# Cooling-Rate Dependence for the Spin-Glass Ground-State Energy: Implications for Optimization by Simulated Annealing

Gary S. Grest
*Corporate Research Laboratories, Exxon Research and Engineering Company, Annandale, New Jersey 08801*

C. M. Soukoulis
*Ames Laboratory and Department of Physics, Iowa State University, Ames, Iowa 50010*

and

K. Levin
*The James Franck Institute, The University of Chicago, Chicago, Illinois 60637*
(Received 20 September 1985)

The zero-temperature ground-state properties of five spin-glass models have been studied as a function of the cooling rate $r = -\Delta T/t$. Here $\Delta T$ is the temperature decrement and $t$ is the time (in Monte Carlo steps) at each temperature $T$. For the 2D $\pm J$ and Gaussian models, $E(r) = E_0 + C_1 r^x$, where $x \cong 0.25$, while for the 3D $\pm J$, a two-layer $\pm J$, and infinite-range models, $E(r) = E_0 - C_2(\ln r)^{-1}$. We speculate that this difference is related to the fact that the 2D models are not *NP*-complete while the other three models are.

PACS numbers: 64.60.Cn, 05.50.+q, 75.10.Hk, 75.50.Kj

In 1983, Kirkpatrick, Gelatt, and Vecchi[1] introduced the concept of optimization by simulated annealing to search for energy minima of complex systems. They showed that for systems with many constraints frustration is important and simple algorithms based on iterative approaches are often not very effective in finding the lowest-energy state. However, by making an analogy between statistical mechanics and combinatorial optimization, they found[1] that by starting at high temperature $T$ and cooling slowly, significant improvement over standard optimization techniques was obtained for a number of problems. By use of a Metropolis criterion[2] for accepting or rejecting randomly generated trial states, the $N$-city traveling salesman problem,[1] the min-cut partitioning problem,[1] global wiring,[3] and spin-glass ground states[4] have been studied.

While it is clear that the Monte Carlo (MC) method will always find the lowest energy given enough computer time, i.e., a slow enough annealing schedule, it is not known how the final state depends on the actual annealing schedule chosen. Usually one starts at high $T$, where the relaxation times are short, then lowers $T$ by a chosen amount $\Delta T$ ($<0$) and runs for a fixed number of MC steps per spin (MCS) $t$. After a time, $T$ is lowered, $T \rightarrow T + \Delta T$, and one has the choice of either beginning at the new $T$ with the last configuration or with the configuration with the lowest energy[1,5] from the run. In either case, the final energy depends on both the size of $\Delta T$ and the number of MC steps $t$. If the cooling rate $r = -\Delta T/t$ is too rapid, the final state will have a higher energy than for smaller $r$. However, too slow a rate is wasteful in computer resources. Since the annealing schedule has remained the least understoood part of this technique, we have carried out a detailed analysis of the ground-state properties for five model spin-glass systems as a function of the cooling rate. We choose to study spin-glasses because of their interest both from an optimization viewpoint as well as intrinsically.

Not surprisingly, the value of the ground-state energy is a sensitive function of $r$. For two two-dimensional (2D) spin-glass models, the nearest-neighbor $\pm J$ and Gaussian models, $E(r) = E_0 + C_1 r^{-x}$, where $x \cong 0.25$, while for the three-dimensional (3D) $\pm J$ model, a two-layer $\pm J$ model, and the infinite-range[6] model, the dependence is slower, $E(r) \cong E_0 - C_2(\ln r)^{-1}$. Here $C_1$ and $C_2$ are constants. We speculate that this dependence is related to the fact that finding $E_0$ for the latter three problems is an intractable problem,[7] i.e., it is *NP*-complete,[8] which means that the amount of computer time needed to find the true ground state grows faster than any polynomial in $N$, where $N$ is the number of sites. By contrast, the 2D spin-glass models are not *NP*-complete, so that $E_0$ can be found in a time which scales as a polynomial[7] in $N$. This difference in the dependence of $E_0$ on $r$ is not related to the existence or nonexistence of a nonzero transition temperature[6,9-12] $T_c$ since the two-layer model has $T_c = 0$, yet a logarithmic dependence on $r$. We think it is clear that an *NP*-complete problem should not depend on the cooling rate $r$ as a power law. If it did, then one could find a solution in a time which scales as a power law in $N$ and thereby have an inconsistency. However, there may be some non-*NP*-complete problems which do not depend on the cooling rate as a power law. The 1D Gaussian

