ELSEVIER

18 March 1994

Chemical Physics Letters 219 (1994) 343–348

CHEMICAL PHYSICS LETTERS

Quantum annealing: a new method for minimizing multidimensional functions

A.B. Finnila, M.A. Gomez, C. Sebenik, C. Stenson, J.D. Doll

Department of Chemistry, Brown University, Providence, RI 02912, USA

Received 29 November 1993

Abstract

Quantum annealing is a new method for finding extrema of multidimensional functions. Based on an extension of classical, simulated annealing, this approach appears robust with respect to avoiding local minima. Further, unlike some of its predecessors, it does not require an approximation to a wavefunction. We apply the technique to the problem of finding the lowest energy configurations of Lennard-Jones clusters of up to 19 particles (roughly 10^5 local minima). This early success suggests that this method may complement the widely implemented technique of simulated annealing.

1. Introduction

Whether you want to plan an efficient bus route, optimize the positions of integrated circuits on a silicon chip, or predict the lowest energy configuration of a cluster of atoms, the problem reduces to one of finding minima or maxima of specified functions. The significance of the optimization field is reflected in the number of papers written on the subject. Over 300 articles on one method alone, simulated annealing, have been published since 1988.

There appears to be no one "perfect" algorithm that will solve every optimization problem. Instead, a host of complementary methods have evolved, each being well suited to specific tasks [1]. For example, Brent's method or the golden section search can be used if the function to be minimized (or maximized) has only one independent variable even if its derivative is unknown. Multidimensional functions are much more difficult to optimize. If the first derivatives can be calculated, conjugate gradient methods or variable metric methods may be used. Otherwise, the downhill simplex method, direction-set methods, simulated annealing, or genetic annealing may be used. None of these techniques can be universally implemented, and the method of choice varies with the details of the specific problem.

Simulated annealing is a particularly promising minimization technique [1,2]. It has, for example, proved effective in finding the global minimum of multidimensional functions having large numbers of local minima. As with other Monte Carlo based approaches, this method is well suited for implementation on both serial and parallel architectures. The mathematical problem, the optimization of a function, is "solved" by relating the task to a corresponding physical process, thermal annealing. As classical systems are slowly cooled, the configuration space density tends to condense in regions where the potential energy is small. If the cooling process is slow enough, then the system ultimately finds the physical arrangement that minimizes its potential energy. Hence, by assigning the function to be minimized to be the analog of the potential energy and some con-

SSDI 0009-2614(94)00117-9

trol parameter as the analog of temperature, the global minimum of a function can be found by simulating the annealing process as the "temperature" is taken to zero. As in the case of the physical system, the rate of cooling is important in determining whether or not such annealing procedures ultimately find their way to the global minimum or become trapped in various local minima. Although this technique is relatively new, it has proved useful for a wide range of optimization problems [1].

Within the protein folding community, various quantum methods have been developed in which the global minimum is found either by tunneling out of local minima or by removing uninteresting minima through smoothing [3]. While this work was in progress, Amara, Hsu and Straub developed a minimization scheme for multidimensional potentials via an approximate solution to the imaginary time Schrödinger equation. The approximate wave function is comprised of a Hartree product of single wave packets. These packets are allowed to move, tunnel, expand, and contract in search of the global minimum. The quantum mechanics of the system is then relaxed. In the limit of $h \rightarrow 0$, the classical minimum is found [4].

We describe in section 2 a new optimization approach, quantum annealing, that is closely related to its classical counterpart. Unlike the other quantum-mechanical approaches discussed, this method does not require an approximation to the wave function. We illustrate its use with two examples in section 3. The first, a pedagogical exercise is designed to illustrate how quantum annealing avoids local minima. The second is a non-trivial determination of the lowest energy configurations of various homogeneous Lennard-Jones clusters. In section 4 we present our summary and discussion.

