BRIEF REPORTS

Brief Reports are accounts of completed research which do not warrant regular articles or the priority handling given to Rapid Communications; however, the same standards of scientific quality apply. (Addenda are included in Brief Reports.) A Brief Report may be no longer than 4 printed pages and must be accompanied by an abstract. The same publication schedule as for regular articles is followed, and page proofs are sent to authors.

Cooling-rate dependence of the ground-state energy using microcanonical simulated annealing

Pam Ocampo-Alfaro and Hong Guo

Department of Physics and Centre for the Physics of Materials, McGill University, Montréal, Québec, Canada H3A 2T8

(Received 19 June 1995)

We have studied the cooling-rate dependence of approaching the ground state in several spin glass models within the microcanonical simulated annealing scheme. The final state energies of the two-dimensional (2D) and 3D Gaussian models and the 3D $\pm J$ model were found to depend on the cooling rate $\Delta E/\tau$, where ΔE is the energy decrement in the microcanonical simulation and τ is the number of Monte Carlo steps per spin at each annealing stage. In particular the ground-state energy dependence on τ for all models studied shows a stretched exponential form for a wide range of the values τ. This behavior is very different from that of the conventional simulated annealing algorithms.

PACS number(s): 64.60.Cn, 75.10.Hk, 75.50.Kj, 05.50.+q

I. INTRODUCTION

Much progress has been achieved in recent years in understanding the behavior of complex systems that involve many parameters and conflicting constraints [1]. A central question of these problems is the search for the ground state where a measurable quantity obtains an extremal value. For example, in the traveling salesman problem one wishes to find the shortest path that connects all the nodes; while in the spin glass models one needs to compute the ground-state energy. A powerful tool for searching for the ground state was proposed in a classic paper by Kirkpatrick, Gelatt, and Vecchi in 1983 where the stochastic model of simulated annealing was introduced [2]. In this approach a cost function is constructed to characterize the optimization process. A controlled thermal treatment followed by slow cooling gives the system the chance to jump out of local minima of the cost function, thereby improving the capability of finding the global minimum.

The idea of simulated annealing is most commonly combined with the conventional Metropolis Monte Carlo simulation. For example, to compute the ground-state energy of a spin glass model, one starts at a high temperature T and anneals down to $T = 0$ in small steps ΔT. At each stage of the annealing schedule a prespecified number of Monte Carlo steps τ, which measures a time, are completed before the temperature is lowered further. Obviously one wishes to use the largest possible ΔT and smallest possible τ to achieve an accurate answer. However, as expected, the minimum energy found for spin glass models depends on the cooling rate $r \equiv \Delta T/\tau$. In particular [3] for two-dimensional (2D) models this dependence is a power law, while for 3D models it is inversely logarithmic. While the reason for these particular kinds of time dependence were not clear, the difference in the time dependence between the 2D and 3D models is believed to be related to the fact that 3D models are NP complete while the 2D models are not. These findings provide useful guidelines for the design of optimal annealing schedules for other applications. (NP refers to a category of problems that are believed to be insoluble by algorithms whose run time grows as a polynomial in the size of the problem.)

It is also possible to perform simulated annealing in other statistical ensembles. Although formulations of equilibrium statistical mechanics using different ensembles are equivalent, the dynamics of the algorithms based on these ensembles can be different. Thus it is possible to exploit different ensembles for optimal dynamical performance of a simulation. For example, multicanonical ensemble simulated annealing [4, 5] was recently applied successfully to compute ground-state properties of a realistic protein. In a previous work we have explored simulated annealing within the microcanonical ensemble [6] and found that this scheme can be advantageous for certain applications.

In this paper we present a detailed numerical study of the cooling rate dependence of ground-state energies of two- and three-dimensional Ising spin glass models, using microcanonical simulated annealing. Similar to what is found in the more conventional Metropolis simulation [3], we also find a dependence of the final energy on the cooling rate. However, the τ dependence of the approach

to the final state (which may or may not be the global minimum energy state) fits very well with a stretched exponential in all cases studied, from the initial time to some late time where the system energy compares well with that of the known values of the ground states. This behavior is very different from the conventional simulated annealing algorithms. Although there is a long tail in the stretched exponential decay (see below) that probably will alter the overall time dependence at very large times, for all practical purposes we believe that this algorithm has a "better" dependence on the annealing time as far as the approach to ground-state spin configuration is concerned. This is because a stretched exponential is a faster decaying function than a power law, or the inverse logarithmic dependence found in the canonical ensemble simulated annealing in the time regime of a typical simulation. In the next section a brief review of the microcanonical ensemble simulated annealing method is given. Section III presents results of our work. Section IV contains a short summary.

