Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 360 (2006) 525–533
PHYSICA A
www.elsevier.com/locate/physa

Hysteretic optimization, faster and simpler

Károly F. Pál

Institute of Nuclear Research of the Hungarian Academy of Sciences, H-4001 Debrecen, P.O. Box 51, Hungary

Received 26 January 2005
Available online 15 June 2005

Abstract

Hysteretic optimization is a recently proposed heuristic optimization method inspired by the demagnetization of magnetic materials by an alternating external field of decreasing amplitude. So far the method has been implemented by changing the field very slowly. Here we show that changing the field very fast is often more effective. We also introduce an external field of the form different from the one applied before, which makes the application of the algorithm as a general-purpose optimizer simpler.

PACS: 02.60.Pn; 07.05.Tp

Keywords: Optimization; Hysteresis; Spin glass; Travelling salesman problem

1. Introduction

Hysteretic optimization [1] is one of the optimization algorithms inspired by a physical process. When a magnetic sample is subjected to an alternating magnetic field of slowly diminishing amplitude, it not only becomes demagnetized, but also ends up in a low-energy state. Finding ground states of some disordered magnetic systems is one of the toughest optimization problems. It has been shown that the simulation of the behaviour of such a sample in the demagnetizing field does lead to a very low-lying state. The result may be further improved by repeatedly shaking up

E-mail address: kfpal@hal.atomki.hu.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2005.05.040

the system by applying the demagnetizing process again and again, but with a smaller maximum amplitude, such that some of the best correlations achieved so far are not destroyed. Instead of a uniform external field, it is better to apply fields whose direction is chosen randomly at each spin site, and to use a new field configuration for each shake-up. Hysteretic optimization is a full demagnetization process followed by shake-ups. The method is very competitive for a spin glass model of high connectivity, but works less well for finite-dimensional cases, although it still outperforms simulated annealing [1], which is a very powerful and perhaps the most widely used general-purpose heuristic optimization method [2,3].

Hysteretic optimization can also be made a general-purpose optimization method. It can be generalized to completely different types of optimization problems by extending the notion of the external field. For magnetic systems the external field pulls the system towards a well-defined state, the one fully aligned with it. In the general case we may define a distance between the members of the configuration space of the problem, and consider a term proportional to the distance of the actual state from a randomly chosen reference state as the external field. Such a term gives the smallest contribution, zero, for the reference state itself, so it does have the effect of pulling the system towards that state. The usefulness of the extended algorithm has been demonstrated on a travelling salesman problem [4].

The demagnetization process usually works best if the field is changed very slowly. However, it turns out that an algorithm with abrupt changes of the external field may still be significantly more effective. In many cases such an algorithm can be made so much faster, that by simply doing more shake-ups we may achieve better results still in less time. This will be demonstrated in the present paper. We will also show that hysteretic optimization does work with applying an external field not based on distance from a reference state, but which has the same mathematical form as the original problem. This makes the application of the algorithm as a general-purpose optimization method conceptually simpler, and often it also makes coding easier. Moreover, we may even obtain a more effective algorithm. In the next section we discuss the algorithm in some more detail. Then we show the results for spin glass systems and for instances of the travelling salesman problem. Finally, we draw conclusions.

2. The algorithm

For disordered magnetic systems the application of hysteretic optimization consists of simulating the behaviour of the system in the appropriately varying magnetic field. The energy of an Ising spin system can be written as

$$\mathscr{H} = -\frac{1}{2}\sum_{i,j}^{N} J_{ij}\sigma_i\sigma_j - H\sum_{i}^{N}\xi_i\sigma_i\,, \tag{1}$$

where the first sum is the expression to be minimized to solve the ground state problem by finding the appropriate $\sigma_i = \pm 1$ spin values. The interactions J_{ij} between

the spins are fixed values defining the particular problem. In the case of the Sherrington–Kirkpatrick model each spin interacts with all the others, and $J_{ij} = z_{ij}/\sqrt{N}$ $(i \neq j)$, where z_{ij} is a random Gaussian number with zero mean and unit variance. In the case of a three-dimensional Edwards–Anderson model the spins are arranged on a cubic lattice, only spins on nearest-neighbour sites interact, and the interaction between them is $J_{ij} = z_{ij}$. The second sum in Eq. (1) is the varying external field with strength H and direction $\xi_i = \pm 1$ at site i. For a homogenous field all $\xi_i = 1$.