 

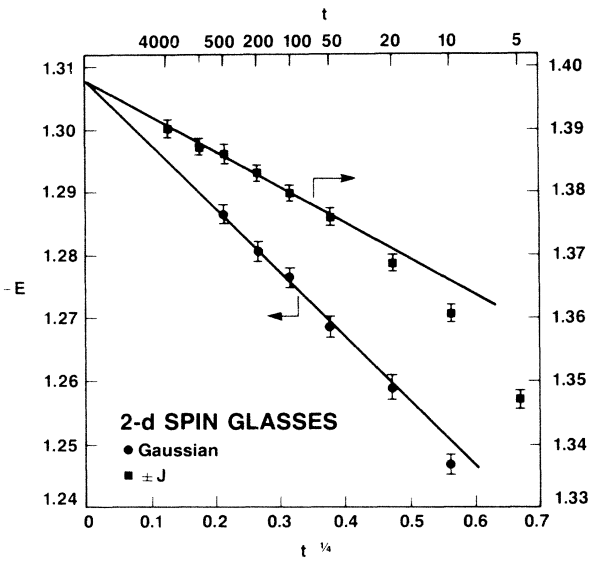


FIG. 1. Ground-state energy $E$ vs $t^{-1/4}$ ($t$ in MCS), for the 2D $\pm J$ (squares) and Gaussian model (circles) with $\Delta T = -0.1$. Solid line through the data is a fit by the form $E(t) = E_0 + C_1 t^{-1/4}$.

spin-glass, studied by Ettelaie and Moore,[13] may be one such example. The only other previous work along these lines is that of Jäckle and Kinzel[14] who found that the residual entropy for the 2D spin-glass model was dependent on the cooling rate, but they did not determine its dependence on $r$.

The Hamiltonian for the short-range spin-glasses is

$$H = -\sum_{ij} J_{ij} S_i S_j, \qquad (1)$$

where the sum is over all nearest-neighbor pairs and $S_i = \pm 1$. The probability distributions $P(J_{ij})$ chosen are the Gaussian distribution of width $J$ and mean $J_0 = 0$ and $\delta(J_{ij} \pm J)$. Because the Gaussian distribution involves considerably more computer time and because there are no reliable, independent estimates of the ground-state energy[12] systems larger than $4^3$, the 3D Gaussian spin-glass model was not studied. We also studied the infinite-range model in which the sum in (1) is over all pairs $ij$ and

$$P(J_{ij}) = \left[\frac{N-1}{2\pi}\right]^{1/2} J^{-1} \exp\left[-\frac{(N-1)J_{ij}^2}{2J^2}\right]. \qquad (2)$$

In all cases we set $J = 1$. In 2D and for the two-layer problem, there is no transition[12] at $T_c > 0$, while the 3D $\pm J$ model on the simple cubic lattice has[10,11] $T_c \cong 1.175$ and the infinite-range model[6-9] has $T_c = 1.0$. In 2D, the simulations were performed on a square lattice of size $100^2$ with periodic boundary conditions. The two-layer model was for a $60 \times 60 \times 2$ lattice, while in 3D we used a $30^3$ simple lattice. The infinite-range model was studied for $N = 200$ and 800. Standard MC techniques were used for the 2D Gaussian and infinite-range models, while a more efficient continuous-time algorithm[15] was used for the three $\pm J$ models. This method is more efficient than standard MC techniques for $T \lesssim 2.0$ in 3D and $T \lesssim 1.5$ in 2D, allowing us to study slower cooling rates. Our results were averaged over $M$ configurations for each cooling rate. The standard deviation in the energy was less than $\pm 0.0010$. In all cases, we started our simulations at high enough $T_i$ where relaxation is known to be rapid. We then lowered $T$ at a fixed rate, usually with $\Delta T = -0.1$. Smaller values of $\Delta T$ showed no significant change in our results. However, much larger decrements do change the results. Similarly, starting at higher $T_i$ did not change the results significantly while lowering $T_i$ produced higher-energy ground states. We started at $T_i = 1.5$ for the 2D models and the infinite-range model, 2.0 for the two-layer model,