II. METHOD

A detailed discussion of the microcanonical ensemble simulated annealing used here can be found in Ref. [6], but for completeness we briefly review it [7]. Consider the nearest-neighbor Ising model described by the Hamiltonian

$$H = -\sum_{i \neq j} J_{ij} S_i S_j, \tag{1}$$

where the sum is over all nearest neighbors, J_{ij} is the exchange interaction strength between spins at site i and j, and $S_i = \pm 1$ is the spin at site i. Two popular Ising spin glass models are for J_{ij} to take random values from a Gaussian distribution (Gaussian model), or to take values of $\pm J$ with equal probability (the $\pm J$ model). When $J_{ij} = J$ we recover the usual ferromagnetic Ising model.

The microcanonical Monte Carlo simulation we will use below was first proposed by Creutz [8] to study the Ising model and has subsequently been applied to a variety of physical systems [9]. To study the equilibrium properties of an Ising model described by (1), a microcanonical ensemble is constructed by letting a demon with energy E_d interact with the spins such that the total energy $E = E_{\text{Ising}} + E_d$ is conserved. In a Monte Carlo simulation step, the energy required to flip a spin, δE, is compared with E_d. If $E_d \geq \delta E$, the flip is permitted and an amount δE is subtracted from E_d. Otherwise the Monte Carlo trial is abandoned. In this algorithm temperature is actually a derived quantity obtained from the fluctuation of the demon energy [8]. This allows large local temperature fluctuations that are essential for models with frustration to bring the system out of its metastable states [6].

The above microcanonical Monte Carlo method is generalized for simulated annealing as the following [6, 7]. The spin system is initialized in the high-temperature phase and allowed to interact with the demon. An annealing schedule is defined as several stages of a Monte Carlo run, say M stages, each consisting of τ Monte Carlo trials per spin. Thus the total number of Monte Carlo trials per spin, which is a measure of the total time of the simulation, is $t = M\tau$. At each stage of the annealing schedule, a maximum value of the demon energy, $E_d^{\max}$, is specified such that if the demon energy $E_d > E_d^{\max}$, E_d is reset to $E_d^{\max}$. In this work we always started with a high $E_d^{\max}$ at the first stage, and reduced it to zero in M stages with equal decrement. This procedure efficiently takes energy out of the spin system via the demon, and brings the system to lower and lower energy states.

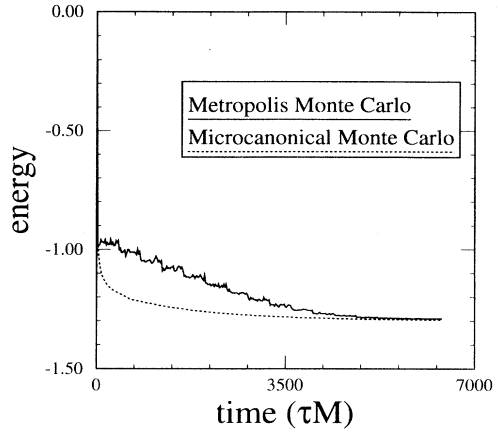

FIG. 1. A typical annealing run of system energy for the 2D Gaussian model as a function of Monte Carlo time, with $M = 16$ and $\tau = 400$. Solid line: Metropolis algorithm; dotted line: Microcanonical algorithm. The unit of energy is δJ, that of time is Monte Carlo trial per spin. System size is 100×100 spins.