2. Formal development

We begin by assuming that our task is the minimization of a specified, many-variable function. As in conventional annealing methods, it is convenient to view this function as being the potential energy of a hypothetical physical system. In contrast to the simulated annealing approach, however, it proves convenient to view our system as quantum-mechanical rather than classical in nature.

Assuming that we can compute the average energy of our system as a function of its temperature and quantum-mechanical character, we can approach our objective, the minimum of the system's potential energy (point A in Fig. 1), in a variety of ways. Starting at the arbitrary point D in Fig. 1, simulated annealing first turns off the system's quantum mechanics and then reduces the temperature of the resulting classical system to zero (path DBA in Fig. 1). It proves useful, however, to invert the order of these two limits first reducing the system's temperature to zero and then following the resulting quantum ground state energy to its classical limit (path DCA in Fig. 1). Since the fictitious mechanical system involved is an artificial construct, we are free to craft its character to suit our purposes. In particular, we can control the degree of its quantum-mechanical behavior by varying the masses of its constituent particles.

The motivation for inverting the usual order of the limits is that, in a sense, it is easier to take the zero temperature limit of a quantum problem than a clas-

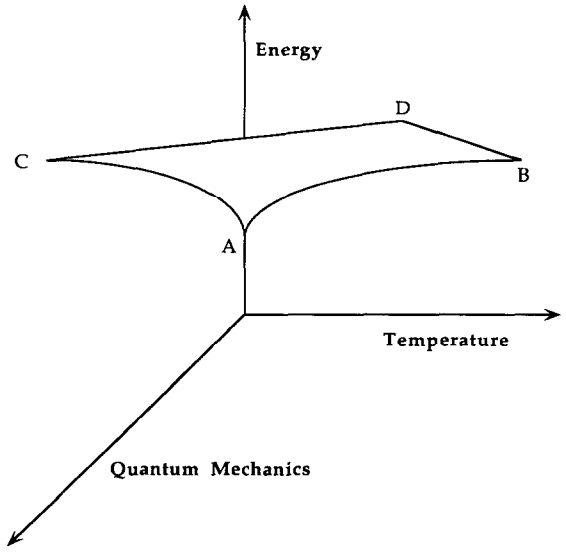

Fig. 1. Shown is an energy–temperature–degree of quantum-mechanics diagram showing the relationship between simulated annealing methods and our quantum annealing process. Simulated annealing is a fully classical method which finds a function's global minimum by slowly cooling temperature (path BA). Quantum annealing is a complementary zero temperature method which finds that same minimum by slowly increasing the mass of the system until the energy approaches its classical limit (path CA).

sical one. In particular, a variety of diffusion based, Monte Carlo methods are available that can be used to compute the ground state energy of general quantum-mechanical systems. Described in detail elsewhere [5], these methods are based on the observation that the Schrödinger equation written in imaginary time is isomorphic to the diffusion equation with a growth/depletion term. In one-dimension the result is

$$\frac{\partial\psi}{\partial\tau} = \frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - [V(x) - E_0]\psi , \qquad (1)$$

where $\tau = it/\hbar$, E_0 is a constant subtracted out for convenience and $V(x)$ is the potential energy.

The diffusion Monte Carlo (DMC) method is one, relatively simple technique for treating such problems. It follows the evolution of a number of random walkers designed to move in such a way to simulate the diffusion, growth and decay processes in Eq. (1). Walkers in regions where $E_0 < V(x)$ are attenuated via first-order decay, while those in regions where $E_0 > V(x)$ undergo analogous growth. In typical applications, E_0 is adjusted iteratively to maintain a steady state population. In terms of the Schrödinger eigenfunctions and energies, $\{\varphi_n\}$ and $\{E_n\}$, respectively, the solution we seek is of the form

$$\psi(x,\tau) = \sum_n \varphi_n(x) \exp[-(E_n - E_0)\tau] . \qquad (2)$$

The ground state wavefunction can thus be identified as the large τ limit of $\psi(x,\tau)$ while the ground state energy is equal to E_0. One advantage of using DMC is that no knowledge of the wavefunction is required. Instead, it can be obtained from the final distribution of walkers in the DMC simulation.