In simulating a full demagnetization process we start from the $\sigma_i = \xi_i$ fully aligned state, which is the lowest energy state for large enough $H > 0$. There is an H_S value of the field strength, where one of the spins becomes unstable. If we flip that spin while keeping H fixed, other spins interacting with the spin flipped may become unstable, whose flip may destabilize further spins. Whenever more than one spin is unstable at the same time, we choose randomly which one to flip next, and carry on until all spins are stable and the avalanche stops. Then we decrease H to a value when a further spin becomes unstable, and simulate the resulting avalanche again. In this way, we may follow the evolution of the system from avalanche to avalanche while we vary H. We decrease H until we reach $-\gamma H_0$, where the $\gamma < 1$ amplitude reduction factor is a parameter of the algorithm (throughout this paper we use $\gamma = 0.95$), and the maximum amplitude H_0 is also predetermined. A value of $H_0 \sim H_S$ is appropriate. Then we start increasing the field up to $+\gamma^2 H_0$, then decrease again to $-\gamma^3 H_0$, and so on, until the amplitude becomes so small that no more spin flip occurs. The system takes very different paths when the field increases than when it decreases: it shows a hysteretic behaviour. The reason for this is that after an avalanche the spins taking part in it end up in a stable configuration. If we reverse the direction of the change of the field the avalanche will not occur backwards; usually another avalanche will start somewhere else at a different field strength. During a shake-up we do almost exactly the same as in the case of the full demagnetization process. The only difference is that we start from zero field and the maximum amplitude is $H_{\text{shake}} \ll H_S$, which is another parameter of the algorithm. A detailed description of the algorithm, both for magnetic systems and in its generalized form, may be found in Ref. [5].

Following the evolution of the system from avalanche to avalanche while keeping H fixed during each avalanche corresponds to the limit of changing the field strength infinitely slowly. Although in practice between avalanches we change the field abruptly, it does not make any difference: nothing would occur, however slowly we did it. In this paper we investigate the other extreme, changing the field strength very fast. The full demagnetization process starts with the fully aligned state again, corresponding to the optimum at large H. We drop H to zero, that is, we do a local optimization of the minimization problem defined by Eq. (1) with $H = 0$ by doing favourable spin flips one by one, choosing randomly which flip to do next, until the system is stable. A local optimization like this is called a (simulated) quench. Then we quench the system with fixing H as $-\gamma H_0$. After that we set the field zero again, and then increase it to $H = +\gamma^2 H_0$, and so on.

We can extend the applicability of the algorithm to optimization problems other than finding ground states of magnetic systems the following way [1,4,5]. We identify the function to be minimized (or minus one times the function to be maximized) as the energy of a system. We define a dynamics on the configuration space of the problem, that is, the space of possible solutions by defining some elementary step between its elements. We will also refer to these elements as states. The choice of dynamics should be such that most of the time one step must not change the energy much. Then we add an extra term to the energy corresponding to the external field. The possibility we suggested earlier is to define a distance between the elements of the configuration space, choose an element randomly as the reference state, and take H times the distance from that state as the external field. In the distance definition we have to make sure that it is consistent with the elementary step, such that states one step away should also have a small distance from each other. The way in which the magnetic systems have been treated fits into this scheme. The elementary step is a spin flip, while the external field—except for an irrelevant additive constant—is proportional to the Hamming distance of the actual state from the fully aligned state, which corresponds to the reference state. A problem with the application of the algorithm to a different type of optimization problem is that pushing the solution away from the reference state, that is, applying a negative field strength, often does not have the desired effect [4,5]. In this case it is a good strategy to apply only positive H values, but to choose a new reference state for each half period of the process. The modified algorithm is the following. Start from the first reference state; decrease the field to zero either slowly or fast. The first means following the evolution avalanche by avalanche, while the second constitutes a quench. Then the ith half period is to choose a new reference state, go up to $\gamma^i H_0$ with the field strength H, and then back to zero.

In this paper we also consider a different form for the external field. We simply take H times an expression having the same mathematical form as the original problem as the external field. Therefore, the demagnetizing field is proportional to the objective function of an auxiliary problem, which is basically another instance of the same kind of problem. In the case of the Ising spin system this kind of external field is $-H/2\sum_{i,j}^{N} J'_{ij}\sigma_i\sigma_j$. Choosing a new external field means choosing a new set of J'_{ij}. For magnetic systems this expression is more complicated than the original one. Nevertheless, demagnetization and shake-ups work equally well with both. Unfortunately, the new scheme is somewhat slower, so less effective. For other problems it does have advantages. The application is conceptually much simpler. We need not bother defining a distance function consistent with the elementary step. The form is given, and whenever the dynamics is appropriate to the original problem, it is always consistent with the elementary step. Furthermore, programming may be simpler as the field and the problem to be solved may be treated analogously. We will show that in some cases we also obtain a more powerful method.