TABLE I. Ground-state energies for different cooling rates for four short-range models. Here $\Delta T = -0.1$, and $M$ is the number of configurations $[J_{ij}]$ which were averaged over.

| | $\pm J$ (2D) | | Gaussian (2D) | | Two-layer $\pm J$ | | $\pm J$ (3D) | |
|---|---|---|---|---|---|---|---|---|
| $t$ (MCS) | $M$ | $E$ | $M$ | $E$ | $M$ | $E$ | $M$ | $E$ |
| 5 | 300 | −1.3472 | 150 | −1.2313 | 600 | −1.5511 | 50 | −1.7231 |
| 10 | 300 | −1.3608 | 150 | −1.2468 | 360 | −1.5666 | 50 | −1.7375 |
| 20 | 300 | −1.3689 | 150 | −1.2590 | 240 | −1.5790 | 50 | −1.7478 |
| 50 | 200 | −1.3763 | 90 | −1.2687 | 120 | −1.5903 | 30 | −1.7576 |
| 100 | 120 | −1.3801 | 90 | −1.2766 | 120 | −1.5970 | 21 | −1.7639 |
| 200 | 120 | −1.3833 | 40 | −1.2812 | 40 | −1.6010 | 12 | −1.7676 |
| 500 | 60 | −1.3863 | 16 | −1.2867 | 40 | −1.6053 | 4 | −1.7706 |
| 1000 | 32 | −1.3974 | ... | ... | 28 | −1.6067 | ... | ... |
| ∞ | ... | −1.398 | ... | −1.308 | ∞ | −1.630 | ... | −1.791 |
| Recursive method[a] | ... | $-1.40 \pm 0.01$ | | $-1.31 \pm 0.01$ | | ... | | $-1.76 \pm 0.02$ |

[a]Reference 12.

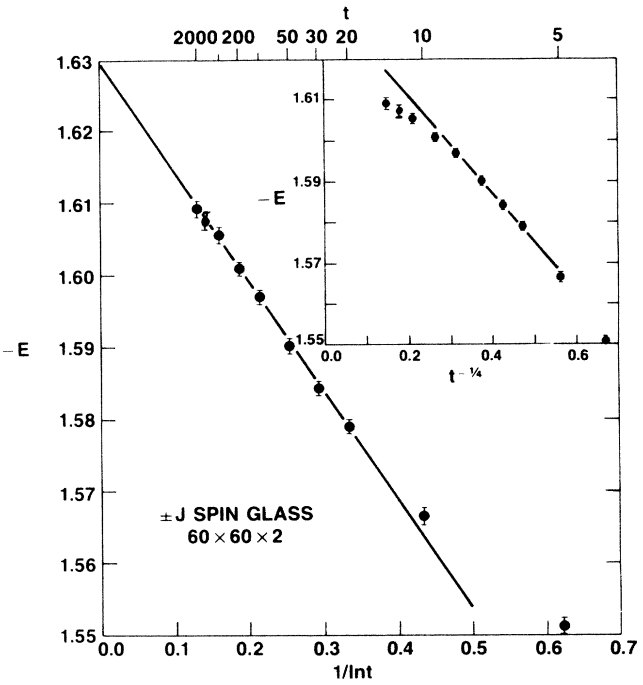


FIG. 2. Ground-state energy $E$ vs $(\ln t)^{-1}$ for the two-layer $\pm J$ model ($60\times60\times2$). Inset: The data plotted vs $t^{-1/4}$.

and 2.2 for the 3D $\pm J$ model. As will be demonstrated, our MC results all extrapolated to the known ground-state energy for infinitely slow cooling rates.