From Eq. (2), it is apparent that the rate of convergence of the DMC method is controlled by the gap between the ground and first excited state energies (ΔE). In fact, the time required for the DMC method to converge to the ground state is $\hbar/\Delta E$. This is the same time scale needed for tunneling between two interacting potential minima. This suggests that DMC finds the ground state wavefunction through tunneling. By gradually increasing the mass of the walkers thereby constraining the wave function, we can follow the ground state energy to its classical limit.

3. Numerical examples

To illustrate quantum annealing, we first consider the problem of finding the minimum of the one-dimensional function shown in Figs. 2 and 3. In Fig. 2, we observe the evolution of transient distributions of DMC diffusers at a fixed mass to demonstrate how the method utilizes tunneling to find the ground state wave function. Initially, all the diffusers are placed in the left well at a ground state energy equal to approximately half the height of the intervening barrier (line (a)). Early in the simulation, some diffusers tunnel through the barrier to the deeper, right well (line (b)). Later, the population of diffusers in the deeper well grows relative to the population in the left well (line (c)) until the energy converges to the ground state and an equilibrium distribution is reached (line (d)). When the ground state is large compared to the difference between the two wells, the distribution is almost symmetric. However, since the right well is marginally deeper, the right population will be correspondingly larger (assuming the number of diffusers is high enough to resolve the slight difference).

In Fig. 3, we show the evolution of the equilibrium

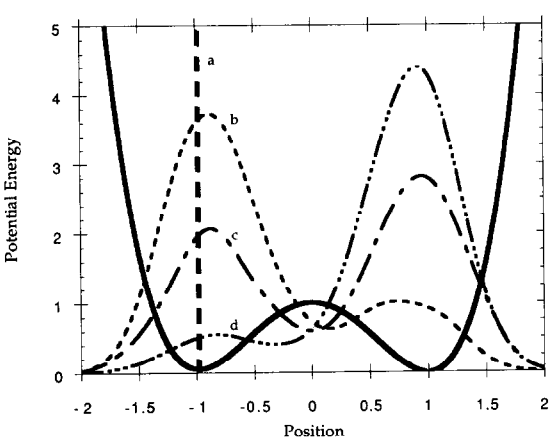

Fig. 2. Diffusion Monte Carlo is used in a one-dimensional double well to illustrate how the method uses tunneling to find the ground state. Lines (a)–(d) show the transient distributions as the diffusers tunnel to the right well and approach an equilibrium distribution. The potential used is of the form $V(x) = 3\delta x^4 + 4\delta(\alpha-1)x^3 - 6\delta\alpha x^2 + 1$ where δ equals $1/(2\alpha+1)$ and α equals -0.98 [6]. The bottoms of the left and right wells are at energies of 0.052 and 0.0 K, respectively, and the height of the barrier is 1.0 K. The mass of the diffusers is 2000 times the proton mass. The ground state energy of the system is 0.388 ± 0.002 K.

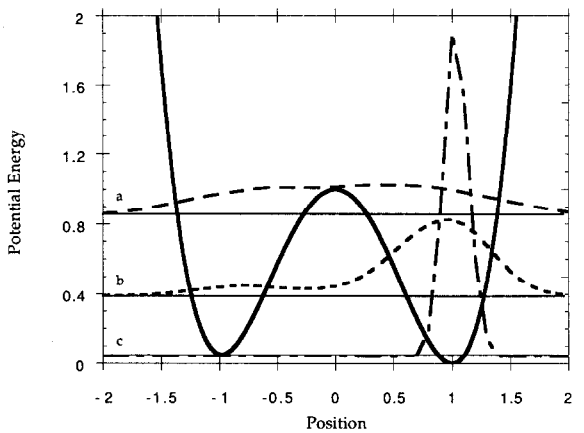

Fig. 3. Quantum annealing is used to isolate the global minimum of the same one-dimensional double well shown in Fig. 2. Some sequential, equilibrium distributions are shown (lines (a)–(c)). The masses of the diffusers are 200, 2000, and 2000000 times the proton mass. The ground state energies are 0.854, 0.388, and 0.043 K, respectively. The ground state energy for each mass is shown by thin solid lines with the distribution of diffusers shown above them.

