizes $\bar{E}$ over the interval [0, 1]. Setting the ground-state energy of the two-level system to zero shows that minimizing $\bar{E}$ is equivalent to minimizing $p_2(n)$. Differentiating (12) with respect to x_n and setting it equal to zero results in the transcendental equation for the optimal x_n

$$p_2(n)D\nu^{1/D}z^{(D-1)/D}=\mathrm{e}^{z} \tag{13}$$

with $z=x_n^{1/\delta}\nu$, $\delta=1/D$ and assuming ν being large enough. For an interval around solution $z_0=x_{n,0}^{1/\delta}\nu_0$ of (13) one obtains the approximate solution:

$$x_n(\nu)=\nu^{-\delta}(z_0+\ln(\nu/\nu_0)^{\delta})^{\delta}. \tag{14}$$

By comparing the coefficients of (14) with those of (6) one obtains the scaling coefficient b:

$$b=\delta=1/D. \tag{15}$$

We calculate $x(\nu)$ for various ν with $\nu_0 = 10\,000$ as a reference point, $D = 2$, and $x_0(\nu_0) = 0.023\,1456$ (obtained numerically from (4)) and find very good agreement with (14) as can be seen in figure 9.

Finally we calculate the optimal annealing schedule of a two state system ($j = 30$, $\nu = 150$) with a barrier height ratio $D^*/D = 7.28/5.28$ (this is also the barrier height ratio between states 1 and 4 of System 1) and compare it with the optimal annealing schedule of System 1 ($j = 30$, $\nu = 600$) by taking into account a rescaling of 'time' ν: while in System 1 four steps are necessary to overcome the barrier, only one step is needed in the two-level system. Figure 10 shows clearly that the two schedules coincide well.

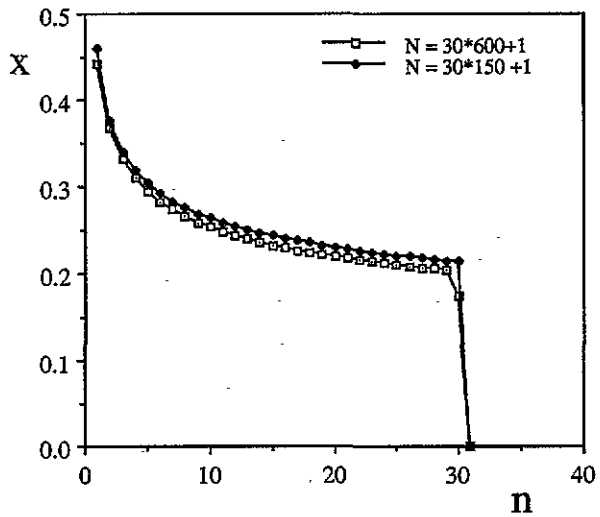

Figure 10. Comparison between the annealing schedule of System 1 ($j = 30$, $\nu = 600$) and the annealing schedule of the two-level system ($j = 30$, $\nu = 150$) whose barrier height ratio $D^*/D = 7.28/5.28$ corresponds to the one of the 'dominating' barrier of System 1.

5. Conclusions

Our calculations show that the optimal annealing schedules follow a power law $x(\nu) \sim \nu^{-b}$ within a finite interval of ν. We find for the two-level system with activation energy D that $b = 1/D$, with corrections due to logarithmic terms as indicated in (14). For larger D the corrections become smaller.

A comparison between the optimal annealing schedule of System 1 and the one of a two-level system suggests that System 1 behaves like a simple two-level system. The fact that the scaling coefficient $b = 0.188$ of System 1 (see figure 5) is in very good agreement with $b = 1/5.28 = 0.189$, where $D = 5.28$ is the activation energy needed to get from state 1 into the ground state, supports this conclusion. Moreover it shows that the scaling behaviour is dominated by the largest barrier provided enough time ν spent at each temperature step (from numerical experience $\geqslant 50$ for the systems considered here). The number of equal-temperature steps ν being large will ensure that all the fast eigenmodes will decay quickly except for the slowest eigenmode and

the stationary one. In this case optimizing the annealing schedule is equivalent to optimizing only the slowest eigenmode.

This is further supported by the following intriguing feature of the optimal schedule. The temperature decreases approximately logarithmically with

$$T_n = \frac{A}{\ln(\nu) + B\ln(n) - C}. \tag{16}$$

Inserting (7) into (6) and comparing it with (16) gives

$$A = 1/b \approx D \qquad B = 0.180\,07/b = 0.953 \qquad C = \ln(1.3122)/b = \ln(4.21). \tag{17}$$

Approximating B by 1 and C by ln(4) the optimal schedule behaves roughly as

$$T_n = \frac{D}{\ln(n\nu/4)}. \tag{18}$$

This dependence reminds very much of the well known schedule which guarantees that the global minimum is achieved with probability one [12]. However, such a dependence has also been found for the long-time dependence of the optimal schedule for a two-level system [18]. Again the 4 in the denominator corresponds to a rescaling of 'time' as in System 1 four steps are necessary to overcome the barrier between states 1 and 4, while only one step is needed in the two-level system. Regarding the approximations made note that the optimization considered in [18] allowed a continuous temperature change while here only a step function is allowed. Nonetheless this observation provides additional evidence that at any time the optimal schedule calculated here is governed by one dominating barrier.