Results for the ground-state energy for the 2D $\pm J$ and Gaussian models are shown in Fig. 1 as a function of the cooling time for $\Delta T = -0.1$. In Table I we present numerical values for energy $E$ and $M$. Though we have plotted the data versus $t^{-1/4}$, a good fit is obtained with $t^{-x}$, $0.20 \lesssim x \lesssim 0.30$. The extrapolated values for $E_0$ given in Table I agree with those obtained from the recursive method.[12]

The dependence on cooling rate is much slower for the other three models. In Figs. 2–4 we show the ground-state energy $E$ vs $(\ln t)^{-1}$ with $\Delta T = -0.1$ for the two-layer $\pm J$, 3D $\pm J$, and infinite-range models, respectively. The values of the data points are presented in Table I for the short-range models.

The inset of Fig. 2 shows the results for the two-layer model plotted versus $t^{-1/4}$, which does not fit the data. The extrapolated value for the $\pm J$ model agrees with that $(-1.78--1.79)$ obtained by Ogielski[16] cooling a $64^3$ lattice very slowly. Note that both our results and those of Ogielski[16] are lower than the exact recursive result, $E_0 = -1.76$ for a $4^3$ lattice.[12] Results for $E_0$ for the infinite-range model agree very well with those determined by use of a slightly modified deterministic procedure.[17] The result[17] for $N = 800$, $E_0 = -0.759 \pm 0.004$ is plotted in Fig. 4. For $N = 200$, $E = -0.742 \pm 0.002$. Both of these results have been rescaled by $N^{1/2}/(N-1)^{1/2}$ to account for

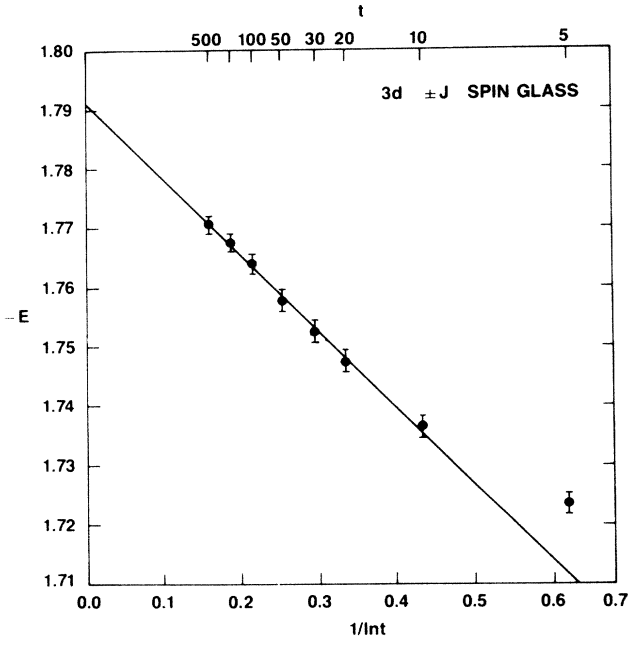


FIG. 3. Ground-state energy $E$ vs $(\ln t)^{-1}$ for the 3D $\pm J$ model with $\Delta T = -0.1$.