3. Application to spin glasses and to the travelling salesman problem

Hysteretic optimization works very well for the Sherrington–Kirkpatrick spin glass model. As far as we know, the largest such systems whose ground states have

been determined and the results have been published [6] contain $N = 1023$ spins. For the algorithm used in that paper, extremal optimization [7], each case took about 20 h of CPU time, and the author considered 190 samples. With hysteretic optimization we can obtain reliable ground states for such sizes in half an hour. Obtaining statistics from ground states of thousands of such systems up to the size of about $N = 1500$ is affordable. With exact methods only much smaller sizes can optimized [8]. Such systems may be handled very reliably with hysteretic optimization in seconds. The algorithm also outperforms hybrid genetic algorithm, which is very good for the Edwards–Anderson case [9,10]. Unfortunately, the modifications considered in the present paper do not lead to an even better algorithm for the Sherrington–Kirkpatrick case. It does not make much difference in the quality of the results if we apply the same number of shake-ups with an external field of the usual form, or if we apply one similar to the internal field. However, the latter takes about a factor of two longer. Changing the field slowly gives better results than changing it fast if the same number of shake-ups is done, and the latter is hardly faster for this problem. The reason for this is the following. To decide whether a spin flip is favourable or not, we have to know the influence of all other spins on that spin, characterized at spin site i by the local internal field $\sum_j^N J_{ij}\sigma_j$. When a spin is flipped, the local field changes everywhere else. Therefore, to update the local fields and the list of favourable moves we need to do operations in the order of the number of spins with each spin flip, and this consumes most of the computer time. The number of spin flips to be done is about the same; either we change the field slowly or fast.

In the case of the Edwards–Anderson model the situation is completely different. Flipping a spin influences only first neighbours, just six spins in three dimensions. Local field only changes at those sites, so the update of that and the list of favourable moves can be done fast. What takes the most time when we change the field slowly is to find out the field strength and the spin position where the next avalanche will start, which has to be done before each avalanche. But there are many of them, because they are typically small. They only involve an average of about 2 spins [5]. We note that in the Sherrington–Kirkpatrick case avalanches tend to be large [11]. When changing the field strength fast, for each quench we only have to create the list of favourable flips once; then, we may update it at each spin flip. This can be done quickly for an Edwards–Anderson spin glass, so the algorithm with abrupt field changes is substantially faster; in our implementation it does about 20 times more shake-ups for an $N = 10^3$ system at the same time, than our code with a slow field change. Moreover, it gives even slightly better results in the same number of shake-ups, so the total gain is quite significant.

A comparison concerning distance from optimum energy per spin averaged for 100 runs versus computation time is shown Fig. 1; a computer with a 2.4 GHz Intel P4 processor was used to measure time. The figure also shows the performance of hysteretic optimization with the auxiliary problem-type external field. In a given number of shake-ups this method works slightly better than the one with the ordinary external field, but each shake-up takes more time. In a given time the algorithm with an abrupt field change and the usual external field performs the best. We note that in the case of the auxiliary problem-type external field we generated a

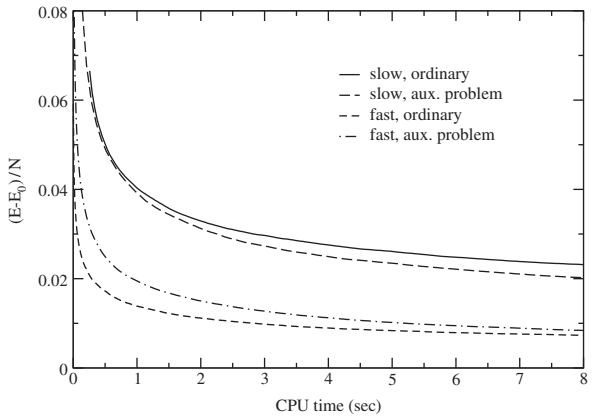

Fig. 1. Distances from optimum energy per spin averaged for 100–100 runs of different versions of hysteretic optimization as functions of computation time (2.4 GHz P4).

new field for each half period, and if the field was changed fast, this required a significant part of all computational needs. To save time, instead of using a Gaussian distribution we actually chose J'_{ij} from an even distribution from the $(-1, 1)$ interval. Otherwise, like the original problem, the auxiliary problem was a three-dimensional Edwards–Anderson spin glass. It has been demonstrated in Ref. [1] that the original hysteretic optimization already outperforms simulated annealing for the present problem. Unfortunately, for the Edwards–Anderson ground state problem, even with the substantial improvement due to the fast change of the field, hysteretic optimization is still not competitive with quite a few other methods [9,10,12–16,7].