population of our model system as we turn off its quantum-mechanical character (lines (a)–(c)). For a small mass, quantum effects are large and the ground state energy is relatively high compared with the slight energy difference between the two wells. In this case, the populations in both wells are essentially equal (line (a)). As we increase the mass, the quantum character is reduced, the associated ground state energy decreases and the relative probability of being near the global minimum increases (line (b)). As we make the mass even larger, the ground state energy drops below the energy of the metastable minimum and the probability of being found anywhere other than near the global minimum quickly dwindles (line (c)). In actual practice, it is generally not necessary to completely converge to the ground state probability distribution before increasing the system mass. Once even a single DMC diffuser reaches a particular potential well, it has the ability to proliferate and, hence, sample that region. As with conventional simulated annealing, however, it is necessary to experiment with annealing rates to be sure that results converge to the global minimum independent of initial configuration chosen.

In our second example, we consider the problem of determining the structure of N-atom homogeneous Lennard-Jones clusters. The particles are assumed to interact pairwise via a two parameter interaction of the form $\epsilon V(\xi)$, where ξ is a dimensionless interparticle separation distance, r/σ. The parameters ϵ and σ are the usual Lennard-Jones energy and length scale variables. Explicitly, the Lennard-Jones potential is,

$$V(\xi)=4(\xi^{-12}-\xi^{-6})\,. \tag{3}$$

In terms of our reduced parameters, the quantum-mechanical Hamiltonian for our system is

$$\frac{1}{\epsilon}H=-\tfrac{1}{2}\eta^2\sum_{n=1}^{N}\nabla_\xi^2+\sum_{m<n=1}^{N}V(\xi_{m,n})\,, \tag{4}$$

where η is a dimensionless constant that controls the scale of the quantum-mechanical effects,

$$\eta=\frac{\hbar}{\sigma(m\epsilon)^{1/2}}\,. \tag{5}$$

For rare gases, η ranges in magnitude from 0.426 for ^{4}He to 0.00995 for Xe. As $\eta\to 0$, the ground state energy approaches its classical limit.

Fig. 4 shows that the calculated ground state energies for Lennard-Jones clusters are relatively simple functions of the parameter η. It is not difficult to relate the limiting slopes and curvatures of the ground state energies in Fig. 4 to corresponding zero-point energies and anharmonic effects, respectively. The simplicity of the ground state energies is in marked

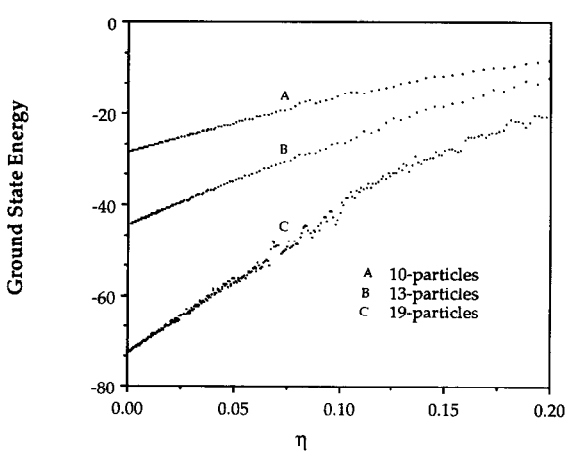

Fig. 4. The energies of clusters of sizes 10, 13, and 19 atoms are shown as a function of the dimensionless parameter η, which shows the degree of quantum effects. As η approaches zero, the ground state energy approaches the potential minimum for the cluster.

contrast to the excited states where the η-dependence is relatively complex even for a 3 atom cluster [7]. We note in particular that the ground state energy of the clusters varies smoothly with η through regions where the behavior of the clusters is changing from delocalized to localized in character. Since η appears in Eq. (4) as the coefficient of the highest order derivative, it is a "singular perturbation", an indication that power series expansions in the parameter η may have a limited (possibly zero) radius of convergence.