The optimal schedules of System 2 consist of three different plateaux. The scaling coefficient b is a different constant for each plateau and each time corresponds to a different barrier to overcome. So, when most of the probability is shovelled over the currently dominating barrier the temperature drops to a new plateau since then the schedule is dominated by the next barrier.

This picture is supported by a thorough eigenmode analysis. We determined all eigenvalues and eigenvectors for a large number of temperature values along the optimal schedule. For those temperature values we were thus able to write down an analytic solution of the dynamics. This enabled us to study the progress of the annealing in detail. For instance the first section of the schedule is dominated basically by the transition from state 5 to the groundstate 1 over a barrier of height $D = 4$. The exponent $b_1 = 0.2063$ corresponds approximately to the expected value $1/D = 0.25$.

As already pointed out the observed scaling behaviour $x_n(\nu) = a_n \nu^{-b_n}$ is only valid within a finite interval of ν. Knowing, however, a few optimal schedules within a limited ν interval the exponent b_n as well as the constant of proportionality can be obtained from a power law fit and thus the optimal schedule for all intermediate ν can be interpolated.

Finally we would like to mention that in principle the n dependence of the constant of proportionality a_n can also be obtained from (13) provided the dominating barrier is unchanged.

Even though the systems studied here are quite small and thus too simple to encompass the complexity of large-scale simulated annealing problems, this work is a first step towards understanding the dependencies of optimal simulated annealing schedules. Further investigations of more complex systems and thus finding out more

about the characteristics relevant to the optimal schedule remains an important task of future research.

Acknowledgments

The authors would like to acknowledge the stimulating environment of the Telluride Summer Research Center, Telluride, Co, USA where part of this work was done.

References

[1] Kirkpatrick S, Gelatt Jr C D, Vecchi M P 1983 Optimization by simulated annealing *Science* **220** 671
[2] Lin S, Kernighan B W 1973 *Oper. Res.* **21** 498
[3] Ablay P 1987 Optimieren mit Evolutionsstrategien *Spektrum der Wissenschaft* July
[4] Wang Q 1987 Optimization by simulating molecular evolution *Biol. Cybern.* **57** 95–101
[5] Davis L 1987 *Genetic Algorithms and Simulated Annealing* (London: Pitman)
[6] Bonomi E and Lutton J L 1984 *SIAM Rev.* **26** 551
[7] Durbin R and Willshaw D 1987 *Nature* **326** 689
[8] Salamon P, Nulton J, Harland J R and Pedersen J 1988 Simulated annealing with constant thermodynamic speed *Computer Physics Comm.* **49** 423 (Amsterdam: North-Holland)
[9] Slarry P and Dreyfus G 1983 *J. Physique Lett.* **45** L39
[10] White S 1984 Concepts of scale in simulated annealing *IEEE Proc. Int. Conf. on Computer Design* p 646
[11] Vecchi M P and Kirkpatrick S 1982 *Research Report* RC9555 (IBM Thomas J Watson Research Center, Yorktown Heights, New York)
[12] Geman S and Geman D 1984 Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images *IEEE PAMI* **6** 721
[13] Huang M D, Romeo F and Sangiovanni-Vincentelli A 1986 An efficient cooling schedule for simulated annealing *IEEE Proc. ICAAD* p 381
[14] Morgenstern I and Würtz D 1987 Simulated annealing for 'spin-glass-like' optimization problems *Z. Phys. B. Condensed Matter* **67** 397
[15] Greene J W and Supowit K J 1984 Simulated annealing without rejected moves *IEEE Proc. ICCD* p 658
[16] Szu H and Hartley R 1987 Fast simulated annealing *Phys. Lett.* **122A** 157
[17] Hoffmann K H 1988 Bounds and optima for irreversible thermodynamic processes and their application to simulated annealing *Habilitationsschrift* Universität Heidelberg
[18] Hoffmann K H and Salamom P 1990 *J. Phys. A: Math. Gen.* **23** 3511
[19] Hoffmann K H, Sibani P, Pedersen J M and Salamon P 1990 Optimal ensemble size for parallel implementation of simulated annealing *Appl. Math. Lett.* **3** 53
[20] Press W H, Flanner B P, Teukolsky S A and Vetterling W T 1986 *Numerical Recipes* (Cambridge: Cambridge University Press)
[21] Hoffmann K H and Sibani P 1988 *Diffusion in hierarchies Phys. Rev.* A **38** 4261
[22] Sibani P, Pedersen J, Hoffmann K H and Salamon P 1990 *Phys. Rev.* A **42** 7080