The other problem we consider in the present paper is the travelling salesman problem, one of the most popular combinatorial optimization problems. It involves N cities with given distances between them. The aim is to find the shortest round tour that goes through each of the cities once. Earlier [4] we demonstrated the usefulness of the generalized hysteretic optimization with a slowly changing external field on some Euclidean travelling salesman problem (the cities are arranged on a plane and their distance is the ordinary Euclidean distance). We considered different dynamics and also different sorts of distances between the elements of the configuration space to define the external field. Now we only consider the combination which turned out to be the best we tried then. The elementary step defining the dynamics is 2-opt. This operation means visiting cities in some part of the original tour in the opposite order. To do this we have to link the subtour left alone and the one reversed differently, which involves deleting 2 edges of the original tour and establishing 2 new ones. (There is an edge between two towns if the salesman travels from one town to the other, or vice versa). As distance does not depend on which direction we go between two cities, this is the only change that affects the tour length, and this is the change that involves the smallest number of edges. In this sense this is the most basic operation, and it is not surprising that accepting this as the elementary step gives a good dynamics. We note that it would even be better to consider a wider class of

steps as elementary, like 3-opt, involving the rearrangement of up to 3 edges. The distance we consider between two tours is simply the number of edges not common in the two tours.

We consider two problems here. One involves 100 cities scattered randomly on the unit square, with even distribution. The other is a 127-city problem, taken from TSPLIB [17], based on the locations of bier gardens in Augsburg. Comparison of performance with the slowly and fast-varying external field is shown in Figs. 2 and 3 for the two problems. The figures show the proportion of independent runs that have not found the global optimum yet as a function of computation time. In the case of the 100-city problem fast field change works much better. It is not much worse in the same number of shake-ups, but a few times faster. For the bier garden problem the difference in speed is even larger, a factor of 7, but that still does not compensate for the difference in the quality achieved in the same number of shake-ups, so the performance in a given time is similar.

We also considered hysteretic optimization with the external field derived from an auxiliary problem. In this case the demagnetizing field is H times the total length of the tour for the same cities but arranged differently than in the original problem. This actually means that the 'energy' at a given field strength is the total tour length of a travelling salesman problem with distances between any pair of towns given as the distance according to the original problem plus H times that of the auxiliary problem. Between two avalanches or after each quench we are always in a local optimum of such a problem with the appropriate H. It is easy to see that the previous external field may be considered as a special case; it is equivalent to an auxiliary problem that takes city to city distances at zero if they are connected in the reference tour, and one otherwise. Now we consider another Euclidean problem instead, with cities distributed randomly on a square. In this case we always obtain much better results in a given number of shake-ups than in the case in which the external field is

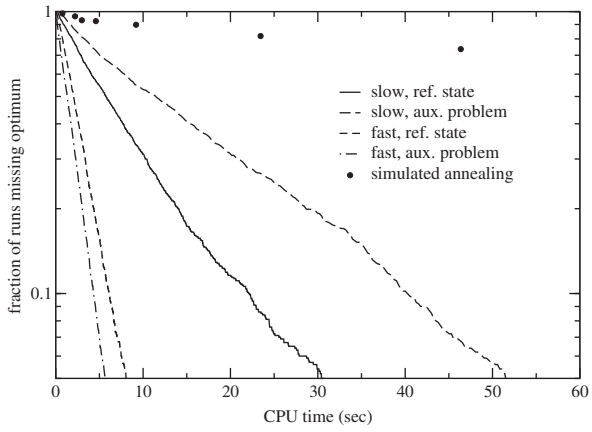

Fig. 2. Proportion of independent runs of different versions of hysteretic optimization that has not found the global optimum yet as a function of computation time for a travelling salesman problem with 100 towns randomly distributed on a square. Results with simulated annealing are also shown.