The results of the quantum annealing calculations for various cluster sizes are shown in Table 1. The energies quoted were obtained by extrapolating a linear fit to the ground state energies for small (<0.01) values of η. The extrapolated potential minima found are in good agreement with the known minimum energy structures of the classical systems [8]. We note that for all cases reported in Table 1, the agreement becomes exact if we refine the geometries of the large mass DMC clusters using traditional methods, a relatively simple task once the quantum annealing methods have successfully located the vicinity of the global minimum. That we obtain the correct structures is non-trivial since the 13 atom clusters has 988 known minima, and the 19 atom clusters has on the order of 10^5 stable isomers [9]. We note that the minimum energies found using the above extrapolation method are upper bounds to the actual minima since we are using linear fits to approximate a curve that is concave downward (see Fig. 4). Decreasing the interval used in the fits (η_{max}) leads to a better extrapolation of the curve and hence a better estimate of the potential minima. For example, in the seven particle cluster, as η_{max} is decreased from 0.05 to 0.02 to 0.01, the estimate for the potential minimum goes from -16.469 to -16.495 to -16.505, respectively, the final value being that reported by Hoare and Pal [8].

Table 1
Minimum potential energy

Number of particles	Quantum annealing	Ref. [8]
2	−0.998	−1.000
3	−2.999	−3.000
4	−5.997	−6.000
5	−9.095	−9.104
6	−12.710	−12.712
7	−16.505	−16.505
8	−19.794	−19.822
9	−24.108	−24.113
10	−28.408	−28.420
13	−44.305	−44.327
19	−72.622	−72.659

The minimum energies are shown in units of ϵ for various Lennard-Jones clusters obtained via quantum annealing and from Hoare and Pal (see ref. [8]). When conjugate gradient methods are used to refine the minimum, the Hoare and Pal results are reproduced.

4. Discussion

Any method of locating the global minimum must address the issue of local minima. Simulated annealing confronts this problem through the device of classical "thermal fluctuations". Quantum annealing and other methods [3,4] use delocalization and tunneling to avoid metastable regions. By utilizing a quantum rather than a classical system, the present approach exploits a number of specialized ground state methods that are not available within classical problems. Quantum annealing has the further advantage of making knowledge of the wavefunction unnecessary. The physically different ways in which quantum and simulated annealing avoid local minima suggests that these type of approaches may complement each other in general optimization applications.

Acknowledgement

This material is based upon work supported under two National Science Foundation Graduate Research Fellowships, National Science Foundation grant CHE-9203498, and the Petroleum Research Fund of the American Chemical Society.

References

[1] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, Numerical Recipes in FORTRAN, 2nd Ed. (Cambridge Univ. Press, Cambridge, 1992).

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

[3] L. Piela, J. Kostrowicki and H. Scheraga, J. Phys. Chem. 93 (1989) 3339;
R.J. Somorjai, J. Phys. Chem. 95 (1991) 4141;
K.A. Olszewski, L. Piela and H.A. Scheraga, J. Phys. Chem. 96 (1992) 4672.
[4] P. Amara, D. Hsu and J.E. Straub, J. Phys. Chem. 97 (1993) 6715.
[5] D. Ceperley and B. Alder, Science 231 (1986) 555.
[6] D.D. Frantz, D.L. Freeman and J.D. Doll, J. Chem. Phys. 93 (1990) 2769.
[7] D.L. Leiter, J.D. Doll and R.M. Whitnell, J. Chem. Phys. 96 (1992) 9239.
[8] M.R. Hoare and P. Pal, Advan. Phys. 20 (1971) 161.
[9] F.H. Stillinger and D. Stillinger, J. Chem. Phys. 93 (1990) 6013.