In a previous work [6], some comparisons of the performance between the microcanonical simulated annealing and the conventional Metropolis algorithm were presented. Figure 1 shows a comparison of the two algorithms in a typical annealing run using the 2D Gaussian model, for given $M = 16$ and $\tau = 400$ steps. It is clear that at the end of the run, i.e., time $t = M\tau = 6400$ where $T \approx 0$ for the Metropolis simulation and $E_d^{\max} \approx 0$ for the microcanonical simulation, both algorithms produce ground-state energies very close to each other, and in this case quite close to the known value as well. However, it is also clear that the actual ground-state spin configuration is approached more quickly by the microcanonical algorithm as the shown by the figure. Namely, before the final annealing stage is reached, spin configurations generated by the microcanonical algorithm are closer to that of the true ground state than those of the Metropolis simulation. The main purpose of this work is to examine the actual τ and M dependence of the final system energy in the microcanonical algorithm. Obviously we expect that larger values of M and τ will lead to lower final system energy.

III. RESULTS

The short-range Ising spin glass described by (1) has been extensively investigated by many authors [10]. For the Gaussian model, transfer matrix calculations [11] give the ground-state energies: $E_0/\delta J \approx -1.31 \pm 0.01$ for 2D; and $E_0/\delta J \approx -1.70 \pm 0.03$ for 3D; where δJ is the width of the Gaussian distribution for $\{J_{ij}\}$. For the 3D$\pm J$

TABLE I. Ground-state energies for different cooling rates of the two-dimensional (2D) Gaussian model. The unit of energy is δJ, which is the width of the Gaussian distribution for the bond strength $\{J_{ij}\}$. M is the number of annealing stages and τ is the number of Monte Carlo steps per annealing stage. The initial demon energy is $E_d = 8$.

	Energy			
τ	M=4	M=8	M=12	M=14
1	-1.0152	-1.1483	-1.1875	-1.1981
5	-1.2003	-1.2407	-1.2519	-1.2550
10	-1.2295	-1.2550	-1.2631	-1.2655
20	-1.2480	-1.2659	-1.2719	-1.2741
30	-1.2564	-1.2710	-1.2763	-1.2777
40	-1.2611	-1.2759	-1.2793	-1.2803
50	-1.2646	-1.2771	-1.2811	-1.2822
80	-1.2711	-1.2810	-1.2847	-1.2857
100	-1.2741	-1.2823	-1.2862	-1.2873
200	-1.2807	-1.2877	-1.2905	-1.2912
400	-1.2857	-1.2917	-1.2938	-1.2940
600	-1.2885	-1.2935	-1.2958	-1.2958
800	-1.2901	-1.2948	-1.2965	-1.2965
1600	-1.2940	-1.2977	-1.2984	-1.2989
2000	-1.2941	-1.2985	-1.2993	-1.3000

model, the transfer matrix [11] gives $E_0/J \approx -1.76 \pm 0.02$. We take these values as a benchmark with which to compare our simulated annealing calculations.

In 2D, we study square lattices with 100×100 spins. In 3D, cubic lattices of 20^3 spins were used. The initial demon energy is arbitrarily set at $E_d = 8$ in dimensionless units, which was also the $E_d^{\max}$ for the first annealing stage, and the spins are initially disordered. Then we cool down the system by reducing $E_d^{\max}$ to zero in equal decrement in M stages, as described in the last section. Various values of M are used, and usually we set M to 4, 8, 12, or 14. At each of the M stages, a number of τ Monte Carlo steps are run per spin. We used τ ranging from 1 to 2000. Finally, for each set of parameters, at least 500 independent bond configurations are averaged to obtain good statistics.

TABLE II. Ground-state energies for different cooling rates for the three-dimensional (3D) Gaussian and $\pm J$ models. The unit of energy is δJ or J for the Gaussian and $\pm J$ models, respectively. M is the number of annealing stages, and τ is the number of Monte Carlo steps per annealing stage. The initial demon energy is $E_d = 8$.

	Gaussian model		$\pm J$ model	
τ	M=8	M=12	M=8	M=12
1	-1.4944	-1.5393	-1.5884	-1.6449
5	-1.6111	-1.6238	-1.7273	-1.7371
10	-1.6292	-1.6389	-1.7421	-1.7474
20	-1.6443	-1.6503	-1.7508	-1.7538
30	-1.6497	-1.6550	-1.7541	-1.7568
40	-1.6540	-1.6580	-1.7565	-1.7585
50	-1.6567	-1.6608	-1.7583	-1.7601
80	-1.6617	-1.6650	-1.7615	-1.7615
100	-1.6635	-1.6666	-1.7625	-1.7642
200	-1.6688	-1.6726	-1.7666	-1.7670
400	-1.6719	-1.6742	-1.7688	-1.7691
800	-1.6755	-1.6772	-1.7713	-1.7715
1200	-1.6766	-1.6782	-1.7722	-1.7723
1600	-1.6775	-1.6797	-1.7730	-1.7730