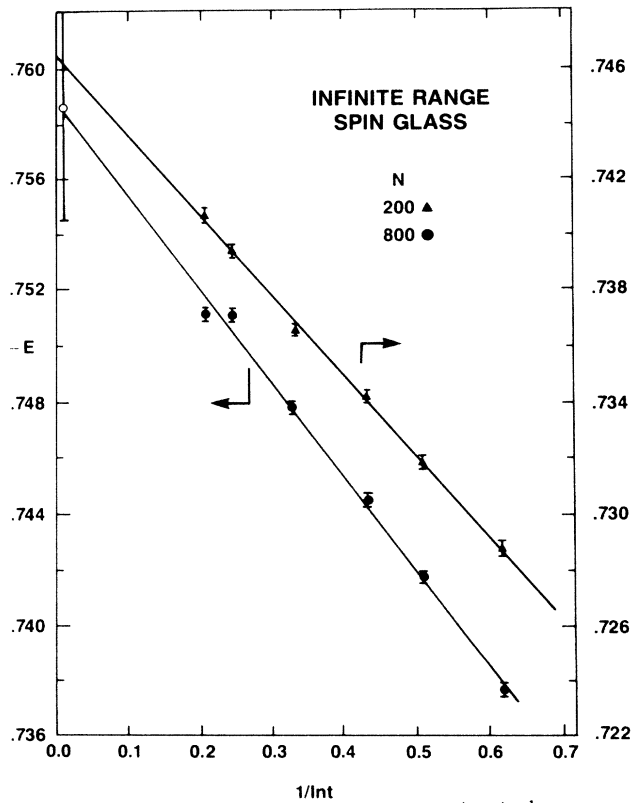


FIG. 4. Ground-state energy $E$ vs $(\ln t)^{-1}$ for the infinite-range model with $\Delta T = -0.1$ for two values of $N$.

the different definition of $P(H)$ in Eq. (2).

In summary, Fig. 1 graphically illustrates the reliability of traditional MC methods in estimating the ground-state energies for 2D systems. Figures 2–4 emphasize the difficulties of obtaining good ground states for *NP*-complete problems. This difference is related to the *NP*-complete problem and not to the existence of a nonzero transition temperature. Our results[18] for the $N$-city traveling salesman problem, which is also *NP*-complete, show that the optimal tour depends logarithmically on the cooling rate.

We thank I. Morgenstern and A. T. Ogielski for several helpful discussions. This work was partially supported by the National Science Foundation (Grant No. DMR 81-15618), the Materials Research Laboratory (Grant No. NSF-DMR 16892), and the U.S. Department of Energy (Grant No. W-7405-Eng-82).

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

[2]N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[3]M. P. Vecchi and S. Kirkpatrick, IEEE Trans. Comput.-Aided Des. **2**, 215 (1983).

[4]C. M. Soukoulis, K. Levin, and G. S. Grest, Phys. Rev. B **28**, 1495 (1983). Here it is shown that a mean-field approach is useful for rapidly generating ground states for the Gaussian model. It has been shown [J. D. Reger, K. Binder, and W. Kinzel, Phys. Rev. B **30**, 4028 (1985)] that it is not useful for the $\pm J$ model.

[5]E. Bonomi and J.-L. Lutton, SIAM Rev. **26**, 551 (1984).

[6]D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

[7]F. Barahona, J. Phys. A **15**, 3241 (1982).

[8]M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).

[9]G. Parisi, Phys. Rev. Lett. **43**, 1754 (1980), and J. Phys. A **13**, 1011, 1887 (1980).

[10]A. T. Ogielski and I. Morgenstern, Phys. Rev. Lett. **54**, 928 (1985); A. T. Ogielski, Phys. Rev. B **32**, 7384 (1985).

[11]R. N. Bhatt and A. P. Young, Phys. Rev. Lett. **54**, 924 (1985).

[12]I. Morgenstern and K. Binder, Phys. Rev. Lett. **43**, 1615 (1979), and Phys. Rev. B **22**, 288 (1980), and Z. Phys. B **39**, 227 (1980).

[13]R. Ettelaie and M. A. Moore [J. Phys. Lett. (to be published)] studied the residual entropy $S_{res}/N$ for the 1D Gaussian spin-glass model as a function of cooling rates. They found that $S_{res}$ decreases to zero only logarithmically slowly with inverse cooling rate. We replotted their data vs both $t^{-x}$ and $(\ln t)^{-1}$ to determine if the data would extrapolate to the known result $S_{res}=0$ for $t\rightarrow\infty$. However, all values of $x$ gave results for $S_{res}$ which were far above 0, suggesting that the asymptotic regime had not been reached. We are presently studying this model in more detail.

[14]J. Jäckle and W. Kinzel, J. Phys. A **16**, L163 (1983).

[15]A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975); G. S. Grest and D.J. Srolovitz, Phys. Rev. B **32**, 3014 (1985).

[16]A. Ogielski, private communication.

[17]S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978); R. G. Palmer and C. M. Pond, J. Phys. F **9**, 1451 (1979).

[18]R. E. Randelman and G. S. Grest, unpublished.