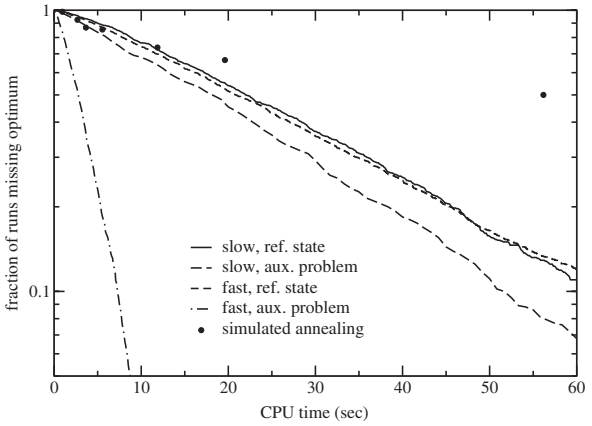

Fig. 3. Proportion of independent runs of different versions of hysteretic optimization that has not found the global optimum yet as a function of computation time for a 127-town travelling salesman problem from Ref. [17]. Results with simulated annealing are also shown.

defined with the distance from a reference state, but unfortunately, a shake-up also takes a few times longer. Nevertheless, the difference is much smaller when the field is switched on and off quickly than when it is varied slowly. For both travelling salesman problem instances we considered, the best strategy is fast field change with fields from auxiliary problems (Figs. 2 and 3). For the bier garden problem this combination is much better than any other one. We also show results from simulated annealing. There we used the same 2-opt operation as the elementary step, we took 4,00,000 (5,00,000) attempted moves or 40,000 (50,000) accepted moves at each temperature for the 100 (127)-city problem, and applied exponential cooling schedules. From each run we accepted the best state reached any time during the run. We can see that we can beat simulated annealing with hysteretic optimization significantly. However, we must admit that these general-purpose algorithms cannot compete with methods specifically designed for the travelling salesman problem [18], especially with the ones that also take advantage of the geometric origin of the problem. Choosing a set wider than 2-opt as an elementary step would certainly help considerably, but we do not believe we can obtain global optima for problems with thousands of cities in an affordable time, which the best methods seem to be able to do.

4. Conclusion

Hysteretic optimization is a recently proposed physics-inspired optimization method. It has been applied for finding low-lying states of disordered magnetic systems and in its generalized form for solving instances of the travelling salesman problem. For a spin glass system of high connectivity the method is very successful. For finite-dimensional spin glasses and for the travelling salesman problem its performance is less impressive, but considering it as a general-purpose optimization

algorithm it works reasonably well, especially with the options introduced in the present paper. In every case we could do significantly better with some form of hysteretic optimization than with simulated annealing, which is probably the most popular general-purpose heuristic optimization algorithm. Simulated annealing is so popular not only because it does quite well on a very wide variety of problems, but because it is extremely simple to implement. The application of the generalized hysteretic optimization required more considerations: to define an appropriate distance measure on the configuration space for the external field is not necessarily obvious for any problem. The idea introduced here to use a field whose form is the same as that of the problem to be solved itself is a step towards making the usage of hysteretic optimization also simple.

Acknowledgements

The author acknowledges the support received from Hungarian grants OTKA T037212 and T037991.

References

[1] G. Zaránd, F. Pázmándi, K.F. Pál, G.T. Zimányi, Phys. Rev. Lett. 89 (2002) 150201.
[2] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi, Nature 220 (1983) 671.
[3] E. Aarts, J. Korst, Simulated Annealing and Boltzmann Machines, Wiley, Chichester, 1989.
[4] K.F. Pál, Physica A 329 (2003) 287.
[5] K.F. Pál, in: A.K. Hartmann, H. Rieger (Eds.), New Optimization Algorithms in Physics, Wiley-VCH, Weinheim, 2004, p. 205.
[6] S. Boettcher, Eur. Phys. J. B. cond-mat/0407130, accepted for publication.
[7] S. Boettcher, A.G. Percus, Phys. Rev. Lett. 86 (2001) 5211.
[8] S. Kobe, cond-mat/0311657.
[9] K.F. Pál, Physica A 223 (1996) 283.
[10] K.F. Pál, Physica A 233 (1996) 60.
[11] F. Pázmándi, G. Zaránd, G.T. Zimányi, Phys. Rev. Lett. 83 (1999) 1034.
[12] A.K. Hartmann, Europhys. Lett. 40 (1997) 429.
[13] A.K. Hartmann, Phys. Rev. E 59 (1999) 84.
[14] J. Houdayer, O.C. Martin, Phys. Rev. Lett. 83 (1999) 1030.
[15] J. Houdayer, O.C. Martin, Phys. Rev. E 64 (2001) 056704.
[16] A.A. Middleton, Phys. Rev. E 69 (2004) 055701 (R).
[17] http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/
[18] D.S. Johnson, L.A. McGeoch, in: G. Gutin, A.P. Punnen (Eds.), The Traveling Salesman Problem and its Variations, Kluwer Academic Publishers, Dortrecht, 2002, p. 369.