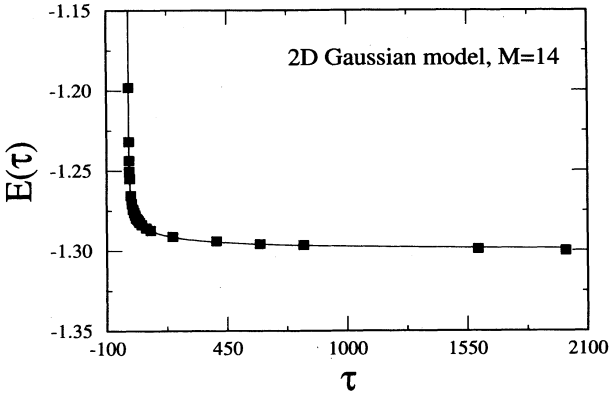

FIG. 2. Final system energy for the 2D Gaussian model as a function of τ, which is the number of Monte Carlo steps per spin at each annealing stage, for $M = 14$. Filled squares are the simulation data and the solid line is the fit to the stretched exponential form, Eq. (2). The unit of energy is δJ.

Similarly, to the results reported in Ref. [3], the final system energy E sensitively depends on the values of τ when τ is small. However, for reasonably large values of τ, E is only weakly dependent on this parameter. This was also observed in our earlier work [6]. Tables I and II summarize the τ dependence of E for the three models studied here. In all cases the lowest energy, which comes from the largest M and τ, compares well with the transfer matrix results quoted above [6]. The exception is the 3D$\pm J$ model, where our result is lower than (though within error bars of) the exact recursive calculation based on small systems [10]. However, this is a known situation which was found in a number of previous calculations [3, 6].

The main results of this work are the τ dependence. Figure 2 shows the final system energy for different τ's, i.e., $E = E(\tau)$ for a value of the annealing stage $M = 14$

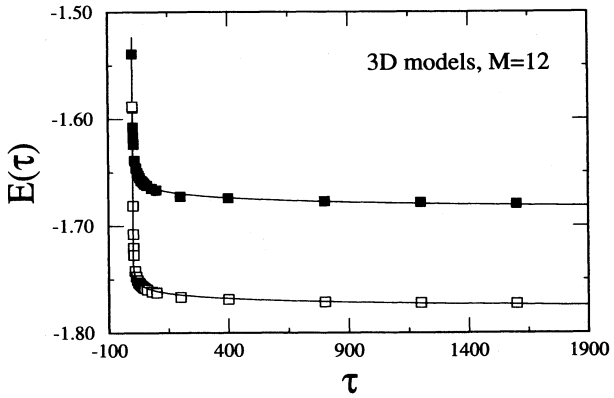

FIG. 3. Final system energy for the 3D models as a function of τ with $M = 12$. (a) Filled squares: Gaussian model; (b) open squares: $\pm J$ model. The solid lines are the fit by Eq. (2). The unit of energy is δJ.

for the 2D Gaussian model (filled squares). It is evident that E varies with τ with a very sharp drop at small values of τ, followed by a very long and slow decaying tail. This is also seen in Table I. We find that the whole data range, for all the M values we have used (Table I), can be well fitted by a stretched exponential form,

$$E(\tau) = E_0 + Ae^{-(\frac{\tau}{\epsilon})^{\gamma}} , \quad (2)$$

where E_0, A, ϵ, and γ are constants. Since our lowest energy is very close to that of the transfer matrix result, we have fixed E_0 to that value in all fits. The solid line of Fig. 2 is the stretched exponential form (2). For all the M's, we found that the time constant $\epsilon \approx 1$, while the stretching exponent $\gamma \approx 0.2$. This form of the energy relaxation is very different from that of Ref. [3], where the conventional Metropolis Monte Carlo method was used in the simulated annealing and a power-law dependence on τ was found. Since a stretched exponential decay is faster than a power-law decay in general, and the lowest energy is indeed very close to that of the known ground-state value, the microcanonical simulated annealing is apparently more efficient than that of the conventional method for the 2D Gaussian model in bringing the spin configuration to that of the true ground state.

It is surprising that our data for the 3D models can also be fitted to the form of (2) for the range of τ that we have used. Figure 3 shows this fitting for the 3D Gaussian (solid squares) and $\pm J$ (open squares) models. Again, the shape of the curves is similar to the 2D Gaussian model case in that there is a sharp drop of the final energy followed by a long slow decaying tail. For the two 3D models, and for different annealing stages M, we found the time constant $\epsilon \approx 1$, and the stretching exponent $\gamma \approx 0.1$. Although the exponent γ is very small, this form of τ dependence is very different from that of the conventional simulated annealing [3] where E was found to decay as $(\ln \tau)^{-1}$. We caution that while the fitting to a stretched exponential form is reasonable up to the largest τ we have tried, the data do possess a very long decaying tail, as seen in Fig. 3. We believe that at even larger τ, the dependence will change from the stretched exponential to something much slower (already signaled by the slow tail). Otherwise the NP-complete problem would be solved in a time, which scales in N equal to or even faster than a power law, thus producing an inconsistency. However, for all practical purposes we obtain final energies quite close to the known ground-state values using the range of τ shown here. Thus we may take the form (2) as an "operational" formula that describes the τ dependence of the present algorithm.

IV. SUMMARY

We have studied the cooling-rate dependence of the approaching ground state for several spin glass models within the microcanonical ensemble simulated annealing scheme. For all models we obtain very good ground-state energies in comparison with the exact recursive calculation on finite systems, and with previous simulations. It is interesting that the final energy dependence on τ, which is the number of Monte Carlo steps at each stage of annealing schedule, can be fitted by a stretched exponential. This is true at least for the range of τ that we have used and that gives reasonable ground-state energies. For the 3D models we believe that the τ dependence will likely change if data for exceedingly large τ are included. But that is beyond our computation capability and is not necessary for practical purposes as very good ground-state energies have already been obtained using smaller τ. Thus for the range of τ we have studied, the stretched exponential effectively describes the cooling-rate dependence of the present simulated annealing algorithm within the microcanonical scheme.

ACKNOWLEDGMENTS

H.G. gratefully acknowledges support by the Natural Sciences and Engineering Research Council of Canada and le Fonds pour la Formation de Chercheurs et l'Aide à la Recherche de la Province de Québec.

[1] See, for example, H.P. Schwefel, *Numerical Optimization of Computer Models* (John Wiley and Sons, New York, 1981); *Combinatorial Optimization*, edited by N. Christofides, A. Mingozzi, P. Toth, and C. Sandi (Wiley-Interscience, New York, 1979).

[2] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, Science, **220**, 671 (1983); S. Kirkpatrick, J. Stat. Phys. **34**, 975 (1984).

[3] G.S. Grest, C.M. Soukoulis, and K. Levin, Phys. Rev. Lett. **56**, 1148 (1986).

[4] B.A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).

[5] Ulrich H.E. Hansmann and Yuko Okamoto (unpublished).

[6] Hong Guo, Martin Zuckermann, R. Harris, and Martin Grant, Phys. Scr. **T38**, 40 (1991).

[7] See also C. Dasgupta, S.K. Ma, and C.K. Hu, Phys. Rev. B **20**, 3837 (1979); N. Sourlas, Europhys. Lett. **6**, 561 (1988); M. Clover (unpublished).

[8] M. Creutz, Phys. Rev. Lett. **50**, 1411 (1983).

[9] G. Bhanot, M. Creutz, and H. Neuberger, Nucl. Phys. **B235**, 417 (1984); M. Creutz, A. Gocksch, M. Ogilvie, and M. Okawa, Phys. Rev. Lett. **53**, 875 (1984); R. Harris, Phys. Lett. **111A**, 299 (1985).

[10] See, for example, K. Binder and A.P. Young, Rev. Mod. Phys. **58**, 801 (1986), and references therein.

[11] I. Morgenstern and K. Binder, Phys. Rev. Lett. **43**, 1615 (1979); Z. Phys. B **39**, 227 (1980); Phys. Rev. B **22**, 288 (1980); J. Appl. Phys. **52**, 1692 (1981); H. Cheung and W.L. McMillan, J. Phys. C **16**, 7027 (1983).